

Evandro Lopes de Oliveira

Previsão de demanda de energia elétrica
com redes neurais artificiais e análise por
série de Fourier

Belo Horizonte - Minas Gerais
2016

Evandro Lopes de Oliveira

Previsão de demanda de energia elétrica
com redes neurais artificiais e análise por
série de Fourier

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica - Universidade Federal de Minas Gerais, para a obtenção de Título de Mestre em Engenharia Elétrica, na Área de Sistemas de Computação e Telecomunicações.

orientador: Cristiano Leite de Castro
co-orientador: Frederico Gualberto Ferreira Coelho

Belo Horizonte - Minas Gerais
2016

de Oliveira, Evandro L.

Previsão de demanda de energia elétrica com redes neurais artificiais e análise por série de Fourier

106 páginas

Dissertação de Mestrado - Universidade Federal de Minas Gerais. Departamento de Engenharia Elétrica.

1. Redes neurais
2. Série de Fourier
3. Demanda de energia
4. Modelo de previsão

I. Universidade Federal de Minas Gerais. Programa de Pós-Graduação em Engenharia Elétrica. Departamento de Engenharia Elétrica.

Comissão Julgadora:

Prof. Dr. Rodney Rezende Saldanha

Prof. Dr. Hilton de Oliveira Mota

Prof. Dr. Cristiano Leite de Castro

Dr. Frederico Gualberto Ferreira Coelho

Aos meus pais que, mesmo sendo humildes residentes do interior do Brasil com pouco recurso financeiro, conseguiram investir na educação dos filhos. Isso por acreditarem que esse era o maior tesouro que poderiam dar a eles.

We are the change that we seek.

Barack Obama

Agradecimentos

Agradeço a Deus por ter me guiado nos caminhos do meio acadêmico. A Ele agradeço a oportunidade e os recursos para que pudesse ser feito um trabalho interessante, gratificante e que a tempos era planejado.

Ao Professor Dr. Cristiano Leite de Castro, pela oportunidade concedida na orientação e dedicação durante o desenvolvimento deste trabalho, sempre apresentando observações que agregaram valor acadêmico/profissional e pessoal.

Ao Co-orientador Dr. Frederico Gualberto Ferreira Coelho que teve fundamental participação no esclarecimento de técnicas empregadas nesse trabalho e contribuiu para que o mesmo se realizasse.

Ao Professor Dr. Rodney Rezende Saldanha, assim como todos os outros que ajudaram na minha evolução acadêmica.

À minha família pela paciência em vários momentos que estive ausente por estar dedicado ao estudo.

À assistência dada pelos profissionais dos laboratórios e salas dos departamentos da UFMG.

Resumo

A estrutura atual do mercado de energia elétrica no Brasil é decorrente da quebra do monopólio estatal no setor de energia. Essa estrutura possui regras que separam o setor de energia em segmentos que oferecem o serviço, criando um cenário de competição.

A tendência no mercado competitivo de energia elétrica é que as empresas busquem formas de prever a necessidade de seus clientes, melhorando a qualidade do serviço prestado. Para isso, as empresas mapeiam o comportamento do mercado, verificam os pontos críticos e montam estratégias de gerenciamento e intervenção nesses pontos.

Considerando que o planejamento da demanda de energia elétrica do Brasil é estratégico, o presente trabalho propõe metodologias de previsão dessa demanda, utilizando indicadores de desempenho sócio-econômicos, indicadores climáticos e históricos de demanda de energia elétrica. Os modelos propostos são baseados em redes neurais artificiais, algoritmos estatísticos, algoritmos de identificação de tendências e análise de variáveis exógenas.

A aplicação dos métodos propostos nesse trabalho possibilita a obtenção do comportamento futuro da demanda do Sistema Interligado Nacional (SIN) com uma média de acerto muito próxima da sua série histórica.

Palavras-chave: redes neurais, série de Fourier, demanda de energia, modelo de previsão

Abstract

The current model of the electricity market in Brazil is due to the breaking up of the state monopoly on the electrical energy sector. This model has regulations that divide the energy sector in sections that offer the service, creating a competitive environment.

The trend in the market that competes for electricity is that companies seek ways to provide the customers' need. In order to improve the quality of service provided, companies analyze market behavior, check the critical points and elaborate control and intervention strategies in these points.

Considering that the electricity demand planning in Brazil is strategic, this work shows some demand forecasting methodologies using socio-economic performance indicators, weather indicators and electricity demand history. The proposed models are based on artificial neural networks, statistical algorithms, trend identification and analysis of exogenous variables.

The application of the forecasting methods, demonstrated in this work, enables the achievement of a future behavior of the National Interconnected System's electrical demand with an average close to its historical serie.

Keywords: neural networks, Fourier series, energy demand, forecasting model

Sumário

1	Introdução	15
1.1	Justificativa	17
1.2	Objetivos do trabalho	17
1.2.1	Objetivos específicos	17
1.3	Estrutura da dissertação	18
1.4	Literatura relacionada	19
1.4.1	Publicações	21
2	Previsão de séries temporais	23
2.1	Séries temporais	23
2.1.1	Componentes não observáveis de uma série histórica	24
2.1.2	Modelos auto-regressivos	25
2.2	Variáveis de entrada	26
2.2.1	Autocorrelações	26
2.2.2	Filtro de Hodrick-Presscott	29
2.3	Extrapolação de tendências	30
3	Redes Neurais Artificiais	33
3.1	Características e aplicações	34
3.2	Modelos	37

3.2.1	Classificação	37
3.2.2	Perceptron	38
3.2.3	Perceptron de múltiplas camadas	39
3.3	Treinamento multi objetivo	42
4	Método de previsão por série de Fourier	45
4.1	Funções periódicas	46
4.2	Séries de Fourier	47
5	Seleção de variáveis	51
5.1	Definição das variáveis	51
5.2	Sistema repositório de dados	54
5.3	Relief	56
5.3.1	Relief-F	58
5.3.2	RRelief-F	60
5.4	Métodos de seleção	61
5.4.1	Método Forward	61
5.4.2	Método Backward	64
6	Realização da previsão	67
6.1	Resultados para o modelo RNA	70
6.1.1	Modelo RNA com seleção de variáveis	74
6.1.2	Seleção de variáveis com Mínimos Quadrados	84
6.2	Resultados para o modelo com série de Fourier	89
6.3	Análise de resultados	90
6.3.1	Análise por série de Fourier	91
6.3.2	Análise do modelo RNA	92

7 Conclusão	97
7.1 Trabalhos Futuros	100
Referências Bibliográficas	101

Lista de Figuras

2.1	Processo Estacionário e Não Estacionário em séries temporais (Davila, 2015)	26
2.2	Gráfico de autocorrelação	27
2.3	Decomposição de série temporal	30
2.4	Tendência para extrapolação Linear	31
3.1	Representação simplificada de um neurônio (Ferneda, 2005)	34
3.2	O Neurônio Genérico em RNAs (Castro and Zuben, 2015)	36
3.3	Representação de uma MLP (Soares and Silva, 2011)	40
3.4	Arquitetura da MLP para única saída	40
3.5	Conjunto Pareto (Teixeira, 2001)	43
4.1	Exemplo de uma função periódica (Santos, 2004)	46
4.2	Esquema de previsão com parâmetros atualizados	50
4.3	Esquema de previsão com parâmetros fixos	50
5.1	Aplicativo de gerenciamento de dados	55
5.2	Fluxo de execução do algoritmo Relief	58
5.3	Fluxograma do método forward	63
5.4	Fluxograma do método backward	65
6.1	Demanda de energia elétrica	68

6.2	Componente cíclico da demanda de energia elétrica	69
6.3	Previsão pela modelagem naïve	70
6.4	Gráfico de autocorrelação	71
6.5	Divisão de dados com atrasos mais correlacionados com a série de demanda	72
6.6	Divisão de dados dos 3 primeiros atrasos da série de demanda	73
6.7	Divisão de dados correlacionados e os 3 primeiros atrasos da série de demanda	74
6.8	Curva com as soluções geradas para os dados correlacionados	75
6.9	Curva com as soluções geradas para os 3 primeiros atrasos da série de demanda	75
6.10	Curva com as soluções geradas para dados correlacionados e os 3 primei- ros atrasos da série de demanda	76
6.11	Solução gerada pela MLP para os dados mais correlacionados comparada com a série original	76
6.12	Solução gerada pela MLP com os 3 primeiros atrasos comparada com a série original	77
6.13	Solução gerada pela MLP com os 3 primeiros atrasos e os atrasos com maior correlação comparada com a série original	77
6.14	Ranking de variáveis independentes	80
6.15	Rede Neural Híbrida Forward	81
6.16	Gráfico de aproximação da RNA por seleção de variável com Forward) .	82
6.17	Rede Neural Híbrida Backward	82
6.18	Gráfico de aproximação da RNA por seleção de variável com Backward)	83
6.19	Fluxograma Forward MMQ	85
6.20	Gráfico de aproximação da RNA por seleção de variável com Forward MMQ)	86
6.21	Fluxograma Backward MMQ	87

6.22 Gráfico de aproximação da RNA por seleção de variável com Backward	
MMQ	88
6.23 Gráfico de aproximação por série de Fourier (série estacionária)	89
6.24 Gráfico de aproximação por série de Fourier (série não estacionária)	90

Lista de Tabelas

6.1	Resultados para RNA	78
6.2	Resultados RNA seleção de variáveis com Forward	81
6.3	Resultados RNA seleção de variáveis com Backward	83
6.4	Resultados RNA seleção de variáveis com Forward MMQ	86
6.5	Resultados RNA seleção de variáveis com Backward MMQ	88
6.6	Resultados para série de Fourier	90
6.7	Comparação de resultados	96

Capítulo 1

Introdução

O início da geração de energia elétrica no Brasil se deu com a instalação da primeira iluminação permanente na estação Central do Brasil em 1879 (CEMIG , 2015). Com o passar dos anos, o uso de tecnologias movidas a eletricidade evoluiu e outras fontes de geração de energia foram surgindo para atender à demanda.

Grande parte da energia elétrica brasileira é gerada por hidrelétricas, sendo essas as responsáveis por mais de 70% da produção nacional (CEMIG , 2015). Como na matriz elétrica do Brasil essa é a fonte de geração predominante, há uma grande dependência de fatores climáticos.

O Sistema Interligado Nacional (SIN) precisa manter o nível dos reservatórios das hidrelétricas altos para atender à demanda, pois o parque de geração é iminentemente hídrico. Mas como os períodos chuvosos tem sido curtos e com baixa precipitação, se faz necessário despachar termoeletricas e outras fontes alternativas.

No contexto econômico do mercado, a eletricidade passou a ser negociada como produto, e o que antes era controlado pelo governo brasileiro, separou-se em setores que hoje disputam o provimento do serviço de geração, transmissão, distribuição e comercialização de energia.

As empresas provedoras dos serviços relacionados a energia elétrica focam na efi-

ciência produtiva. Essas empresas se organizam para minimizar os custos e perdas de produção, com objetivo de aumentar as margens de lucro. As empresas investem em tecnologia que melhoram a qualidade do serviço prestado, procurando manter a competitividade. Para que o controle seja garantido, essa competição entre empresas passa a ser conduzida em um ambiente regulado.

O setor energético brasileiro exige que as empresas distribuidoras contratem geração de forma a atender a 100% da demanda dos consumidores, com margem de erro dentro da faixa de 0% a 3% (Azevedo et al., 2005). Então, conforme o decreto no. 5.163/04, as empresas devem ter a capacidade de produção garantida, com novos empreendimentos. Considera-se, na negociação desses novos empreendimentos, o prazo médio de construção de usinas hidrelétrica de grande porte de 5 anos e de usinas termelétricas de 3 anos (Ferreira, 2006). Esse é o tempo necessário para viabilizar a operação dessas usinas.

A estimativa necessária para o planejamento das concessionárias, com qualidade de resultados, pode ser feita com técnicas estatísticas e de inteligência computacional, previsão da demanda de energia elétrica utilizando análise por série de Fourier e redes neurais artificiais (RNA), que possuem metodologias próprias para resolução dos problemas.

A série de Fourier é a representação de uma função como séries infinitas de senos e cossenos. A análise pela série de Fourier é aplicada neste trabalho utilizando o histórico da demanda de energia elétrica como entrada do modelo. Os resultados obtidos com os experimentos são comparados à previsão feita pelo modelo RNA.

As RNAs vêm sendo aplicadas extensivamente à previsão de séries temporais. Essas redes são desenvolvidas com base na modelagem de sistemas de neurônios biológicos. O modelo RNA, aplicado ao trabalho proposto, obteve resultados baseados na entrada de dados históricos da demanda de energia elétrica e algumas de suas variáveis de influência.

São vários os fatores que influenciam na demanda de energia, e identificar todas essas

variáveis é um desafio. O impacto de algumas delas na demanda pôde ser analisado nesse trabalho.

Ao identificar as causas de eventuais variações na demanda de energia elétrica, os seus pontos críticos são expostos. Os resultados alcançados, considerando esses pontos críticos, obtiveram uma média de acerto de 96,78% na previsão comparada aos dados do SIN.

1.1 Justificativa

Em 2001, o Brasil enfrentou a falta de energia causada pela escassez de chuvas. Mesmo com aumento de 1.835 MW na potência instalada do parque hidrelétrico, a oferta de energia hidráulica reduziu-se em 1,9% devido às condições hidrológicas observadas no ano seguinte. Neste mesmo ano, também se teve um aumento de 3,8% do consumo final de eletricidade. São esses dados que influenciam na estimativa da demanda de energia elétrica das concessionárias que provêm os serviços.

O planejamento deficiente no setor energético pode ter um grande peso na inflação. As concessionárias de energia elétrica podem resolver esse problema econômico com estimativas de demanda eficientes, mantendo a disputa de espaço no mercado.

1.2 Objetivos do trabalho

O objetivo geral do trabalho é fazer a previsão da demanda de energia elétrica do Brasil construindo modelos de previsão com RNA e série de Fourier, baseado em um repositório com dados de influência sobre o comportamento da série temporal.

1.2.1 Objetivos específicos

As atividades e os experimentos realizados no trabalho estão associados aos seguintes objetivos específicos:

- Realizar estudo de características e variáveis que possam ter influência na demanda de energia do Brasil.
- Construir um sistema de leitura da base de dados contendo as informações de influência na demanda de energia elétrica do Brasil.
- Gerar uma aproximação da série de demanda de energia elétrica com um modelo baseado na análise da série de Fourier.
- Gerar uma aproximação da série de demanda de energia elétrica com um modelo baseado em Perceptron Múltiplas Camadas.
- Incluir um método de seleção de características usando RRelief-F para qualificar as variáveis.
- Fazer uso dos métodos forward e backward para filtrar as variáveis de influência.
- Comparar os resultados obtidos com o histórico da demanda de energia elétrica e argumentar sobre a solução.

1.3 Estrutura da dissertação

Inicialmente, é feita uma apresentação sobre séries temporais onde são mostradas algumas técnicas e possibilidades de geração de variáveis a partir da própria série.

Na sequência, são incluídas informações em um capítulo sobre redes neurais artificiais e um outro sobre série de Fourier. Esses modelos são utilizados nos experimentos apresentados nos capítulos finais.

Posteriormente, são demonstrados os métodos de seleção de variáveis, classificação e filtros de informação.

Nos capítulos finais, aplica-se os modelos de previsão já apresentados, explicando as soluções de implementação e expondo os resultados em análise.

O trabalho é encerrado com um capítulo de conclusões, onde são incluídas ideias e linhas de investigação para o futuro.

1.4 Literatura relacionada

Diversos trabalhos realizam previsão usando RNA e poucos são os que realizam previsão usando análise da série de Fourier. Aqueles que comparam esses tipos de previsão desmerecem o modelo de Redes Neurais frente à série de Fourier sem apresentar resultados. Atualmente nenhum desses trabalhos é focado na seleção de variáveis comparando os dois métodos.

Os trabalhos citados a seguir são fontes relacionadas com o presente trabalho. Com as técnicas observadas, foi possível elaborar a estrutura para obtenção e melhora dos resultados

([Diniz, 2008](#)) estudou e modelou o comportamento do regime de precipitação em dois municípios da baixada cuiabana/MT, com uma série histórica de 95 anos e outra de 20 anos. O trabalho contou com o auxílio das séries de Fourier e ajuste pelo método dos mínimos quadrados.

Para realizar a previsão da demanda para um sistema de abastecimento de água, ([Brentan, 2015](#)) apresentou um método de correlação de dados, Série de Fourier, aplicado à previsão de demanda em tempo real com atualização simultânea dos coeficientes da série. No trabalho, ele defende sua técnica argumentando as vantagens frente à RNA.

([Ferreira, 2006](#)) apresentou um modelo RNA que abrange parcialmente o território brasileiro baseado no SIN. Em seu trabalho, foi usado um modelo estatístico para aproximação da função real de demanda através da descrição da série com uma função matemática de variáveis endógenas e exógenas correlacionadas. Foi feita uma otimização para encontrar a combinação de coeficientes da função que minimiza o erro de

previsão e isso foi combinado ao modelo de redes neurais ADALINE.

Com o fim de prever o comportamento hidrológico de rios, (Batista, 2009) investigou duas metodologias distintas para realizar previsão de séries temporais. Ele construiu dois modelos em seu trabalho originados da metodologia Box & Jenkins e RNA. O modelo estatístico que melhor se adequou ao seu conjunto de dados foi o SARIMA, e o modelo RNA com melhores resultados foi o MLP.

(Castro, 2004) apresentou um trabalho que usa dados para previsão da bolsa de valores de São Paulo com redes neurais, e tem por objetivo a construção de um sistema para calcular e disponibilizar o índice previsto por meio web. Foram usadas redes MLP para previsão da série Ibovespa, onde as entradas dessas redes foram determinadas a partir da análise estatística individual realizada na série, e também com análises de dependência temporal entre a própria série e as séries de mercado nacional e internacional.

Para fazer a previsão de cargas elétricas em um período de 24 e 48 horas a frente, considerando-se os dados históricos de uma companhia do setor elétrico, (Alves, 2013) propôs uma rede neural híbrida composta pelo método de regressão linear com seleção de variáveis stepwise juntamente com a rede neural artificial perceptron multicamadas com o algoritmo backpropagation.

(Kononenko, 1994) lidou com o problema de estimar a qualidade de atributos com e sem dependências entre eles. Em seu trabalho, ele apresenta um algoritmo chamado ReliefF para tratar de problemas de classificação e (Robnik and Kononenko, 2003) apresenta a forma modificada desse algoritmo para problemas de regressão, o RReliefF.

(Kira and Rendell, 1992) caracterizou a seleção de atributos em seus pontos fortes e fracos. O trabalho fez uma análise do algoritmo Relief para selecionar características relevantes. São relatados os resultados dos testes comparados a outros algoritmos de seleção. Os resultados que sustentam a análise teórica do trabalho são avaliados e é sugerida uma abordagem prática para problemas reais.

1.4.1 Publicações

Durante a construção do trabalho de dissertação de mestrado, foram publicados dois artigos científicos que envolvem a matriz elétrica. Esses trabalhos estão descritos a seguir:

- Uso de redes neurais RBF e GRNN para comparação de previsões de carga de longo prazo de energia elétrica - 12^o Congresso Brasileiro de Inteligência Computacional: Esse trabalho faz a predição de carga de longo prazo de energia da região sudeste do Brasil comparando o resultado da rede RBF à rede GRNN em um espaço de previsão de sete anos, com estimativas por método de autorregressão e por método naïve.
- Otimização de custos do sistema hidrotérmico do Brasil no controle de emissão de dióxido de carbono - Simpósio Brasileiro de Automação Inteligente (SBAI/2015): Esse trabalho busca tratar a redução de custos de energia do Brasil, assim como a emissão de gases efeito estufa na atmosfera.

Capítulo 2

Previsão de séries temporais

Neste capítulo, serão mostrados os componentes da série temporal e as técnicas que se aplicam para a previsão das mesmas. Serão abordados também os fundamentos teóricos de técnicas que são usadas ao longo desta dissertação para extração de variáveis de entrada dos modelos de previsão e extrapolação de tendência da série.

2.1 Séries temporais

Uma série temporal, ou série histórica, é uma sequência de dados obtidos em intervalos regulares de tempo durante um período específico.

Existem duas aproximações fundamentais para a análise de séries temporais sendo elas no domínio do tempo e no domínio de frequência. Essas aproximações são métodos complementares que são ligados matematicamente.

Os métodos de domínio temporal procuram caracterizar as séries de dados nos mesmos termos em que são observados e reportados. A ferramenta para caracterização de relações entre valores na aproximação do domínio temporal é a função de autocorrelação. As análises do domínio temporal operam no mesmo espaço dos valores dos dados ([Chatfield, 1996](#)).

As análises no domínio de frequência representam as séries de dados em diferentes frequências características. Cada escala é representada por um par de funções seno e co-seno. A série completa é considerada como resultante de efeitos combinados de uma coleção de ondas senoidais e co-senoidais oscilando em diferentes taxas. A soma destas ondas reproduz os dados originais ([Chatfield, 1996](#)).

A série temporal se caracteriza pela dependência existente entre as observações vizinhas, sendo esse o seu aspecto de importância em relação ao tipo de dado trabalhado ([Ehlers, 2009](#)).

A importância está na ordem das informações nos modelos de regressão. Essa ordem não tem relevância para a análise, mas sim nas coleções com dados históricos, que levam o tempo em consideração.

A maneira tradicional de analisar uma série temporal é através da sua decomposição nos componentes de tendência, ciclo e sazonalidade. A tendência e variação sazonal são fatores complicadores na resolução e estimativa de previsão.

2.1.1 Componentes não observáveis de uma série histórica

Uma série temporal, segundo a decomposição clássica, pode ser decomposta em quatro componentes não observáveis: tendência, sazonalidade, ciclo e erro. Segundo essa abordagem, as séries podem ser representadas como a combinação desses quatro componentes básicos ([Morettin and Tolo, 1981](#)).

- A tendência de uma série é o indicativo do comportamento dos dados no tempo, isto é, se a sequência de dados é crescente, decrescente ou permanece estável, e qual a velocidade destas mudanças. Nos casos comuns, trabalha-se com tendência constante, linear ou quadrática.
- A sazonalidade indica a repetição de um padrão na série dentro do período. Esse padrão corresponde a um movimento oscilatório de curta duração que traduz a

influência de fatores cuja atuação é periódica.

- Ciclo é um movimento oscilatório de longa duração que exprime a influência de fatores aleatórios. Ele indica as fases de expansão e contração na série.
- O componente irregular, erro, é um movimento instável que exprime a influência de fatores casuais.

2.1.2 Modelos auto-regressivos

A definição da estacionariedade da série é feita antes da sua análise. Em um processo estacionário, a média, a variância e a estrutura de autocorrelação não mudam no decorrer do tempo, ou seja, a estacionariedade numa série temporal significa que os dados oscilam sobre uma média constante, independente do tempo, com a variância das flutuações permanecendo essencialmente a mesma (Diniz et al., 1998).

Os modelos auto-regressivos são modelos estatísticos lineares para análise de séries temporais. Esses modelos usam correlações entre as observações em diversos instantes. Sua ideia envolve filtros lineares que identificam sua estrutura automaticamente, evitando a complexidade da análise. O modelo autorregressivo (AR) e médias móveis (MA) compõem o modelo ARMA, que é utilizado para séries estacionárias (Box and Pierce, 1970) (Box and Jenkins, 1970).

Quando o processo é não estacionário, ou seja, possui tendência, uma das maneiras de analisá-lo é incorporando um processo de diferenças no modelo ARMA. O resultado é o modelo conhecido como ARIMA (modelo autorregressivo integrado de médias móveis) (Diniz, 2008). A Figura 2.1 mostra a diferença gráfica entre um processo estacionário e não estacionário.

Os modelos AR, MA e ARMA são utilizados apenas para séries temporais estacionárias. Caso uma série seja não estacionária, é possível diferenciá-la até torná-la estacionária aplicando o modelo ARIMA.

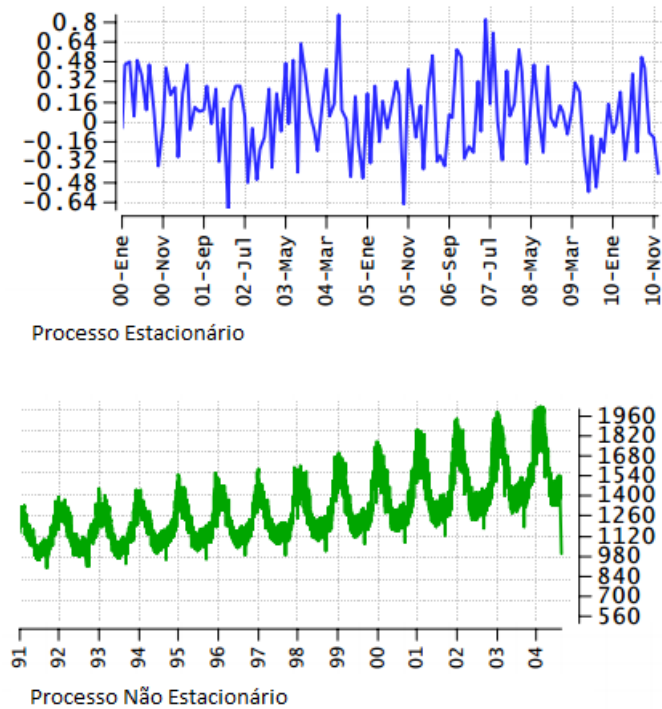


Figura 2.1: Processo Estacionário e Não Estacionário em séries temporais (Davila, 2015)

2.2 Variáveis de entrada

2.2.1 Autocorrelações

A função de autocorrelação amostral mede o grau de correlação de uma variável, consigo mesma, em um dado instante, ou em um instante posterior. Ela permite que se analise o grau de irregularidade de um sinal. Ela mostra o quão forte o valor atual observado está correlacionado com os valores observados no passado e como choques atuais afetam valores futuros da variável estocástica.

Um processo estacionário normal X_t é caracterizado por sua média $\mu_x(t)$ e sua função de auto covariância $\gamma_x(h)$, onde h é o intervalo constante ou atraso. Seja X_t uma série temporal estacionária. A função de auto covariância dessa série é descrita pela equação (2.1) (Brockwell and Davis, 1996).

$$\gamma_x(h) = \gamma_x(h,0) = \gamma_x(t, t+h) = Cov(X_t, X_{t+h}) \quad (2.1)$$

A função de autocorrelação de X_t é dada conforme a equação (2.2):

$$\rho_x(h) = \frac{\gamma_x(h)}{\gamma_x(0)} \quad (2.2)$$

A título ilustrativo, a função de autocorrelação, a seguir, é usada como exemplo na série temporal MA(2) denominada, processo de médias móveis de ordem 2 especificada por

$$y_t = \varepsilon_t - 0.5\varepsilon_{t-1} + 0.4\varepsilon_{t-2} \quad (2.3)$$

onde ε_t é Gaussiana com média 0 e variância 1.

A autocorrelação pode ser visualizada pela Figura 2.2 em uma simulação de 1000 observações. O retorno é um número padrão de 20 atrasos, com as linhas horizontais delimitando os limites de confiança superior e inferior.

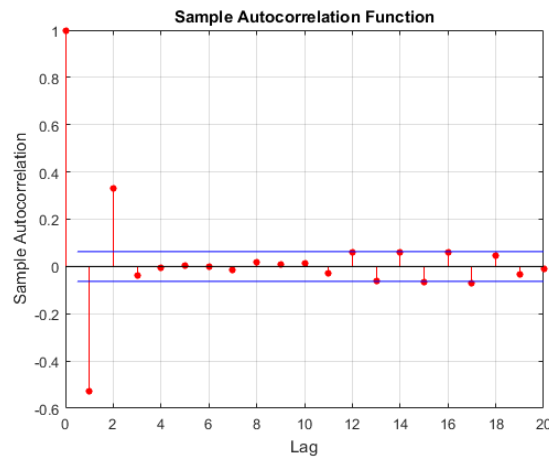


Figura 2.2: Gráfico de autocorrelação

A função de autocorrelação parcial mede a correlação entre as observações que estão h períodos afastados ponderados por um peso φ , entretanto remove o efeito das

correlações intermediárias, ou seja, é a correlação entre x_t e x_{t-h} depois de remover a influência de $x_{t-1}, x_{t-2}, \dots, x_{t-h+1}$. Esse coeficiente de correlação é obtido através da regressão múltipla conforme equação (2.4) (Junior and Caten, 2004).

$$x_t = \varphi_{h1}x_{t-1} + \varphi_{h2}x_{t-2} \dots \varphi_{hh}x_{t-h} \quad (2.4)$$

A função de correlação cruzada entre o processo $X(t)$ e um outro processo $Y(t)$, sendo $X(t)$ e $Y(t)$ séries temporais estacionárias, é definida pela equação (2.5). Novamente, se o processo for estacionário, apenas a diferença de tempo entre os pontos amostrados é relevante e a função de correlação reduz-se à equação (2.6). A função de correlação cruzada reflete o quanto os dois processos estão correlacionados entre si (Lages, 2004).

$$RXY(t_1, t_2) = E[X(t_1)Y(t_2)] \quad (2.5)$$

$$RXY(\tau) = E[X(t)Y(t + \tau)] \quad (2.6)$$

As séries temporais podem ser tratadas como uma amostra dos processos subjacentes. Portanto, as autocorrelações são calculadas pelas estimativas dos valores reais, sujeitos aos erros de amostragem.

As autocorrelações também fazem um importante papel para decidir se uma série temporal é estacionária, para a classe de modelos a qual ela pertence, e quantos coeficientes são caracterizados por ela.

Para que as entradas sejam extraídas para um modelo de previsão, é preciso transformar a série original em uma série estacionária, retirando assim a sua tendência. Desse modo, é feita a autocorrelação, e os atrasos que possuem o índice de correlação mais próximos da série estacionária original são usados como entradas endógenas.

2.2.2 Filtro de Hodrick-Presscott

Um filtro é a combinação das observações originais de uma variável para distintos momentos do tempo, que se realiza com a finalidade de remover algum componente de uma série temporal original.

Filtrar e dessazonalizar são procedimentos adotados para as relações entre as séries temporais. São procedimentos tomados para isolar as frequências de interesse antes de analisar as séries.

O filtro de Hodrick-Presscott (HP) é um método padrão para a remoção de movimentos de tendência na leitura de ciclos das séries. Esse filtro pode ser aplicado aos dados reais no intuito de tornar a série estacionária e nos estudos em que são comparados dados artificiais de um modelo com os dados reais.

O filtro HP busca extrair a tendência, que é considerada estocástica, mas com variações suaves ao longo do tempo e não correlacionadas com o ciclo, através da minimização com respeito a τ_t , como mostrado na equação (2.7) em que a série original é representada por d :

$$d = \left[\sum_{t=1}^T c_t^2 + \lambda \sum_{t=3}^T (\Delta^2 \tau_t)^2 \right] \quad (2.7)$$

onde $c_t = y_t - \tau_t$ e $\Delta^2 = 1 - L^2$, com L sendo o operador de defasagens, $Lx_t = x_{t-1}$. T é o tamanho da amostra e λ é um parâmetro que penaliza a variabilidade da tendência. λ é o parâmetro de suavidade com o qual se controla as variações na taxa de crescimento do componente de tendência.

O primeiro termo da equação (2.7) é a soma dos desvios das séries com respeito à tendência ao quadrado, e é uma medida do grau de ajuste. O segundo termo é a soma de quadrados das segundas diferenças dos componentes de tendência e é uma medida do grau de suavidade.

A Figura 2.3 mostra um exemplo da decomposição de uma série temporal separando

o componente de tendência e o ciclo, função pela qual o filtro é aplicado.

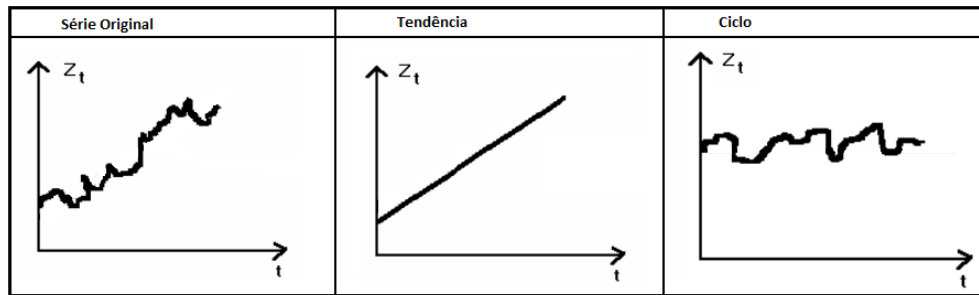


Figura 2.3: Decomposição de série temporal

2.3 Extrapolação de tendências

A extrapolação de tendências é o ajuste de curvas mais confiáveis para séries temporais (Mendonça, 2011). Os modelos de séries temporais determinísticas são aqueles que se prestam a fazer a previsão a partir desse procedimento de extrapolação de tendência com a incorporação de variáveis explicativas ou com a tentativa de captura de processos estocásticos embutidos na série histórica.

A análise de tendências é a forma mais simples de prospecção. Este método é baseado na hipótese de que os padrões do passado serão mantidos no futuro e usa técnicas matemáticas e estatísticas para extrapolar séries temporais para o futuro.

Existem tendências diferentes e é importante caracterizá-las como tendências típicas que são lineares, exponenciais, parabólicas, curvas de logística / curvas de crescimento e curvas ciclo de vida (Miles, 2004).

A tendência linear, foco para extrapolação no presente trabalho, aumenta ou diminui ao longo do tempo a taxas absolutas de acordo com a fórmula básica (2.8):

$$Y = a + b * X \quad (2.8)$$

onde Y = a variável analisada, a = o valor inicial da tendência (em $X = 0$), b =

mudança absoluta de Y por período de observação, e $X =$ a unidade de tempo (ano, mês, trimestre, etc.), em que $x = 0$ no ponto de partida e as unidades de tempo são continuamente atualizados (1, 2, 3, ... n).

A tendência pode ser considerada linear supondo que, por um período de tempo, o processo em estudo represente um componente constante, ou seja, que apareça no gráfico como da Figura 2.4, mas que em algum momento, uma tendência linear comece a emergir. A emergência dessa tendência seria a extrapolação a partir do comportamento da série temporal.

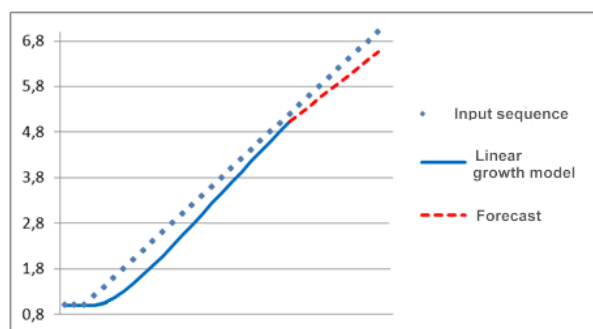


Figura 2.4: Tendência para extrapolação Linear

O exemplo da extrapolação linear pode ser analisado, como na equação (2.8), com a extração dos coeficientes a e b do componente de tendência. O componente a é o valor inicial da série (o valor no período zero), e o componente b é a diferença entre o valor no último período pelo valor do primeiro período, dividido pelo número de períodos. Substituindo os valores encontrados na equação e atualizando o valor para o período X , a extrapolação da tendência é encontrada.

Capítulo 3

Redes Neurais Artificiais

As redes neurais artificiais (RNAs) correspondem a modelos matemáticos com capacidade de aprendizado por meio computacional e são baseadas em redes de neurônios biológicos. Essas redes trabalham com as informações por meio de estruturas de armazenamento e processamento, sendo que cada elemento processador é um neurônio. Esses neurônios recebem um conjunto de entradas ponderadas por pesos e sua saída é obtida pela aplicação de uma função de ativação.

Para determinar entradas do modelo de previsão baseado em RNAs, uma metodologia geralmente adotada quando os dados são baseados em séries temporais é a transformação dos próprios dados da série extraíndo os componentes de tendência e cíclicos. A previsão é feita a partir das séries de ciclos e o resultado final é obtido somando-se o componente de tendência, o qual deve ser extrapolado.

Os resultados encontrados são comparados com os dados históricos de demanda, e na avaliação final da previsão se aproxima dos dados reais. Isso mostra que o modelo de redes neurais é uma opção para que a previsão seja feita, porque possui capacidade de aprendizagem através da generalização, usando os registros da série histórica.

As próximas seções descreverão com detalhes as redes perceptron com múltiplas camadas e o treinamento com algoritmo multi-objetivo.

3.1 Características e aplicações

O trabalho precursor empregando técnicas de Redes Neurais foi realizado em 1945, por *McCulloch e Pitts*, que desenvolveram um estudo comportamental do neurônio biológico com objetivo de criar um modelo matemático para o mesmo. Concluíram, desta pesquisa, a importância da área para implementação computacional do neurônio formal ([Eberhart and Duvins, 1990](#)).

As RNA's ([Braga et al., 2000](#)) são partes do campo científico da computação ligadas à inteligência computacional, e buscam implementar modelos matemáticos que se assemelhem às estruturas biológicas, como mostrado na Figura 3.1. Sendo assim, as redes apresentam capacidade de adaptação dos seus parâmetros como resultado que advém da interação com o meio externo. Sua melhora é gradativa ao seu desempenho na busca da solução de um determinado problema ([Ferneda, 2005](#)).

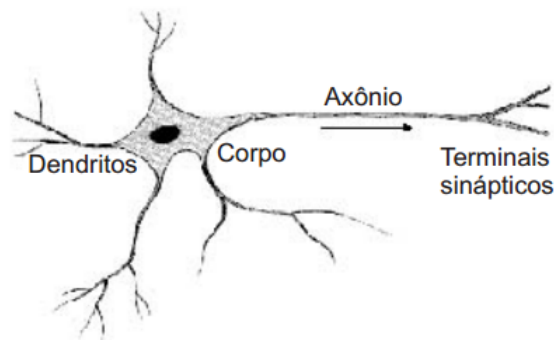


Figura 3.1: Representação simplificada de um neurônio ([Ferneda, 2005](#))

A RNA pode ser definida como uma estrutura de processamento passível de implementação em dispositivos eletrônicos, compondo-se por uma quantidade estabelecida de unidades interconectadas chamadas de neurônios artificiais, sendo que cada unidade representa um comportamento específico de entrada ou saída.

As RNAs são capazes de aprender com exemplos e capturam os relacionamentos entre os dados da série sem que haja conhecimento inicial acerca do problema estudado.

Assim, elas são indicadas para problemas que requerem conhecimentos de especificação mais complexa. (Castro, 2004).

A capacidade de uma RNA de aprender através de um conjunto limitado de exemplos, mapeando relações de entrada, saída e dando respostas para dados não-conhecidos, demonstra o poder dessas redes como ferramenta computacional para solução de problemas complexos. Este atrativo é um diferencial em relação à computação convencional que necessita de algoritmos com regras rígidas e pré-fixadas (Krose and Smagt, 1996).

Os aspectos de componentes físicos com maior evidência que diferenciam RNAs são a conexão e os elementos de processamento. Essa combinação cria a rede propriamente dita, onde pode-se fazer uma analogia a uma estrutura de dados de grafos, em que os nodos seriam os elementos de processamento e as arestas, as conexões. Existem também componentes não físicos, como padrões e funções, que são os dados de entrada da rede com codificação correspondente a uma modelagem da realidade de alguma aplicação definida para a rede (Simpson, 1990).

Os elementos básicos de uma RNA são as conexões, os elementos de processamento, os padrões e as funções, como mostrado na Figura 3.2. Essa terminologia não é uma regra, pois não há uma norma definida que denomine seu uso, porém a visão de funcionamento dos componentes pode ser assim compreendida (Cardon and Müller, 1994).

RNAs estendem a representação de dígrafo por incluir um peso em cada aresta de conexão do sinal de saída, que é passado para o nodo adjacente. Os pesos são os responsáveis pela memorização do padrão, pois são ajustados durante o processo de aprendizado que captura as informações. Desta forma, uma conexão define a informação quanto à modulação que é passada (Cardon and Müller, 1994).

Definiu-se o conceito de atualização de pesos sinápticos no estudo de quatro pontos importantes (Eberhart and Duvins, 1990).

- numa rede neural a informação é armazenada nos pesos;

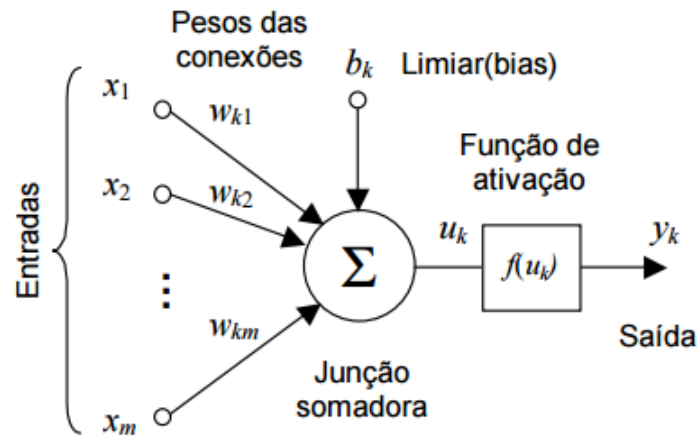


Figura 3.2: O Neurônio Genérico em RNAs (Castro and Zuben, 2015)

- o coeficiente de aprendizagem é proporcional ao produto dos valores de ativação do neurônio;
- os pesos são simétricos;
- quando ocorre o aprendizado, os pesos são alterados.

Existem aplicações de redes neurais na área de finanças, recursos humanos, marketing, medicina, engenharia e outras. A primeira dessas aplicações foi a respeito de técnicas no prognóstico de mercados financeiros, mas a aplicação na previsão de demanda de energia elétrica também está se tornando importante para as empresas distribuidoras. Isso pelo fato de que o auxílio no planejamento está relacionado à própria demanda, análise, controle, necessidade de eventual expansão e fluxo de potência. Então, para garantia da referida previsão de cargas, diversos métodos têm sido utilizados como a inteligência computacional, através do uso de Redes Neurais Artificiais, com dados anteriores, para obtenção das futuras cargas elétricas (Oliveira, 2015).

3.2 Modelos

3.2.1 Classificação

Não há uma norma para que possam ser classificados os modelos de redes neurais, mas existe diferenciação entre eles. Pode-se realizar uma distinção entre os modelos por meio de características básicas, como a entrada e a forma de conexão, ou até mesmo o tipo de aprendizado.

Por tipo de entrada, a classificação pode se dar por binária ou intervalar. Os modelos binários são aqueles que aceitam entradas discretas, e nesse caso englobam modelos como o de Hopfield e o ART. Já os intervalares são os que aceitam valores numéricos reais, e para tais modelos se enquadram o Backpropagation e o Kohonen ([Cardon and Müller, 1994](#)).

A classificação quanto a conexão é definida de acordo com o elemento representado e pela transmissão de sinal na rede ([Cardon and Müller, 1994](#)).

- alimentação à frente, onde os sinais de entrada são simplesmente transformados em sinais de saída;
- retro-alimentação, no qual os sinais são alterados em diversas transições de estado, sendo a saída também alimentadora da entrada;
- competitiva, que realiza a interação lateral dos sinais recebidos na entrada entre os elementos dentro de uma zona de vizinhança.

A maneira pela qual uma rede neural se relaciona com o ambiente também é um fator considerado na classificação. Nesse contexto, existem os paradigmas de aprendizado.

O aprendizado supervisionado é aquele onde a rede neural é treinada com auxílio de um supervisor. Para tanto, a rede deverá possuir pares de entrada e saída. Isso significa um conjunto de entradas e um conjunto com as saídas desejadas para cada entrada. Assim, toda vez que uma entrada for apresentada à rede, deverá ser verificado se a saída

gerada confere com a saída desejada para aquela. Sendo diferente, a rede deverá ajustar os pesos de forma que armazene o conhecimento desejado. Esta interatividade do treino deverá ser repetida com todo conjunto de treinamento para as entradas e saídas, até que a taxa de acerto esteja dentro de uma faixa considerada satisfatória (Alvarez and Luque, 2003).

O aprendizado não supervisionado, ou auto-organização, ocorre quando não existe um agente externo indicando a resposta desejada para os padrões de entrada. Este tipo de aprendizado também é conhecido como aprendizado auto-supervisionado ou de auto-organização por que não requer saída desejada e/ou não precisa usar supervisores para seu treinamento (Alvarez and Luque, 2003).

3.2.2 Perceptron

O primeiro modelo de rede neural implementado foi o perceptron, por Frank Rosenblatt em 1958. Esse modelo é uma rede neural simples constituída de uma camada de entrada e uma camada de saída (Russel and Norvig, 1995).

A rede neural perceptron se caracteriza por possuir entrada intervalar, aprendizado supervisionado e alimentação à frente. Ela se utiliza de um combinador adaptativo linear em que a saída de um elemento processador é a combinação linear das entradas, resultado em um vetor multiplicado por pesos (Cardon and Müller, 1994).

$$S(t) = \sum_{i=0}^{n-1} x_i(t) w_i(t) + b_i \quad (3.1)$$

onde para equação (3.1) o x_i representa as entradas, o b representa o bias para controlar o neurônio e o w_i são os pesos.

Existem modelos de RNA tais como Multilayer Perceptron (MLP), Redes de função de Base Radial (RBF) e Redes de Função Sample (SFNN). Analiticamente, o processo de seleção do modelo tende a ser uma tarefa complexa diante do número de possibili-

dades.

É provável que nem todos os modelos forneçam uma solução excelente, porém uma vantagem no uso de redes MLP é o seu rápido processo de treinamento.

3.2.3 Perceptron de múltiplas camadas

A arquitetura escolhida para modelar a série do presente trabalho se vale de redes em camadas, que são constituídas por uma camada de entrada e uma, ou mais, camadas ocultas, seguidas por uma camada de saída. O sinal se propaga de forma direcional dentro da rede, camada a camada, até a saída.

Existem três características básicas das redes em camadas que são (Velasco, 2007):

- O modelo de cada neurônio da rede inclui uma função de ativação não-linear. Uma das formas não lineares usadas que satisfazem essa exigência é a não linearidade sigmoide.
- Esse tipo de rede possui uma ou mais camadas de neurônios ocultos que não compõe a saída da rede
- A rede exibe um grau elevado de conectividade determinado pelos seus pesos. Uma modificação na conectividade da rede requer modificações nos pesos.

Em decorrência identificação das limitações relativas ao perceptron de camada simples que tinha fronteiras de decisão lineares e funções de lógica simplificadas, foi desenvolvido o multilayer perceptron (MLP), mostrado na Figura 3.3. Este consiste em uma camada de entradas, uma ou mais camadas intermediárias, ou escondidas, e uma camada de saída.

O processamento da informação no MLP possui a fase de propagação, quando o sinal de entrada é propagado através da rede, camada por camada, até produzir uma saída. Há também a fase de adaptação, quando ocorrem os ajustes dos pesos da rede.

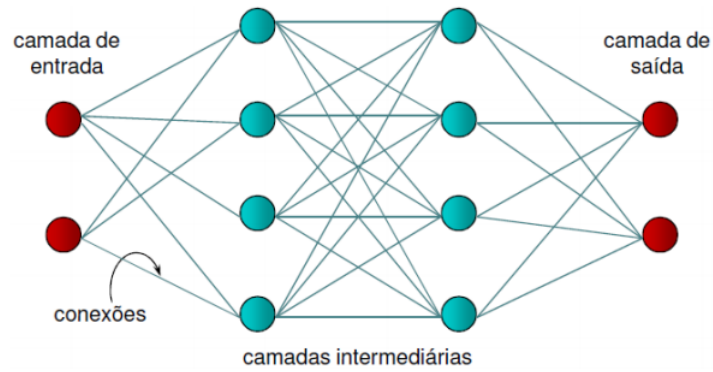


Figura 3.3: Representação de uma MLP (Soares and Silva, 2011)

A Figura 3.4 mostra a arquitetura da MLP com uma única saída tendo uma função sigmoide para a camada intermediária e uma função de ativação linear para a camada de saída.

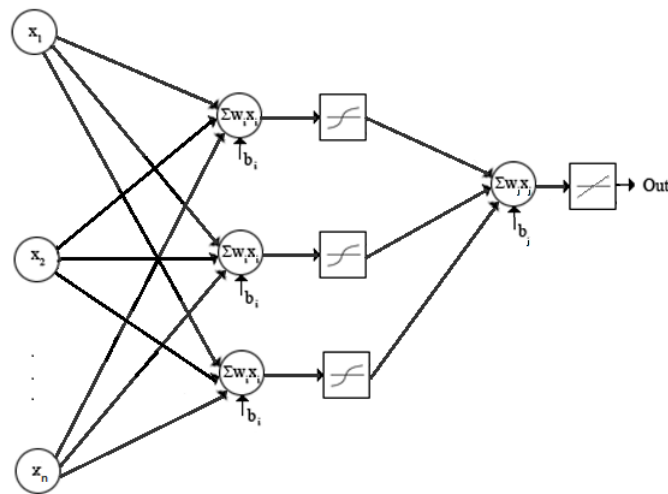


Figura 3.4: Arquitetura da MLP para única saída

A expressão que representa a rede utilizada para esse trabalho é apresentada como

$$f(x, w) = \sum_{j=0}^m w_j \phi_j(S) + b_j \tag{3.2}$$

sendo a representação de

$$S = \sum_{i=0}^n x_i w_i + b_i \quad (3.3)$$

onde x é o vetor de entrada, b_i e b_j representam o bias associados à camada intermediária e à camada de saída, w_i e w_j são os pesos associados à camada intermediária e à camada de saída e $\phi_j(S)$ é a função de ativação tangente hiperbólica aplicada na forma

$$\phi(S) = \frac{2}{(1 + \exp(-2 * S)) - 1} \quad (3.4)$$

As MLPs são aplicadas a problemas através do treinamento de forma supervisionada, onde a rede tem uma resposta que é comparada com a saída desejada, recebendo informações sobre o erro da resposta atual. Para minimização dos erros quadráticos, os algoritmos conhecidos capazes de fazer o treinamento são *RegraDelta* e o *Backpropagation*.

Uma solução através da utilização de algoritmos de treino supervisionado pode se basear na heurística do aprendizado por correção do erro da camada de saída, retropropagando para as camadas intermediárias da RNA. Na literatura, essa técnica é conhecida como *Backpropagation*.

Na especificidade de cada camada, a camada de saída recebe os estímulos da camada intermediária e constrói o padrão que será a resposta. As camadas intermediárias funcionam como extratoras de características. Seus pesos seguem a codificação de características, pois são apresentados nos padrões de entrada e permitem que a rede crie sua própria representação (Tonsig, 2000).

3.3 Treinamento multi objetivo

O treinamento de RNA é um ponto de investigação para pesquisadores da área de aprendizado de máquina. Técnicas para que a generalização dos padrões gere respostas com alta capacidade, fundamentam a importância do treinamento adequado.

O processo de aprendizagem da rede neural tem por objetivo a redução do erro para os padrões de treinamento, mas também procura novos padrões de entrada. Com o algoritmo multi objetivo se faz possível o uso da alta capacidade de generalização oferecida a partir do controle de complexidade do modelo, que independe da sua dimensão com relação aos dados. É possível fazer o controle dos pesos e bias (Medeiros, 2004).

Uma abordagem que pode-se propor para o uso da técnica de otimização multi objetivo é encontrar as soluções a partir do erro quadrático e a função de custo, utilizando-se também a função norma do vetor de pesos.

Uma grande parte dos algoritmos para treinamento supervisionado de RNAs fazem o somatório dos erros quadráticos e o usam como padrões de treinamento da função de custo. Isso não faz com que eles encontrem soluções bem generalizadas, e para que essa capacidade seja atingida deve haver um nível de complexidade médio em que o ruído presente se modere e também possibilite a modelagem completa das características da função que gera os dados (Teixeira, 2001).

O relacionamento entre o comportamento da magnitude dos pesos e o comportamento do erro para os padrões de treinamento são conflitantes, e o modelo apresenta um comportamento como um sistema em que o erro para os padrões de treinamento cresce à medida que a magnitude dos pesos decresce e vice versa como mostrado na Figura 3.5. O objetivo para o caso em questão é o encontro da solução que apresente erros menores para padrões de treinamento que se ajustem aos dados modelando o ruído normalmente presente nos dados. Existe uma importância no equilíbrio entre o erro no treinamento e a complexidade da rede para que os efeitos de *over fitting* e *under fitting*

sejam minimizados (Teixeira, 2001).

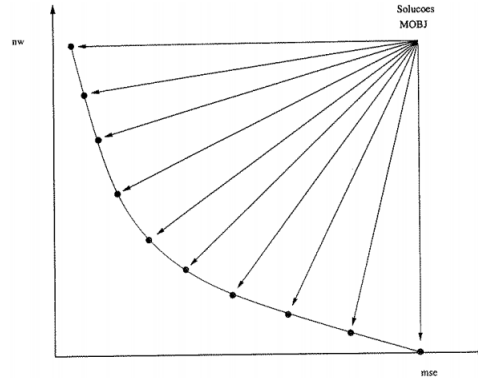


Figura 3.5: Conjunto Pareto (Teixeira, 2001)

A técnica consiste em controlar a complexidade das redes por meio da minimização simultânea do erro para os padrões de treinamento e da norma do vetor de pesos, sendo que o processo de minimização dos objetivos é realizado buscando um ponto que equilibre as soluções que compõe o conjunto Pareto.

O treinamento Multiobjetivo pode ser representado pelas equações

$$\min_{w \in W} f_1(w) \quad (3.5)$$

Sujeito a: $f_2 \leq \epsilon$

$$f_1(w) = \frac{1}{N_T} \sum_i^{N_T} (y d_i - y s_i)^2 \quad (3.6)$$

$$f_2(w) = \|w\| \quad (3.7)$$

onde w representa o vetor de pesos da rede, $y d_i$ é a saída desejada para o padrão de treinamento i , $y s_i$ é a saída simulada para o mesmo padrão e N_T é o número de padrões de treinamento.

Capítulo 4

Método de previsão por série de Fourier

A série de Fourier é a representação de uma função periódica através de uma soma de senoides. Para o estudo de séries temporais que apresentam variações periódicas sazonais, pode-se recorrer à análise harmônica por séries de Fourier para o desenvolvimento de modelos.

Considerando uma série temporal em um intervalo definido, a preparação para previsão dessa série é feita extraíndo sua tendência por meio do filtro de Hodrick-Prescott. A tendência separada e extrapolada é usada no ajuste dos coeficientes.

Com a série temporal preparada, ela então é dividida em dados para ajuste dos coeficientes conforme série de Fourier e para comparação dos resultados.

Assim que os coeficientes da série de Fourier são calculados para a faixa de valores iniciais, os novos dados são previstos reajustando esses coeficientes aplicando-os ao passo $n+1$. Isso é feito para que a saída do modelo possa ser comparada aos dados reais, mas ainda deve ser adicionada a tendência.

Quando é adicionada a tendência extrapolada da série temporal para ajuste dos coeficientes e aplicada à fórmula da série de Fourier, os dados da previsão estarão

prontos para serem comparados aos dados reais no mesmo intervalo a partir da divisão feita anteriormente.

A série de Fourier, pode ser usada não apenas com a promessa de bons resultados, mas também pela agilidade de cálculo dos parâmetros e sua aplicabilidade para futuros modelos.

A seguir, estão os tópicos detalhados, relacionados às séries temporais e funções periódicas que compõem a série de Fourier, assim como as funções trigonométricas e a análise harmônica.

4.1 Funções periódicas

As funções periódicas são aquelas nas quais os valores da função $f(x)$ se repetem para determinados valores da variável x (Oliveira, 2015). Uma função f é dita periódica se existe um número real positivo P , chamado período de f , onde

$$f(x) = f(x + P), \quad (4.1)$$

para todo x no domínio de f . O gráfico de uma função periódica é obtido pela repetição de qualquer intervalo P (Santos, 2004) como mostra a Figura 4.1.

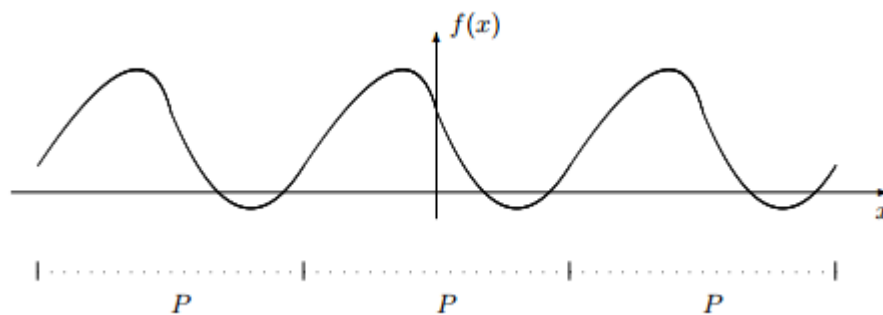


Figura 4.1: Exemplo de uma função periódica (Santos, 2004)

As funções seno e co-seno constituem exemplos familiares de funções periódicas.

Nota-se que a função f é periódica no sentido da definição, porque ela satisfaz $f(x+P) = f(x)$ para qualquer P positivo (Pereira, 2005).

Caso a série seja analisada conforme a tendência, normalmente existe um ciclo dentro de um período. O comportamento da série pode mostrar esses ciclos repetidamente e sua periodicidade.

4.2 Séries de Fourier

A série de Fourier tem aplicações na engenharia elétrica para previsão de demanda assim como para outras áreas.

O modelo da série de Fourier pode ser representado pela função

$$f(x_i) = a_0 + \sum_{j=1}^M [a_j \cdot \cos(j \cdot x_i) + b_j \cdot \sin(j \cdot x_i)]. \quad (4.2)$$

A série representa funções a partir de uma constante a_0 mais um somatório para o cosseno e seno da frequência fundamental representada pela equação (4.3). $f(x_i)$ é o valor da função aproximada pela série de Fourier, j é o j -ésimo termo (Harmônica) da série com M termos e a_j e b_j são os coeficientes ajustáveis da série. Se obedece o critério de $M = \frac{N}{2}$, onde N é o número total de pontos da série temporal.

O ajuste trigonométrico com espaçamento regular entre os pontos é o ajuste por Série de Fourier. Assim, para pontos igualmente espaçados, normalizados tem se a equação (4.3) (Brentan, 2015) em que x_i é a abscissa normalizada no intervalo $[0, 2\pi]$ e i é a posição do ponto na série temporal.

$$x_i = \frac{2\pi i}{N}. \quad (4.3)$$

As harmônicas são as frequências múltiplas de uma frequência fundamental e essa frequência caracteriza a função periódica. Os coeficientes da série de Fourier são ajus-

tados a partir dessas harmônicas, para fazer a representação da função. Nesse modelo de previsão, é feita a extração desses coeficientes dependendo do tamanho da série.

Considerando que d seja o somatório dos valores da série original e d^* o somatório dos valores calculados pelo método para aproximação dessa série, o erro quadrático pode ser escrito pela equação

$$e = (d - d^*)^2. \quad (4.4)$$

Substituindo d^* na forma apresentada em (4.2) na equação (4.4), é obtida a equação

$$e = (d - a_0 + \sum_{j=1}^M [a_j \cdot \cos(j \cdot x_i) + b_j \cdot \sin(j \cdot x_i)])^2. \quad (4.5)$$

Aplicando o método dos mínimos quadrados para obter os termos da série, o sistema linear a ser resolvido pode ser encontrado pela equação

$$\begin{bmatrix} \frac{\partial e}{\partial a_0} \\ \frac{\partial e}{\partial a_j} \\ \frac{\partial e}{\partial b_j} \end{bmatrix} = [0]. \quad (4.6)$$

A ortogonalidade entre senos e cossenos resulta nas condições expressas pelas equações (4.7), (4.8) e (4.9) (Ralston, 1978).

$$\sum_{i=1}^N \text{sen}(j \cdot x_i) \cdot \text{sen}(k \cdot x_j) = \begin{cases} 0, & \text{para } j \neq k, \\ \frac{N}{2}, & \text{para } j = k \neq 0, \\ 0, & \text{para } j = k = 0. \end{cases} \quad (4.7)$$

$$\sum_{i=1}^N \cos(j \cdot x_i) \cdot \cos(k \cdot x_j) = \begin{cases} 0, & \text{para } j \neq k, \\ \frac{N}{2}, & \text{para } j = k \neq 0, \\ 0, & \text{para } j = k = 0. \end{cases} \quad (4.8)$$

$$\sum_{i=1}^N \text{sen}(j.x_i) \cdot \text{cos}(k.x_j) = 0 \quad \forall \quad j, k \quad (4.9)$$

Com as quais, a equação (4.6) resulta no sistema linear apresentado pela equação (4.10).

$$\begin{bmatrix} N & 0 & 0 \\ 0 & \frac{N}{2} & 0 \\ 0 & 0 & \frac{N}{2} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_j \\ b_j \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N d_i \\ \sum_{i=1}^N d_i \cdot \text{cos}(j.x_i) \\ \sum_{i=1}^N d_i \cdot \text{sen}(j.x_i) \end{bmatrix} \quad (4.10)$$

O cálculo dos coeficientes é dado pelas equações (4.11), (4.12) e (4.13) que, aplicados na série de Fourier na equação (4.2), geram o resultado da aproximação (Brentan, 2015).

$$a_0 = \frac{\sum_{i=1}^N d_i}{N} \quad (4.11)$$

$$a_j = \frac{\sum_{i=1}^N d_i \cdot \text{cos}(j.x_i)}{N} \quad (4.12)$$

$$b_j = \frac{\sum_{i=1}^N d_i \cdot \text{sin}(j.x_i)}{N} \quad (4.13)$$

Com os coeficientes da série calculados e conhecendo o seu comportamento, a equação (4.2) pode ser usada para a previsão. Essa equação tem a informação até o passo de tempo N e, observando o seu comportamento, pode ser verificada a possibilidade de ajuste dos coeficientes aplicando-os ao passo N+1. Isso pode ser feito caso não haja expressivas alterações entre um passo de tempo N e um passo de tempo N-1.

Quando a diferença real entre um passo de tempo N e um passo de tempo N+1 é pequena, o método não leva a um erro significativo. Quando a distância entre o intervalo de ajuste dos coeficientes da série e o passo de tempo a ser previsto aumenta, os erros de previsão certamente aumentarão (Brentan, 2015).

Assim, para evitar o erro em cada passo de tempo, atualizam-se os coeficientes da série, a informação é lida no passo de tempo anterior e a janela de dados é movida.

A nova janela de dados considera o valor real da demanda medida no tempo $N+1$. Para ilustrar esse processo de atualização, a Figura 4.2 mostra a previsão para dois passos de tempo seguidos.

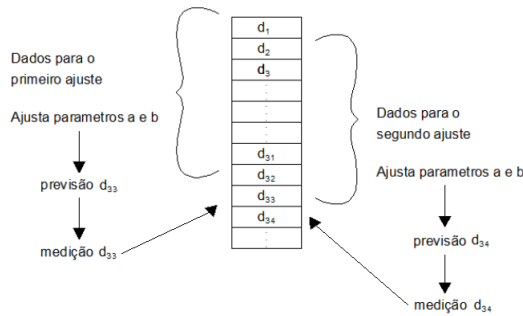


Figura 4.2: Esquema de previsão com parâmetros atualizados

Há também a previsão onde é feito o ajuste dos parâmetros e esses são mantidos para os passos seguintes. A equação (4.2) é aplicada duas vezes, sendo a primeira vez com os parâmetros ajustados para representação da série original, e a segunda vez para previsão, como mostra o esquema de previsão na Figura 4.3.

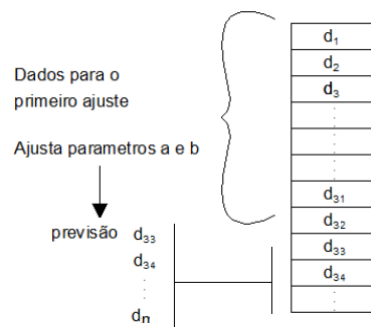


Figura 4.3: Esquema de previsão com parâmetros fixos

Para o uso da previsão com parâmetros fixos, devem ser usados todos os $n/2$ harmônicos (termos j), para obter exatamente a representação da função através de todos os pontos da série. Dessa forma, o ajuste será conforme a série original.

Capítulo 5

Seleção de variáveis

O método de seleção de variáveis tem como objetivo selecionar as variáveis mais importantes em um modelo de regressão, descartando as variáveis menos significativas.

Considerando o modelo baseado em Redes Neurais Artificiais, o objetivo foi determinar o subconjunto de variáveis exógenas que melhor explique a variável resposta, isto é, dentre todas as variáveis disponíveis, deve-se encontrar um subconjunto de variáveis que quando combinadas são importantes para o modelo.

Construir um modelo que inclui um subconjunto de variáveis envolve obter o máximo de informação a partir de inúmeras combinações possíveis e, desta forma, obter um equilíbrio avaliando o erro e a correlação entre essas combinações.

5.1 Definição das variáveis

A análise de dependência entre as variáveis auxilia na obtenção dos resultados. O objetivo da busca pelas variáveis é justamente ir ao encontro daquelas que podem ter influência no comportamento da série de demanda de energia elétrica. Desse modo, foram coletadas variáveis socio-econômicas e climáticas.

As séries analisadas, que têm dependência, foram extraídas de vários repositórios

de dados estatais na Internet, como IBGE, CEMIG, SIN e INMET. Esses dados foram extraídos desses repositórios porque eles são públicos, estão constantemente atualizados e possuem conteúdo parcialmente completo.

Contrárias às variáveis de influência, o uso de variáveis excedentes, não informativas, pode prejudicar o desempenho do modelo baseado em RNAs. Por isso, é importante a etapa de seleção de variáveis.

As 32 variáveis da base de dados utilizadas, em que foi feita a relação com a série de demanda de energia elétrica, estão listadas a seguir:

- Balanço Energético Energia: este apresenta a contabilização da oferta, transformação e consumo final de produtos energéticos no Brasil.
- Capacidade de Geração: potencial de geração de energia no decorrer dos anos.
- Capacidade industrial: capacidade de produção industrial no Brasil.
- Clima - Direção do vento: comportamento do vento no território brasileiro.
- Clima - Nebulosidade: acúmulo de gases atmosféricos no território brasileiro.
- Clima - Precipitação: índice de chuvas no Brasil.
- Clima - Pressão: média de pressão atmosférica anual.
- Clima - Temperatura: média de temperatura anual.
- Clima - Umidade relativa do ar: umidade relativa do ar média anual para o território brasileiro.
- Clima - Velocidade do vento: velocidade do vento no território brasileiro.
- Cotação do Dolar: o valor da moeda americana frente à moeda brasileira.
- Índice Setorial - Agropecuário : índices específicos para medir a evolução nos preços de itens relativos à agricultura e à pecuária brasileira.

- Índice Setorial - Comercial: índices específicos para medir a evolução nos preços de itens relativos ao comércio brasileiro em geral.
- Índice Setorial - Industrial: índices específicos para medir a evolução nos preços de itens relativos à indústria brasileira em geral.
- Índice Setorial - Público: índices específicos para medir a evolução nos preços de itens relativos ao setor público brasileiro.
- Índice Setorial - Residencial: índices específicos para medir a evolução nos preços de itens relativos aos domicílios dos brasileiros.
- Índice Setorial - Transporte: índices específicos para medir a evolução nos preços de itens relativos ao setor de transporte brasileiro.
- PIB anual: produto interno bruto do Brasil.
- População: crescimento populacional brasileiro.
- Porcentagem de Inflação: taxa de aumento no nível de preços no decorrer dos anos.
- Salário mínimo: mínimo valor de salário que os empregadores pagaram legalmente aos seus funcionários no decorrer dos anos.
- SELIC - Meta: meta de qual taxa de juros seria cobrada pelos bancos no Brasil.
- SELIC - Taxa: índice pelo qual as taxas de juros cobradas pelos bancos se balizam no Brasil.
- Valor tarifário - Comercial: tarifas de energia aplicadas ao comércio brasileiro.
- Valor tarifário - Iluminação pública: tarifas de energia para iluminação pública.
- Valor tarifário - Industrial: tarifas de energia aplicadas a indústria brasileira.

- Valor tarifário - Poder público: tarifas de energia aplicadas ao poder público brasileiro.
- Valor tarifário - Próprio: tarifas de energia aplicadas para o próprio setor elétrico brasileiro.
- Valor tarifário - Residencial: tarifas de energia aplicadas ao setor residencial brasileiro.
- Valor tarifário - Rural: tarifas de energia aplicadas ao setor rural sem irrigação.
- Valor tarifário - Rural irrigante: tarifas de energia aplicadas ao setor rural com tecnologia de irrigação.
- Valor tarifário - Saneamento: tarifas de energia aplicadas ao setor de saneamento.

A base de dados com as variáveis foi processada com o objetivo de padronizar a quantidade de registros em relação a série de demanda no intervalo dos anos de 1970 a 2016. As séries extraídas estavam parcialmente completas, então foram usadas técnicas de extrapolação de dados e também a análise por série de Fourier para completá-las.

O Balanço Energético e a Capacidade de Geração têm a base original com dados de 1974 a 2006. Essas séries são estacionárias, então os dados de 1970 a 1974, assim como os dados de 2006 à 2016, foram completados usando a análise por série de Fourier.

Os dados de valor tarifário e índice setorial têm a base original com dados de 1970 a 2004. Essas séries são não estacionárias, então os dados de 2004 a 2016 foram completados usando extrapolação linear.

5.2 Sistema repositório de dados

A quantidade de registros armazenados para uso nesse trabalho foi buscada de forma descentralizada, ou seja, esses dados não foram coletados originalmente em uma única

base.

A base de dados foi estruturada a partir de pesquisa e extração de dados de repositórios on-line, onde foram criados códigos na linguagem de programação java para importação dos dados, e os mesmos foram inseridos em uma base MySQL.

Levando-se em conta que 32 variáveis foram selecionadas para entrada do modelo, sendo que cada uma delas possui registros de 1970 a 2016, o total de dados, considerando os dados da própria demanda de energia, é de 1518. Essa não é uma base com grandes proporções, mas com dados específicos que precisam ser gerenciados.

Para que se pudesse agilizar o gerenciamento dos dados foi então criada uma aplicação própria para controle da informação.

A aplicação pode ser visualizada na Figura 5.1

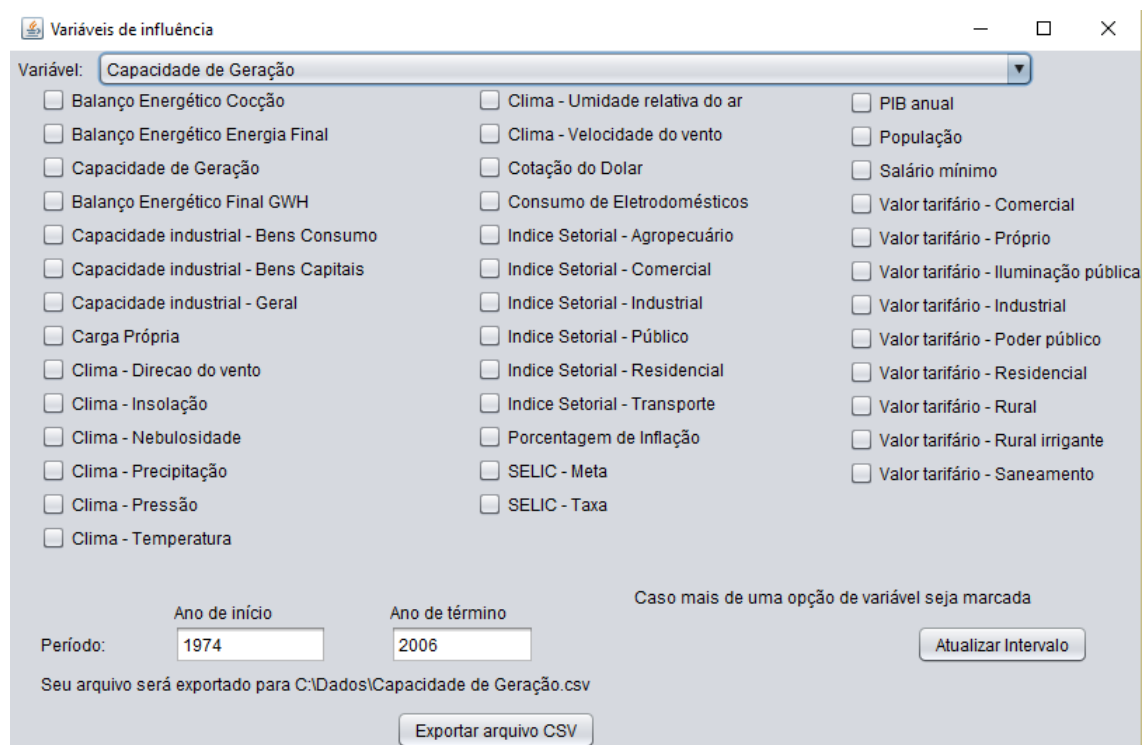


Figura 5.1: Aplicativo de gerenciamento de dados

A aplicação gera relatórios que podem ser importados para o modelo. Esses relató-

rios, no formato CSV, podem ser gerados a partir de uma variável ou um conjunto de variáveis selecionado na interface gráfica.

O sistema funciona da seguinte forma:

- O usuário seleciona uma variável no componente de lista ou nos componentes de marcação, sendo que o componente de marcação possibilita escolha de várias variáveis.
- O usuário escolhe um intervalo para gerar o relatório com os registros das variáveis selecionadas
- O relatório é gerado e aberto para que o usuário visualize os dados.

O ponto importante a ser observado no sistema criado é o componente de escolha de intervalos dos dados em anos, que também podem ser alterados pelo botão correspondente. Esse componente exibe o limite de informações presente na base à medida que as variáveis são selecionadas, e isso é refletido no relatório gerado.

5.3 Relief

Para definir uma ordem de qualidade entre as variáveis e selecionar as mais importantes para o modelo, foi necessário o uso de um filtro de seleção de características.

O Relief foi um algoritmo proposto para estimar a relevância de atributos discretos e contínuos em dados que caracterizam duas classes (Kira and Rendell, 1992). O algoritmo basicamente utiliza métodos estatísticos para selecionar as características relevantes, dando pesos a cada uma dessas características. (DASH, 1997). Ele realiza a busca de maneira sequencial e utiliza medidas de distância para poder realizar a avaliação.

O Relief é o primeiro algoritmo da família Relief-F. Ele foi proposto para dados binários, e o seu princípio de funcionamento está em destacar atributos para vizinhos

mais próximos em classe diferente, bem como penalizar atributos para vizinhos mais próximos na mesma classe. Para isso, o algoritmo busca vizinhos próximos para cada exemplo R_i escolhido aleatoriamente: um de mesma classe - *nearesthit* H-, e o outro de classe diferente - *nearestmiss* M. Desse modo, o algoritmo considera $K = 1$ vizinhos mais próximos de cada classe de um problema (Demsar, 2010). O pseudo-código que explica com mais detalhes o Relief pode ser visto em Algorithm 1.

Algorithm 1 Algoritmo Relief (Robnik and Kononenko, 2003)

Entrada: O conjunto de treinamento e o número de iterações m
 Saída: o vector W de estimativas das qualidades dos atributos
 inicializa o vector de pesos $W[A] := 0:0$;
for $i := 1$ to m **do**
 seleciona aleatoriamente uma instancia R_i ;
 encontra hit mais próximo H e o miss mais próximo M ;
 for $j := 1$ to a **do**
 $W[A_j] := W[A_j] - diff(A_j; R_i; H)/m + diff(A_j; R_i; M)/m$;
 end for
end for

Relief usa a medida *diff* para calcular a diferença entre os valores. Ela é constituída pelas medidas de *overlap* para atributos categóricos e *RangeNormalizedDifference*(RND) para atributos numéricos, como nas equações do exemplo $diff(A_j; R_i; M)$

$$diff(A_j, R_i, M) = \begin{cases} \text{overlap}(A_j, R_i, M) & \text{se } A_j \text{ é categórico} \\ \text{RND}(A_j, R_i, M) & \text{se } A_j \text{ é numérico} \end{cases} \quad (5.1)$$

$$\text{overlap}(A_j, R_i, M) = \begin{cases} 0 & \text{se } A_j, R_i = A_j, M \\ 1 & \text{se } A_j, R_i \neq A_j, M \end{cases} \quad (5.2)$$

$$\text{RND}(A_j, R_i, M) = \frac{A_j, R_i - A_j, M}{\max(A_j) - \min(A_j)} \quad (5.3)$$

Com exemplo mostrado na Figura 5.2 pode ser acompanhado o fluxo de funcionamento do algoritmo Relief em 2 iterações.

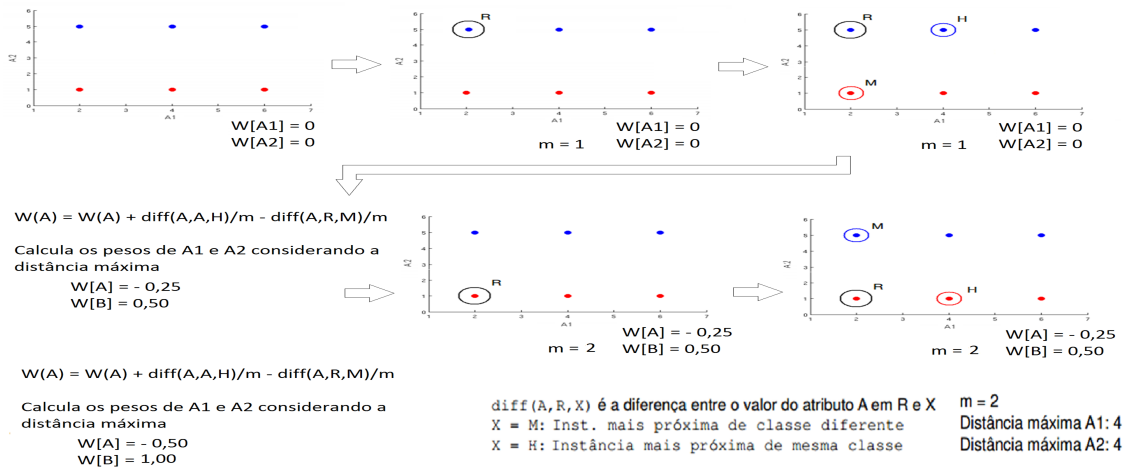


Figura 5.2: Fluxo de execução do algoritmo Relief

5.3.1 Relief-F

O Relief apresenta restrições, como o tratamento de dados com muito ruído e também a incapacidade de manipular bases de dados incompletas. O tratamento de domínios com mais de duas classes também não pode ser feito diretamente por esse algoritmo. Essas deficiências foram contornadas pelos algoritmos da família Relief propostos por (Kononenko, 1994), que são as variantes Relief-A, Relief-B, Relief-C, Relief-D, Relief-E e, sendo a mais eficiente e robusta, a variante Relief-F.

O Relief-F é a variante final que trata os problemas de múltiplas classes de forma mais abrangente que as variações anteriores da família Relief (Kononenko, 1994). Nela é encontrado um M próximo para cada classe diferente, exigindo um maior processamento computacional. A média entre todos os M próximos encontrados é utilizada na alteração dos pesos de relevância dos atributos. O pseudo-código que explica com mais detalhes o Relief-F pode ser visto em Algorithm 2.

As melhorias com o Relief-F são obtidas da seguinte maneira:

- Uso de $k > 1$ vizinhos mais próximos para cada rótulo;
- Tratamento probabilístico de valores faltantes durante o cálculo de dissimilaridade

entre exemplos no espaço de atributos.

A busca por $k > 1$ vizinhos mais próximos garante maior robustez em dados com ruídos, desse modo, para cada R_i escolhido aleatoriamente de uma classe y_i , o Relief-F trabalha com os *hits* k mais próximos $H, k(R_i)$ e os *misses* k mais próximos $M, k(R_i, y_k)$ para cada classe y_k , $y_k \neq y_i$.

Algorithm 2 Algoritmo Relief-F (Robnik and Kononenko, 2003)

Entrada: O conjunto de treinamento e o número de iterações m

Saída: o vector W de estimativas das qualidades dos atributos

Número de K vizinhos mais próximos

inicializa o vector de pesos $W[A]$ com zero;

for $i := 1$ to m **do**

 seleciona aleatoriamente uma instancia R_i ;

 encontra k hits mais próximos (H, k);

for cada classe $y_k \neq$ da classe de R_i , que é y_i **do**

 para a classe y_k encontra misses k mais próximos;

end for

for $j := 1$ to a **do**

$W[A_j] := W[A_j] + \sum_{z=1}^k \text{diff}(A_j, R_i, Hk(R_i, y_k, z)) / (m \cdot k) + \sum_{y_k \neq y_i}$

$[\frac{P(y_k)}{1 - P(y_i)} \sum_{z=1}^k \text{diff}(A_j, R_i, Mk(R_i, z))] / (m \cdot k)$;

end for

end for

Para normalizar os pesos, o algoritmo divide cada probabilidade previamente calculada pela soma das probabilidades envolvendo atributos de classes diferentes. O Relief modifica a medida *diff* para avaliar a probabilidade de R_i e M terem valores diferentes em um atributo A_j .

Se R_i não é conhecido em A_j

$$\text{diff}(A_j, R_i, M) = 1 - P(A_j, R_i | y = y_i) \quad (5.4)$$

Se M não é conhecido em A_j

$$\text{diff}(A_j, R_i, M) = 1 - P(A_j, M | y = y_i) \quad (5.5)$$

Se R_i e M não são conhecidos em A_j , sendo *valores* A_j o número de valores de A_j

$$\text{diff}(A_j, R_i, M) = 1 - \sum_{V}^{\text{valores } A_j} P(V|y = y_i) \cdot P(V|y = y_k) \quad (5.6)$$

5.3.2 RRelief-F

O algoritmo RRelief-F é a variação da família Relief usado em problemas de regressão. Ele faz a modelagem da probabilidade de valores contínuos que sejam de instâncias diferentes. A partir da probabilidade, o algoritmo aproxima a estimativa de importância dos atributos.

A estimativa obtida pelo Relief pode ser aproximada pela diferença entre probabilidades relacionadas às instâncias. Se considerado O_{dA} o evento no qual duas instâncias possuem valores diferentes para um atributo, O_{dY} o evento no qual duas instâncias possuem classes diferentes e $O_{\bar{dY}} = 1 - O_{dY}$ o evento no qual duas instâncias tem a mesma classe, tem-se

$$\text{Relief}(A_j) = P(O_{dA}|O_{dY}) - P(O_{dA}|O_{\bar{dY}}) \quad (5.7)$$

Os cálculos que reescrevem a probabilidade $(O_{dA}|O_{dY})$ e fazem a reformulação da probabilidade $P(O_{dA}|O_{\bar{dY}})$ podem ser conforme

$$P(O_{dA}|O_{dY}) = \frac{P(O_{dY}|O_{dA})P(O_{dA})}{P(O_{dY})} = \frac{P(O_{dY} \cap P(O_{dA}))}{P(O_{dY})} \quad (5.8)$$

Substituindo os termos na equação tem-se o RRelief-F definido como

$$\text{RReliefF}(A_j) = \frac{P(O_{dY} \cap P(O_{dA}))}{P(O_{dY})} - \frac{P(O_{dA}) - P(O_{dY} \cap P(O_{dA}))}{1 - P(O_{dY})} \quad (5.9)$$

$P(O_{dY} \cap P(O_{dA}))$ é a probabilidade em que duas instâncias mais próximas possuem

diferentes valores de rótulo e do atributo A_j . Assim como ocorre com as classificações, essa probabilidade é modelada por meio de uma função de dissimilaridade para o RRelief-F. Para tanto, é utilizada a função *diff*, mencionada anteriormente.

Os pesos específicos para o algoritmo RRelief no pseudo-código do Algorithm 3 são definidos conforme a reformulação por:

- W_{dY} para rótulos diferentes, relacionado à probabilidade $P(O_{dY})$;
- W_{dA} para valores de atributos diferentes, relacionado à probabilidade $P(O_{dA})$;
- W_{dYA} para rótulos e valores de atributos diferentes, relacionado à probabilidade $P(O_{dY} \cap P(O_{dA}))$.

O termo $d(R_i, R_i k_z)$ no Algorithm 3 pondera cada vizinho mais próximo de $R_i k_z$ de acordo com a dissimilaridade entre esse vizinho e a instância aleatoriamente selecionada R_i . Quanto mais próximas as instâncias, maior a influência de seus vizinhos.

5.4 Métodos de seleção

5.4.1 Método Forward

O procedimento de seleção forward é baseado no princípio de que as variáveis preditoras devem ser adicionadas ao modelo uma de cada vez até que não haja mais variáveis.

O método consiste na inclusão de variáveis satisfazendo algum critério estatístico para a seleção (Alves, 2013). Neste trabalho, o critério de seleção utilizado foi a correlação linear e o erro médio quadrático (MSE), da seguinte forma:

- a variável que estiver na primeira posição após a qualificação pelo algoritmo RRelief-F é a variável selecionada para iniciar o método;
- com a variável inicialmente adicionada, calcula-se o coeficiente de correlação e MSE da saída do modelo;

Algorithm 3 Algoritmo RReliefF (Robnik and Kononenko, 2003)

Entrada: O conjunto de treinamento D_{tr} e o número de iterações m
 Saída: o vector W de estimativas das qualidades dos atributos
 Número de K vizinhos mais próximos
 inicializa o vector de pesos $W[A]$ com zero;
 inicializa o vector W_{dY} com zero;
 inicializa o vector W_{dA} com zero;
 inicializa o vector W_{dYA} com zero;
for $i := 1$ to m **do**
 seleciona aleatoriamente uma instância R_i ;
 encontra k vizinhos mais próximos (R_{ik}) com D_{tr} ;
 for $j := 1$ to z **do**
 $W_{dY} = W_{dY} + \text{diff}(Y, R_i, R_{ik_z}).d(R_i, R_{ik_z})$;
 for $j := 1$ to a **do**
 $W_{dA}(A_j) = W_{dA}(A_j) + \text{diff}(A_j, R_i, R_{ik_z}).d(R_i, R_{ik_z})$;
 $W_{dYA}(A_j) = W_{dYA}(A_j) + \text{diff}(Y, R_i, R_{ik_z}).\text{diff}(A_j, R_i, R_{ik_z}).d(R_i, R_{ik_z})$;
 end for
 end for
end for
for $j := 1$ to a **do**
 $W(A_j) = W(A_j) + \frac{W_{dYA}(A_j)}{W_{dY}} - \frac{W_{dY}(A_j) - W_{dYA}(A_j)}{m - W_{dY}}$
end for

- se o coeficiente de correlação gerado com a adição for maior do que o valor do coeficiente de correlação ou o MSE for menor que no modelo anterior, então a variável fica no modelo, caso contrário, sai do modelo;
- o processo continua até que não existam mais variáveis a serem introduzidas no modelo.

Na Figura 5.3, é apresentado o fluxograma do método forward.

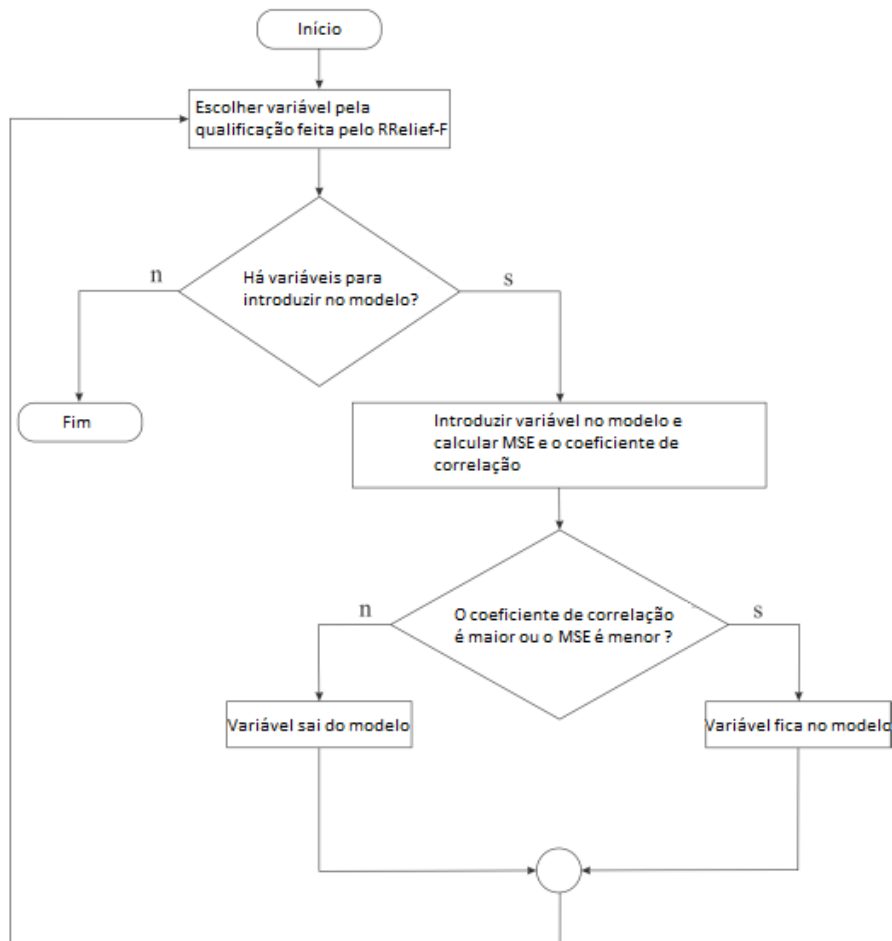


Figura 5.3: Fluxograma do método forward

5.4.2 Método Backward

O procedimento de seleção Backward é baseado no princípio de que as variáveis preditoras devem ser todas adicionadas ao modelo de uma só vez e assim, vão sendo removidas uma por vez até que não haja mais variáveis.

O método consiste na inclusão de todas as variáveis satisfazendo algum critério estatístico para identificar se as variáveis ficam ou saem do modelo (Alves, 2013). Neste trabalho, assim como para o método forward, o critério de seleção utilizado foi a correlação linear e o erro médio quadrático (MSE), em que o procedimento é da seguinte forma:

- é retirada a variável do modelo conforme raqueamento pelo RRelief-F iniciando pela última até a primeira, no qual é calculado o coeficiente de correlação e o MSE com as variáveis que ficaram no modelo;
- se o coeficiente de correlação calculado com a retirada for maior do que o coeficiente de correlação ou o MSE for menor que o MSE anterior, então fica fora do modelo, caso contrário, volta ao modelo;
- o processo é repetido até que todas as variáveis sejam testadas.

Na Figura 5.4, é apresentado o fluxograma do método Backward.

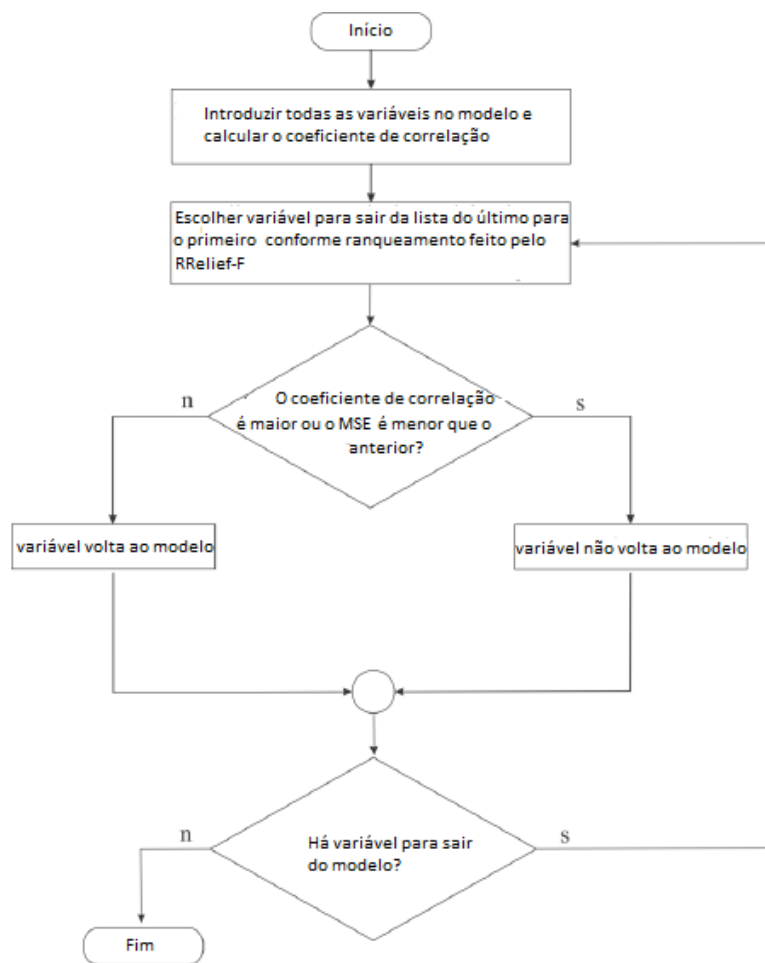


Figura 5.4: Fluxograma do método backward

Capítulo 6

Realização da previsão

A proposta desse trabalho baseia-se no uso de uma rede neural com entradas previamente trabalhadas com técnicas específicas e também um modelo de análise por série de Fourier.

Alguns experimentos consideram o modelo contendo variáveis endógenas (atrasos da série de demanda) e outros consideram variáveis endógenas e exógenas selecionadas via métodos de seleção descritos no Capítulo 5.

O modelo que considera as entradas endógenas e exógenas para a rede neural é composto pela qualificação com método RRelief-F, por testes com Forward e Backward para filtro de entrada e pela arquitetura da rede perceptron multicamadas com treinamento multiobjetivo, aplicada ao problema de previsão da demanda de energia elétrica.

As entradas para o modelo por análise de Fourier usaram a própria série histórica de demanda para adequação e previsão.

Neste capítulo, são apresentados resultados para os modelos de previsão de demanda de energia elétrica do Brasil. Os dados usados para criação da base dos modelos são do ano de 1970 a 2016. A Figura 6.1 mostra a evolução da demanda no passar dos anos.

Os dados de entrada dos modelos não foram exatamente os dados de demanda de energia elétrica mostrados na série na Figura 6.1. Essa série foi trabalhada extraindo

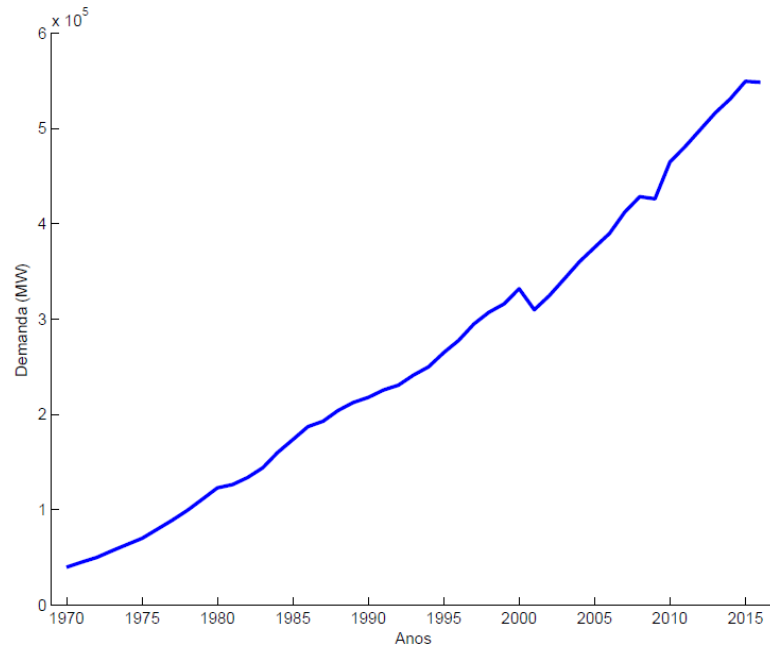


Figura 6.1: Demanda de energia elétrica

a sua tendência a partir do filtro de Hodrick-Presscott, deixando o componente cíclico. Então, esse componente foi interpolado linearmente para ser usado como entrada.

A Figura 6.2 mostra a série de demanda de energia estacionária, sem o seu componente de tendência.

Para avaliação do desempenho das previsões obtidas tanto pelos modelos RNA como pela análise por série de Fourier, foram consideradas as métricas do erro percentual médio absoluto *Mean Absolute Percentage Error* (MAPE) e a raiz do erro médio quadrático *Root Mean Square Error* (RMSE), além da correlação entre as séries original e predita.

Usando o MAPE como um dos critérios de avaliação, comparou-se os valores reais da demanda com os valores previstos através dos modelos, calculado com a equação (6.1).

$$MAPE = \frac{1}{N} \sum_{j=1}^N \left| \frac{alvo_j - pred_j}{alvo_j} \right| * 100\% \quad (6.1)$$

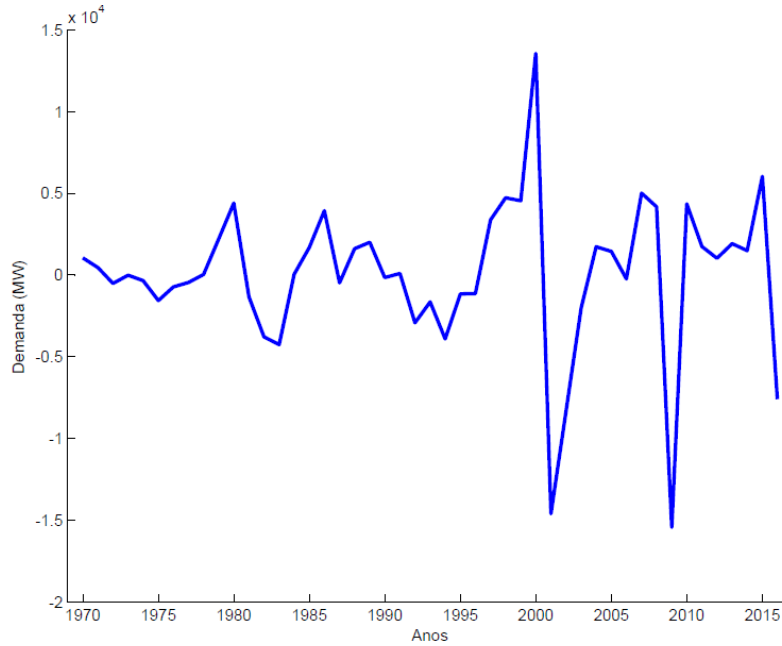


Figura 6.2: Componente cíclico da demanda de energia elétrica

onde N é o número de amostras previstas, $alvo_j$ são os valores reais desejados na previsão e $pred_j$ representa os valores previstos pelos modelos.

Outra métrica utilizada para avaliação foi a RMSE. Essa é uma das métricas mais utilizadas para avaliação de erro em modelos de previsão. A avaliação de um previsor informa a variância do erro no intervalo de observação definida pela equação (6.2).

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (alvo_j - pred_j)^2} \quad (6.2)$$

em que, da mesma forma que na equação (6.1), N é o número de amostras previstas, $alvo_j$ são os valores reais desejados na previsão e $pred_j$ representa os valores previstos pelos modelos.

Para avaliar o desempenho, foram confrontados neste estudo os resultados da modelagem naïve. Essa previsão ingênua é aquela em que se toma o último registro conhecido como sendo a previsão para o próximo período. Tal técnica foi escolhida por ser uti-

lizada para a determinação da capacidade preditiva mínima de modelos de previsão (MAKRIDAKIS et al., 1983).

A Figura 6.3 mostra a previsão pela modelagem naïve para a série de demanda não estacionária utilizada no trabalho.

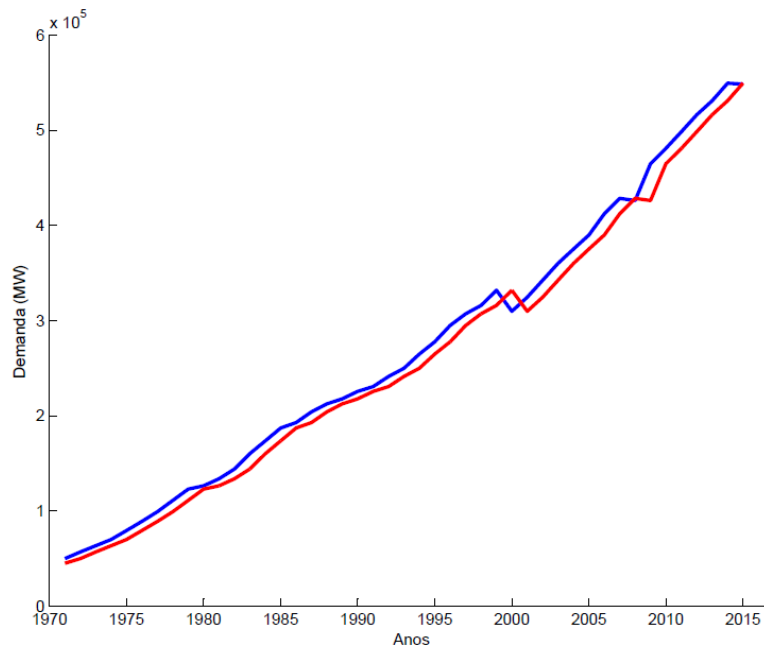


Figura 6.3: Previsão pela modelagem naïve

6.1 Resultados para o modelo RNA

Inicialmente, a seleção dos dados de entrada para a RNA foi feita com base na autocorrelação da série. Para obter um valor de entrada diversificado, considera-se os atrasos correlacionados dentro de um intervalo de confiança de 95%, comumente usado na estatística. Nesse trabalho, considerou-se 93,01%, o que possibilitou o uso de dois atrasos e não apenas 1.

Como pode ser visto na Figura 6.4, os dois atrasos com correlação acima de 93,01% e que se aproximam da linha limite de confiança são os atrasos 7 e 8.

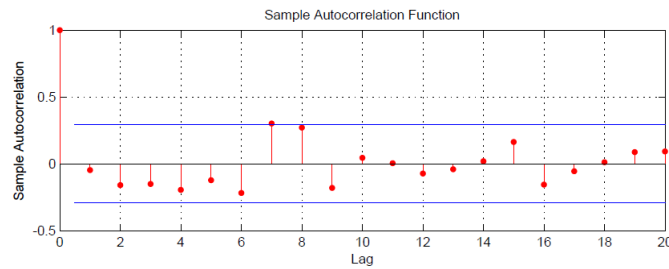


Figura 6.4: Gráfico de autocorrelação

As séries de demanda de energia autocorrelacionadas foram separadas em três categorias e normalizadas entre -1 e 1. Essas categorias incluem os dados utilizados para o treinamento da rede, a categoria de dados de validação e a categoria de dados de teste. Os dados de validação foram utilizados para verificar a eficiência e capacidade de generalização durante o treinamento e os dados de teste, para verificar o desempenho quando submetidos dados não conhecidos pela rede.

O conjunto de treinamento foi composto por 70% da série, sendo que os 30% restantes foram destinados à validação e teste. Esses conjuntos não foram selecionados em sua porção total na série. Cada um deles ainda foi subdividido em outros 3 subconjuntos intercalados, como mostra a Figura 6.5.

Para comparação de resultados, a seleção dos dados de entrada também considerou os primeiros 3 intervalos de atraso da série original, e também a combinação desses atrasos com as entradas com base na autocorrelação, como pode ser visto nas Figuras 6.6 e 6.7. Esses 3 primeiros atrasos foram utilizados com base no conceito naïve.

Com o treinamento da rede multiobjetivo, assim com dito anteriormente, é feito o controle da complexidade das redes através da minimização simultânea do erro para os padrões de treinamento e da norma do vetor de pesos. A Figura 6.8 mostra a solução escolhida na curva Pareto conforme decisor pelo menor erro do conjunto de validação.

Nos outros experimentos, as soluções escolhidas mostradas na Figura 6.9 e 6.10 são para as entradas usando os 3 primeiros atrasos da série e também a combinação dessas

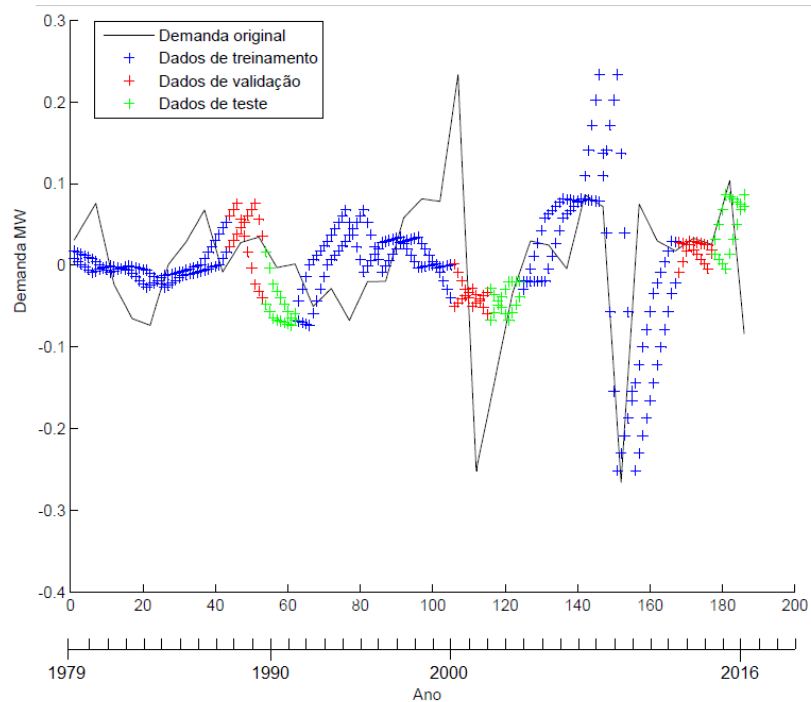


Figura 6.5: Divisão de dados com atrasos mais correlacionados com a série de demanda

entradas com os atrasos mais correlacionados.

Os parâmetros do algoritmo multiobjetivo foram:

- 20 = número de neurônios na camada escondida
- 1 = passo do funcional Norma
- 40 = tamanho máximo para o funcional Norma
- 6000 = número máximo de iterações

É usado um número elevado de neurônios para que o processo de otimização possa variar do menos complexo até o mais complexo. Por esse motivo, foram escolhidos 20 neurônios para a camada escondida. Através dos testes foi observado que, para os dados utilizados neste trabalho, houve necessidade de normas mais elevadas, como 40, para

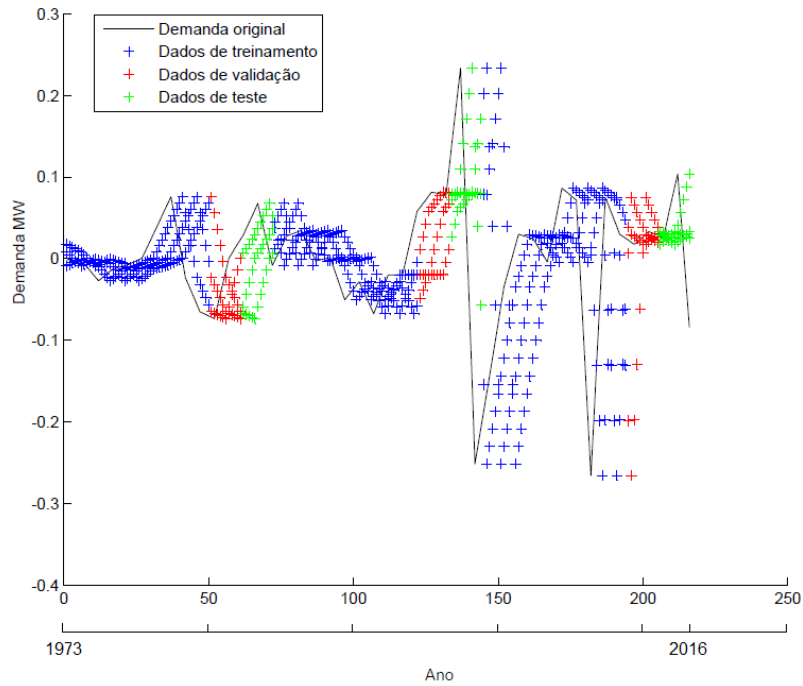


Figura 6.6: Divisão de dados dos 3 primeiros atrasos da série de demanda

conseguir encontrar soluções que se mais se aproximassem dos dados reais. Para gerar uma curva pareto que abrangesse 40 soluções, foi utilizado um passo de 1.

A resposta da MLP nos 3 experimentos feitos pode ser visualizada nos gráficos das Figuras 6.11, 6.12, 6.13 onde as mesmas são comparadas com a série original estacionária e não estacionária.

Para previsão com as séries não estacionárias, foi feita a extrapolação da tendência da série original, a qual foi somada à série estacionária.

Os resultados obtidos nesses experimentos podem ser observado na Tabela 6.1, onde:

- Teste 1 é o teste feito com os dois atrasos como entradas de maior autocorrelação. $D(t - 8)$ e $D(t - 9)$, sendo D a demanda de energia.
- Teste 2 é o teste feito com os três primeiros atrasos como entradas. $D(t - 1)$, $D(t - 2)$ e $D(t - 3)$, sendo D a demanda de energia.

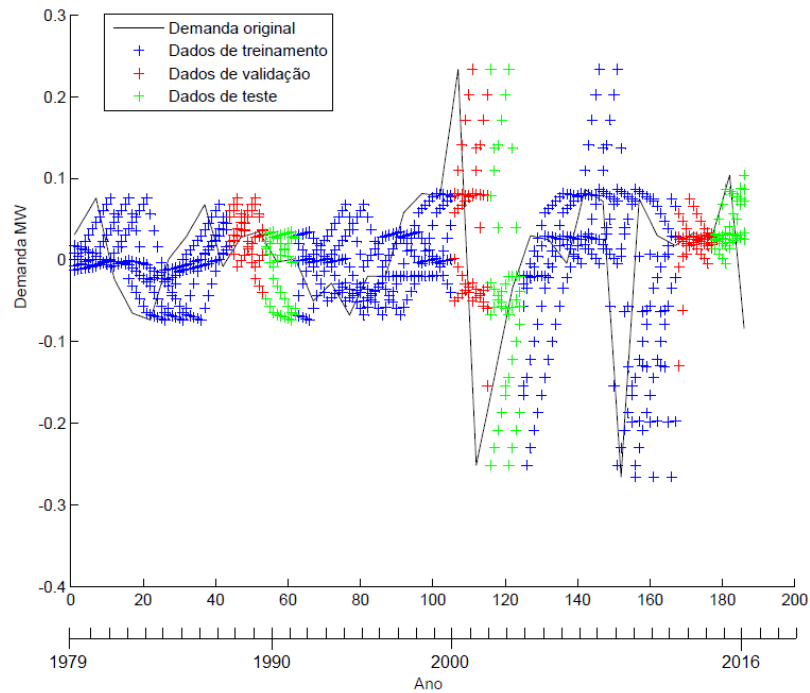


Figura 6.7: Divisão de dados correlacionados e os 3 primeiros atrasos da série de demanda

- Teste 3 é o teste feito com os dois atrasos como entradas de maior autocorrelação e também com os três primeiros atrasos.

6.1.1 Modelo RNA com seleção de variáveis

A metodologia usada na seleção com variáveis exógenas selecionadas baseia-se no uso de uma rede neural híbrida, composta pelo teste com método forward e backward para seleção de variáveis de entrada a partir de um ranking feito pelo algoritmo RRelief-F. A arquitetura da RNA não se altera na suas funções já estabelecidas nesse trabalho.

Usando o RRelief-F com as 32 variáveis, o ranqueamento ilustrado pela Figura 6.14 foi feito na ordem estabelecida a seguir:

1. Índice Setorial - Transporte

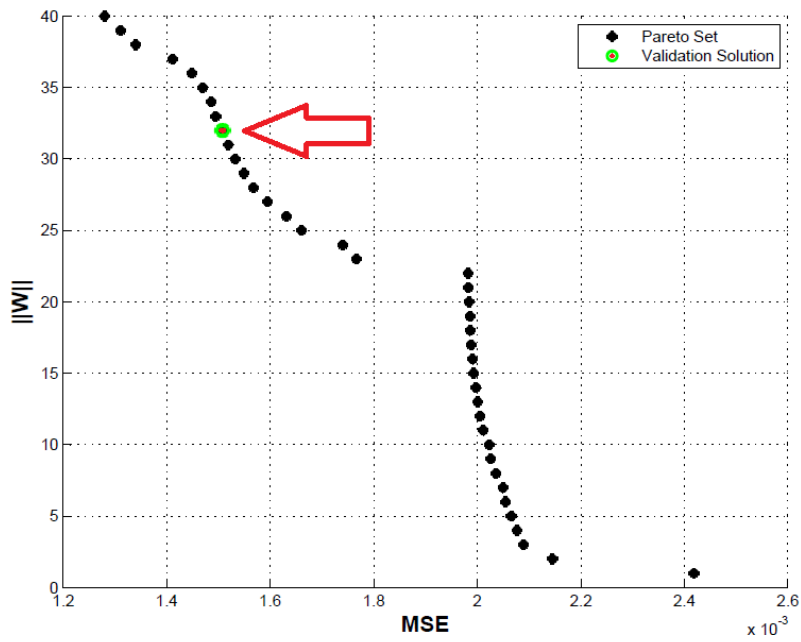


Figura 6.8: Curva com as soluções geradas para os dados correlacionados

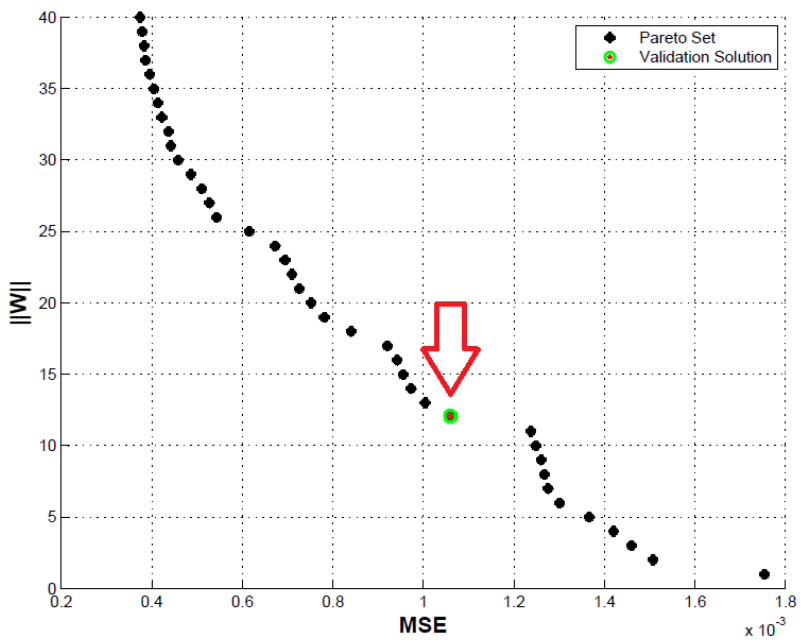


Figura 6.9: Curva com as soluções geradas para os 3 primeiros atrasos da série de demanda

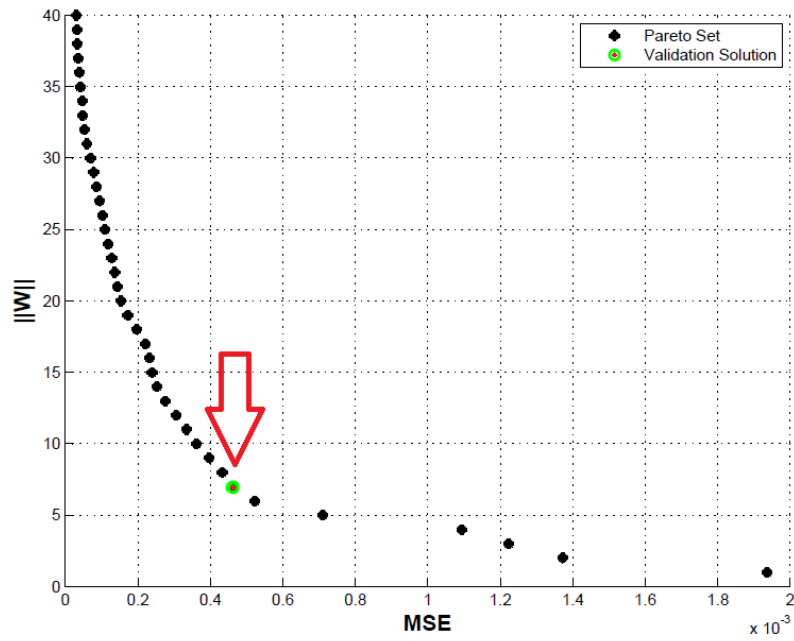


Figura 6.10: Curva com as soluções geradas para dados correlacionados e os 3 primeiros atrasos da série de demanda

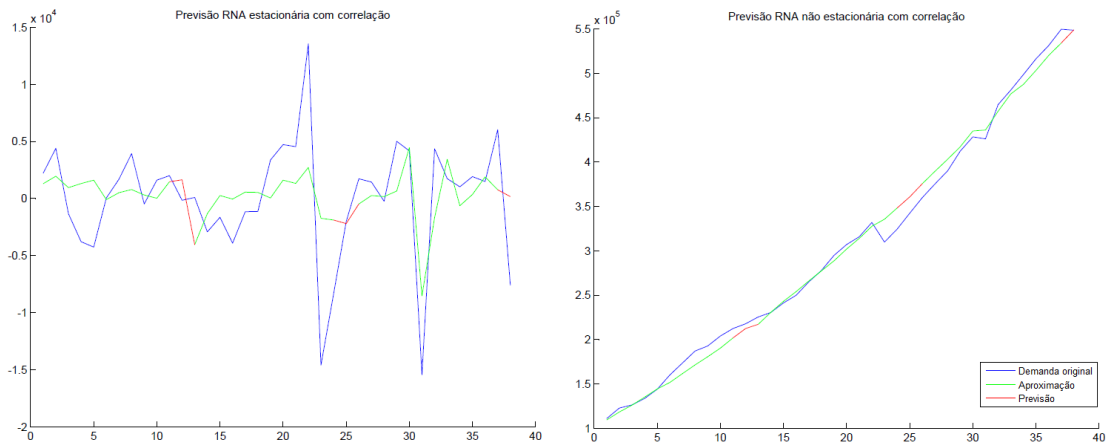


Figura 6.11: Solução gerada pela MLP para os dados mais correlacionados comparada com a série original

2. Balanço Energético Energia Final
3. PIB anual
4. Capacidade industrial - Geral

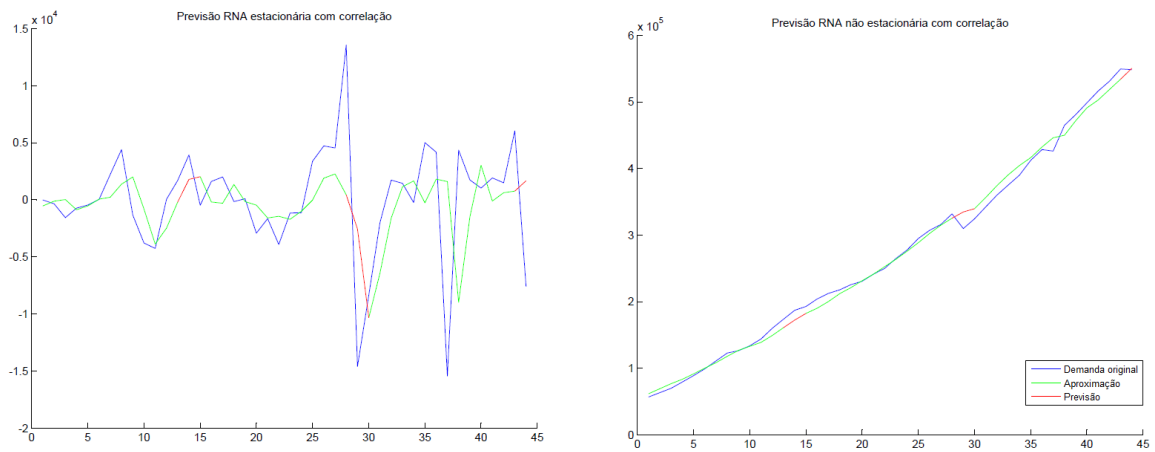


Figura 6.12: Solução gerada pela MLP com os 3 primeiros atrasos comparada com a série original

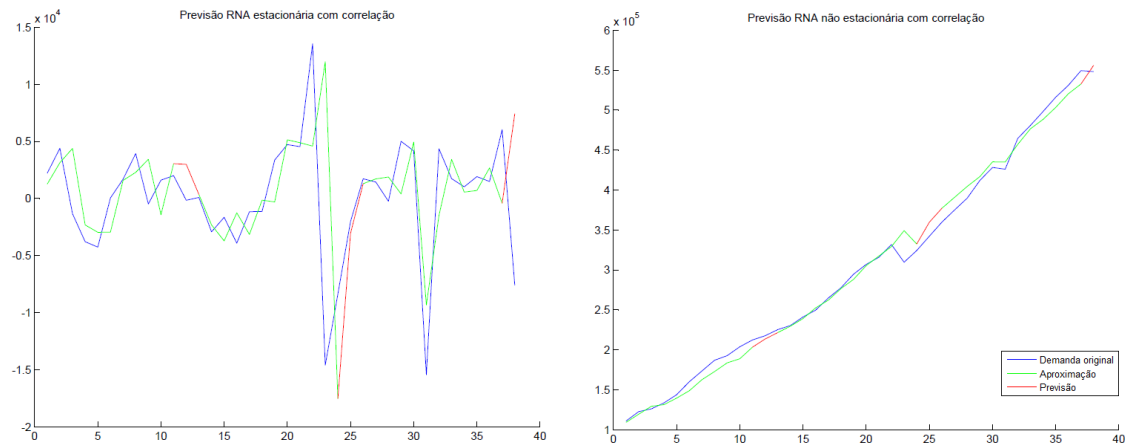


Figura 6.13: Solução gerada pela MLP com os 3 primeiros atrasos e os atrasos com maior correlação comparada com a série original

5. Índice Setorial - Industrial
6. Índice Setorial - Público
7. Índice Setorial - Residencial
8. Índice Setorial - Agropecuário
9. População

Tabela 6.1: Resultados para RNA

Série	Avaliação	Divisão	RNA		
			Teste 1	Teste 2	Teste 3
Estacionária	RMSE (MW)	Geral	4413,7231	5108,8748	5987,9079
		Previsão	4495,2735	8015,9793	7225,2624
	Correlação	Geral	0,5451	0,0412	0,3003
		Previsão	0,0908	0,0162	-0,4452
Não Estacionária	RMSE (MW)	Geral	10550,9371	10657,0772	10821,4209
		Previsão	12779,2423	16474,3131	10961,9692
	MAPE (%)	Geral	2,9545	3,5153	2,8810
		Previsão	3,0209	5,4586	2,6173

10. Índice Setorial - Comercial
11. SELIC - Meta
12. Valor tarifário - Rural irrigante
13. Valor tarifário - Rural
14. Valor tarifário - Industrial
15. Clima - Velocidade do vento
16. Valor tarifário - Iluminação pública
17. SELIC - Taxa
18. Valor tarifário - Saneamento
19. Clima - Precipitação
20. Valor tarifário - Próprio
21. Valor tarifário - Poder público
22. Valor tarifário - Residencial

23. Valor tarifário - Comercial
24. Clima - Direção do vento
25. Clima - Nebulosidade
26. Clima - Temperatura
27. Porcentagem de inflação
28. Capacidade de geração
29. Cotação do dólar
30. Clima - Umidade relativa do ar
31. Salário mínimo
32. Clima - Pressão

Depois que as variáveis foram normalizadas, para que as séries analisadas possuíssem as mesmas condições de estacionariedade, foi feita a extração da tendência usando o filtro HP.

Forward

Inicialmente, todas as variáveis de entrada são avaliadas usando o método forward e as variáveis mais correlacionadas ou com menor MSE são utilizadas na fase de treinamento da rede neural. O diagrama da Figura 6.15 descreve o processo da rede neural híbrida.

A saída com método forward passando pelo modelo selecionou as seguintes variáveis independentes:

1. Balanço Energético Energia Final
2. PIB anual

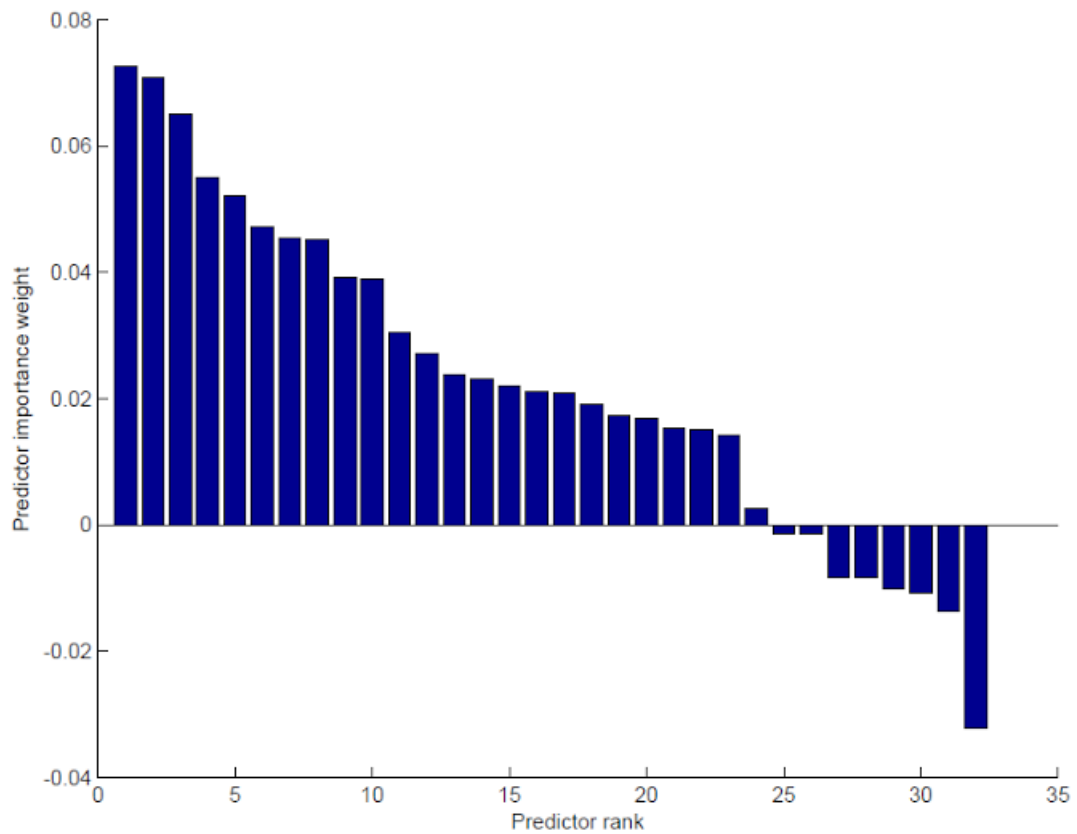


Figura 6.14: Ranking de variáveis independentes

3. Capacidade industrial - Geral
4. Índice setorial - Industrial
5. Índice setorial - Agropecuário
6. População
7. SELIC - Meta

Os resultados obtidos nesses experimentos podem ser observado na Tabela 6.2.

A resposta do experimento feito pode ser visualizada no gráfico das Figuras 6.16 onde são comparadas com a série original estacionária e não estacionária.

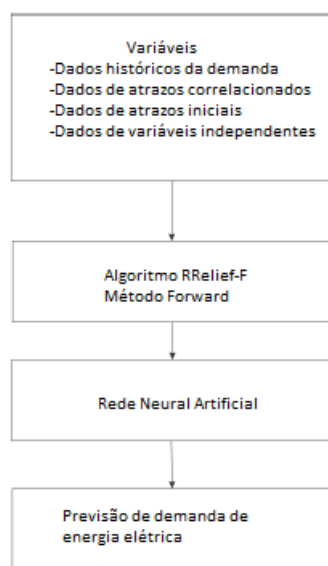


Figura 6.15: Rede Neural Híbrida Forward

Tabela 6.2: Resultados RNA seleção de variáveis com Forward

Série	Avaliação	Divisão	Resultado
Estacionária	RMSE (WM)	Geral	5181,3515
		Previsão	4109,0728
	Correlação	Geral	0,2106
		Previsão	-0,5249
Não estacionária	RMSE (MW)	Geral	11180,236
		Previsão	11732,373
	MAPE(%)	Geral	3,246
		Previsão	3,146

Backward

Todas as variáveis de entrada são submetidas ao modelo usando o método Backward e a saída mais correlacionada ou com menor MSE, com as variáveis avaliadas, são utilizadas na fase de aprendizado da rede neural. O diagrama da Figuras 6.17 descreve o processo da rede neural híbrida.

A saída com método Backward passando pelo modelo selecionou as seguintes variáveis independentes:

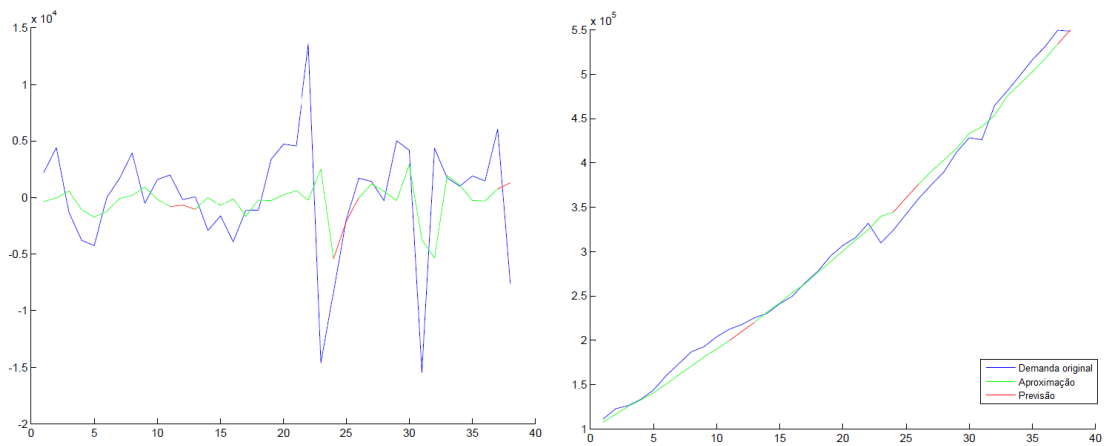


Figura 6.16: Gráfico de aproximação da RNA por seleção de variável com Forward)

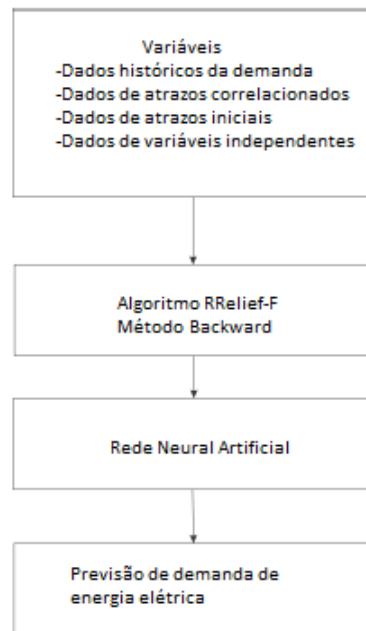


Figura 6.17: Rede Neural Híbrida Backward

1. Clima - Umidade relativa do ar
2. Valor tarifário - Residencial
3. Valor tarifário - Poder público

Tabela 6.3: Resultados RNA seleção de variáveis com Backward

Série	Avaliação	Divisão	Resultado
Estacionária	RMSE (WM)	Geral	5251,9943
		Previsão	4634,6182
	Correlação	Geral	0,2103
		Previsão	-0,5269
Não estacionária	RMSE (MW)	Geral	11095,647
		Previsão	12309,791
	MAPE(%)	Geral	3,0005
		Previsão	3,0781

4. Clima - Precipitação
5. Valor tarifário - Rural irrigante
6. SELIC - Meta
7. Índice setorial - Residencial

Os resultados obtidos nesses experimentos podem ser observado na Tabela 6.3.

A resposta do experimento feito pode ser visualizada no gráfico das Figuras 6.18 onde são comparadas com a série original estacionária e não estacionária.

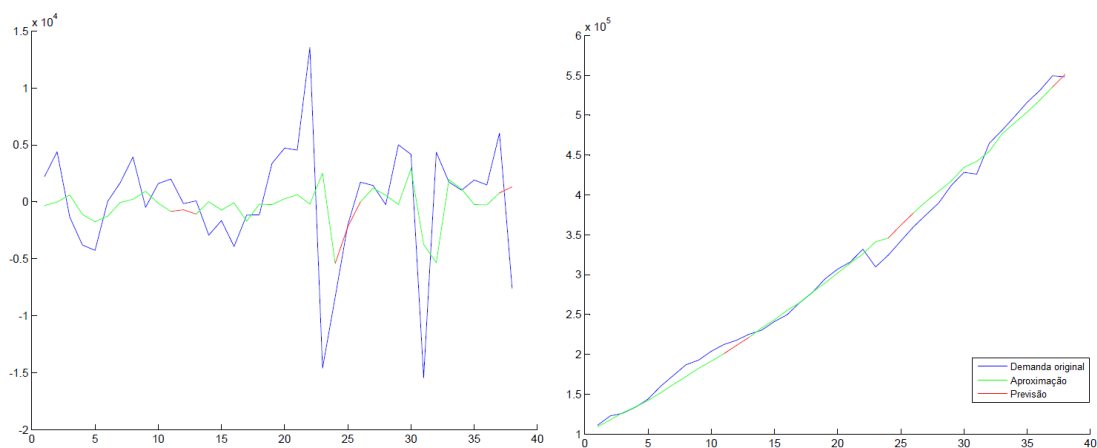


Figura 6.18: Gráfico de aproximação da RNA por seleção de variável com Backward)

6.1.2 Seleção de variáveis com Mínimos Quadrados

Foi feito um experimento em que as variáveis foram selecionadas pelos métodos forward e backward, submetendo-as a uma regressão linear.

Para encontrar a equação de regressão linear, foi utilizado o Método dos Mínimos Quadrados. Esse método define uma solução que minimiza a soma das distâncias ao quadrado entre os pontos. Ele encontra uma função de aproximação dos dados observados ([Ruggiero and Lopes, 1996](#)).

Forward

Inicialmente, todas as variáveis de entrada são avaliadas usando o método forward da mesma forma que no experimento que as envia pelo modelo, porém nesse primeiro momento, o modelo é substituído pelo Método dos Mínimos Quadrados(MMQ).

As variáveis mais correlacionadas ou com menor MSE conforme saída do método forward com MMQ são utilizadas na fase de aprendizado da rede neural. Na Figura 6.19 é apresentado o fluxograma do método forward no uso do MMQ. A arquitetura da rede neural híbrida permanece a mesma descrita anteriormente.

A saída com método Forward passando por MMQ selecionou as oito variáveis independentes a serem introduzidas no modelo:

1. Índice setorial - Transporte
2. PIB anual
3. Índice setorial - Industrial
4. Índice setorial - Público
5. Índice setorial - Residencial
6. Índice setorial - Agropecuário

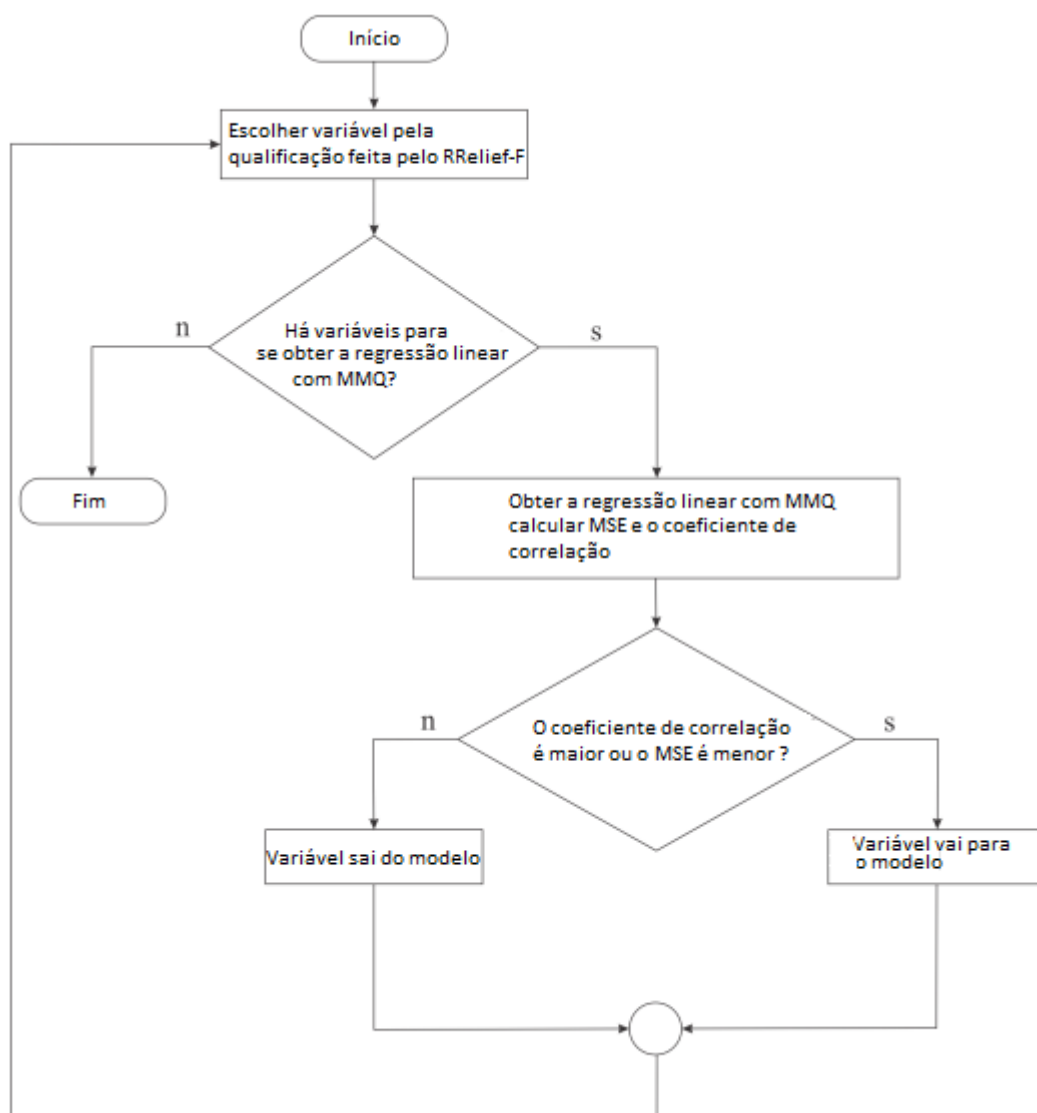


Figura 6.19: Fluxograma Forward MMQ

7. População

8. Índice setorial - Comercial

Os resultados obtidos nesses experimentos podem ser observado na Tabela 6.4.

A resposta do experimento feito pode ser visualizada no gráfico das Figuras 6.20 onde são comparadas com a série original estacionária e não estacionária.

Tabela 6.4: Resultados RNA seleção de variáveis com Forward MMQ

Série	Avaliação	Divisão	Resultado
Estacionária	RMSE (WM)	Geral	2903,007
		Previsão	2663,628
	Correlação	Geral	0,8368
		Previsão	0,6885
Não estacionária	RMSE (MW)	Geral	9672,44
		Previsão	11436,17
	MAPE(%)	Geral	2,7471
		Previsão	2,8958

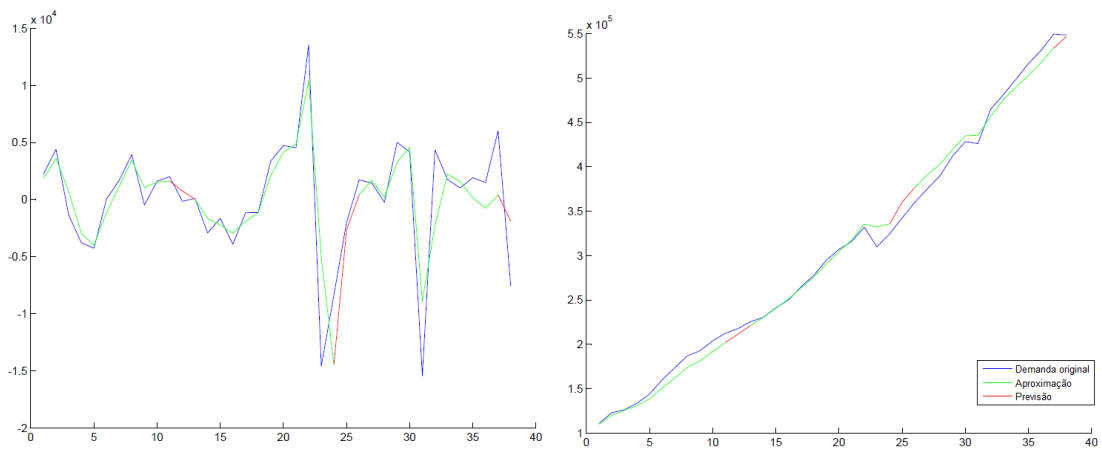


Figura 6.20: Gráfico de aproximação da RNA por seleção de variável com Forward MMQ)

Backward

Nesse primeiro momento, comparado ao fluxograma apresentado para o backward, o modelo é substituído pelo MMQ. Todas as variáveis de entrada são submetidas ao uso do MMQ com método Backward e as saídas mais correlacionadas ou com menor MSE são utilizadas na fase de aprendizado da rede neural.

Na Figura 6.21 é apresentado o fluxograma do método backward no uso do MMQ. A arquitetura da rede neural híbrida permanece a mesma descrita anteriormente.

A saída com método Backward passando pelo MMQ selecionou as seguintes variáveis independentes:

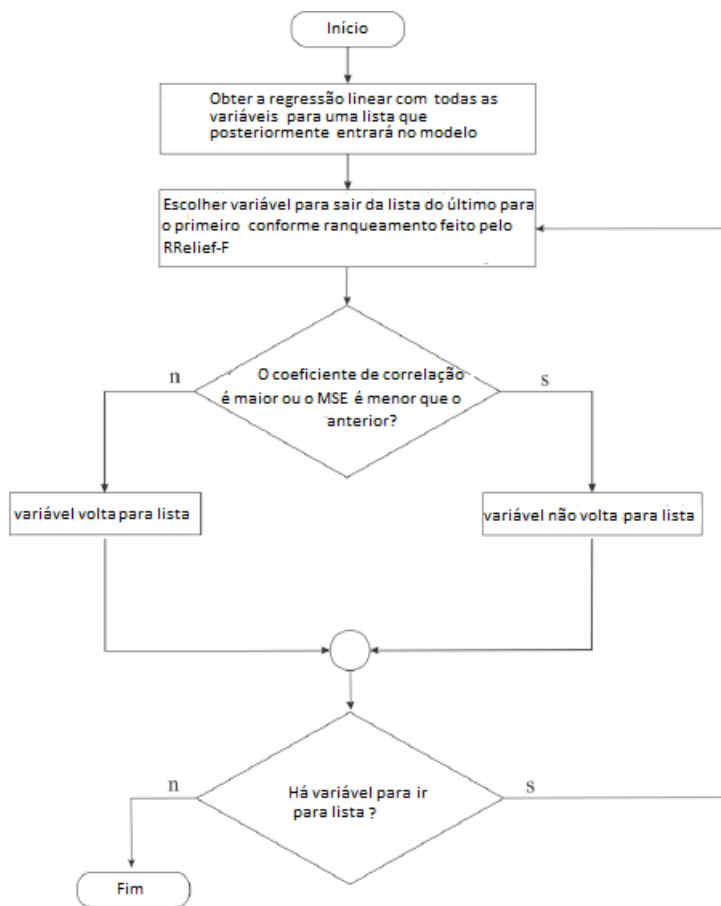


Figura 6.21: Fluxograma Backward MMQ

1. Clima - Pressão
2. Salário mínimo
3. Clima - Umidade relativa do ar
4. Cotação do dólar
5. Capacidade de geração
6. Porcentagem de inflação
7. Clima - Nebulosidade

Tabela 6.5: Resultados RNA seleção de variáveis com Backward MMQ

Série	Avaliação	Divisão	Resultado
Estacionária	RMSE (WM)	Geral	4881,927
		Previsão	3964,045
	Correlação	Geral	0,3758
		Previsão	-0,2541
Não estacionária	RMSE (MW)	Geral	11045,86
		Previsão	12377,71
	MAPE(%)	Geral	3,213
		Previsão	3,2338

8. Clima - Direção do vento

9. Valor tarifário - Comercial

10. Valor tarifário - Residencial

Os resultados obtidos nesses experimentos podem ser observado na Tabela 6.5.

A resposta do experimento feito pode ser visualizada no gráfico das Figuras 6.22 onde são comparadas com a série original estacionária e não estacionária.

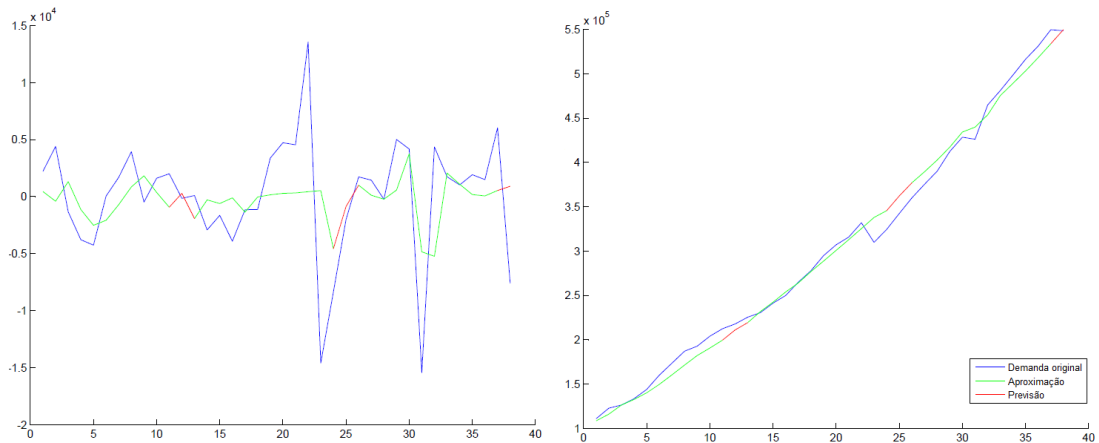


Figura 6.22: Gráfico de aproximação da RNA por seleção de variável com Backward MMQ

6.2 Resultados para o modelo com série de Fourier

Para o uso do modelo da série de Fourier, a série original foi dividida em 3 partes. Para cada parte, considerou-se o intervalo de demanda de energia de 85%, e esses dados foram os utilizados para o primeiro ajuste dos coeficientes. Os outros 15% foram usados para previsão usando coeficientes ajustados fixando os mesmos sem fazer a atualização.

A aproximação da série original usando a análise harmônica da série de Fourier foi testada, com a abordagem citada anteriormente, com coeficientes fixos, e os resultados apresentados podem ser vistos na Figura 6.23.

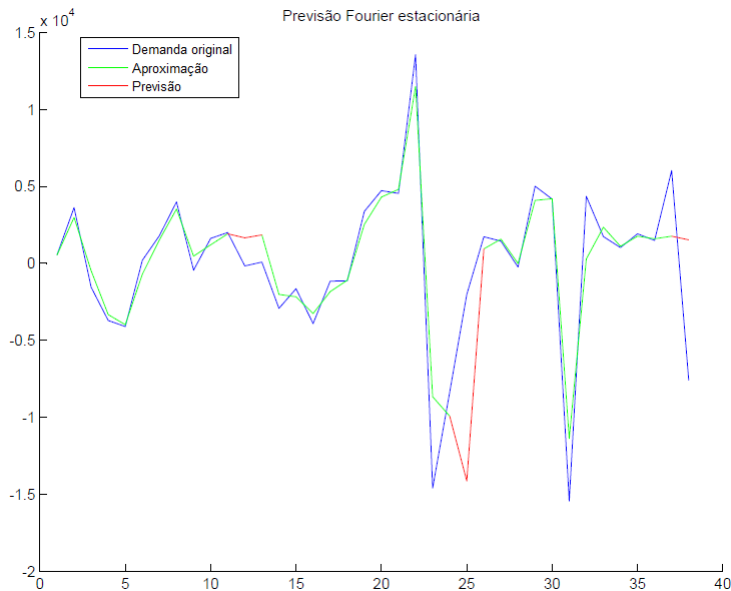


Figura 6.23: Gráfico de aproximação por série de Fourier (série estacionária)

A previsão foi feita fixando os coeficientes após eles serem ajustados a partir das equações (4.11), (4.12) e (4.13), e depois foram aplicados à equação (4.2), citadas nos capítulos anteriores.

A Figura 6.24 mostra o resultado do modelo aplicando os dados de tendência extrapolada.

As métricas podem ser observados na Tabela 6.6. Nessa tabela, são mostrados os



Figura 6.24: Gráfico de aproximação por série de Fourier (série não estacionária)

Tabela 6.6: Resultados para série de Fourier

Série	Avaliação	Divisão	Resultado
Estacionária	RMSE (WM)	Geral	2985,379
		Previsão	6888,096
	Correlação	Geral	0,8236
		Previsão	0,0531
Não estacionária	RMSE (MW)	Geral	8095,084
		Previsão	6596,105
	MAPE(%)	Geral	2,8603
		Previsão	1,9133

resultados para a série estacionária e não estacionária para os dados de aproximação de forma geral e para comparação dos dados que foram previstos separadamente.

6.3 Análise de resultados

Essa sessão trata da avaliação dos resultados quanto aos critérios de aproximação com os dados históricos da demanda de energia elétrica do SIN. As técnicas aplicadas para previsão dessa demanda são comparadas entre si, sendo possível uma análise de quais

tiveram resultados mais promissores.

São discutidos os resultados para solução adotada em todas as etapas de processamento dos dados, desde sua extração e tratamento até a passagem pelo modelo com a geração da saída.

Os resultados foram colocados na Tabela 6.7. Sua indicação em vermelho é para o pior resultado e em verde para o melhor resultado, quando feito um paralelo entre os experimentos por categoria de avaliação.

As separação dos resultados foram entre as séries estacionaria, não estacionária, as métricas de avaliação, a divisão geral e a divisão de previsão. Na divisão geral, foram considerados os resultados comparados com todas as faixas da série, incluindo dados de treinamento, validação e teste. Na divisão de previsão, foi considerada a faixa de teste para comparação de resultados.

6.3.1 Análise por série de Fourier

A vantagem da série de Fourier frente a RNA é a não necessidade de um estudo detalhado das variáveis e uma melhor arquitetura para a obtenção de resultados confiáveis.

([Brentan et al., 2015](#)) diz que podem haver casos em que não exista uma correlação real entre as variáveis exógenas e a série a ser prevista pelas RNAs. A resposta obtida pode ser correta do ponto de vista matemático, mas sem real ligação com o problema físico.

As vantagens da análise por série de Fourier podem ser comprovadas pela sua implementação nesse trabalho. O modelo usa somente os dados do próprio histórico de demanda do SIN.

A aproximação da série histórica com o modelo para os dados de treinamento foram superiores ao método ingênuo em todos os resultados da série não estacionária. Esse comportamento era esperado, já que essa aproximação usa senos e cossenos baseados nos dados originais a serem aproximados.

Como a técnica de aproximação da série de Fourier é periódica, a previsão feita infere um período. O modelo tenta repetir dados do passado em passos futuros assim como o método ingênuo, respeitando a periodicidade. Como a série histórica estacionária não possui essa periodicidade definida, o método não foi superior ao naïve nessa avaliação.

6.3.2 Análise do modelo RNA

Para resolver o problema de casos em que não exista uma correlação real entre variáveis exógenas, são usadas as técnicas de qualificação e filtragem de variáveis. Com esse problema resolvido, o modelo RNA acaba por ter vantagem frente ao modelo de análise por série de Fourier.

Para o modelo RNA foram feitos 7 experimentos que terão seus resultados discutidos a seguir. Três desses experimentos utilizaram apenas variáveis endógenas como entrada.

Entradas correlacionadas com a série histórica

A escolha das entradas mais correlacionadas com a série histórica de demanda trouxe um problema que poderia prejudicar o experimento frente aos outros. O problema da curta faixa de valores para serem trabalhados motivou o uso de interpolação de valores da série.

A série histórica no SIN, utilizada no trabalho, é anual e se limita a 46 registros, sendo 1 registro de frequência de demanda por ano. Quando escolhidos os atrasos mais correlacionados $D(t - 8)$ e $D(t - 9)$, o limite de valores cai para 38.

Considerando a resolução do problema de curta faixa de valores, com a interpolação de dados, e que os atrasos são os mais correlacionados com a série, os resultados foram superiores ao método ingênuo. Pode ser observado que os valores do experimento não foram superiores somente quando comparados à correlação.

Por serem dados de $D(t - 1)$ no modelo de previsão naïve, a correlação com a série original é realmente alta.

Entradas dos três atrasos iniciais da série histórica

Quando considerada que a previsão mais óbvia e simples é a previsão naïve, pode se pressupor que os atrasos mais próximos para série histórica de demanda sejam os mais óbvios para entrada e obtenção de bons resultados. Nesse sentido, foram escolhidos os 3 primeiros atrasos da série $D(t-1)$, $D(t-2)$ e $D(t-3)$ como entrada do modelo RNA.

A série histórica no SIN, utilizada no trabalho, é anual e se limita a 46 registros, sendo cada registro de frequência de demanda por ano. Quando foram escolhidos os 3 atrasos, o limite de valores caiu para 45 registros.

Trabalhar com 45 registros usando interpolação de dados não foi um problema. O que fez os resultados serem os piores, para RMSE da previsão das séries estacionárias, não estacionárias e a correlação da previsão, foram os atrasos com uma baixa correlação com a série histórica de demanda. Mesmo assim, os atrasos foram utilizados como entrada do modelo.

Entradas correlacionadas combinadas aos atrasos iniciais

Os 3 primeiros atrasos da série têm uma correlação baixa com a série histórica de demanda de modo geral, mas quando combinada com os atrasos mais correlacionados, alguns resultados acabam sendo melhores do que os resultados dos experimentos originais.

O RMSE geral da série estacionária foi o pior resultado, porém dados como MAPE% e RMSE da previsão, para série não estacionária, foram os melhores dentre os 3 experimentos.

Entradas selecionadas pelo método Forward

Para os 3 experimentos anteriores usando variáveis endógenas e as variáveis exógenas selecionadas com o método forward, observou-se que a série histórica de demanda teve

seus valores extremos todos condicionados nas variáveis de entrada. Graficamente, todos os picos e vales da série foram alcançados pelas variáveis.

Optou-se por manter as entradas endógenas dos experimentos anteriores e usar o método forward para selecionar as variáveis exógenas.

O experimento utilizando o método forward, obteve resultados parecidos com o experimento que utilizou as entradas mais correlacionadas com a série histórica.

O método ingênuo obteve melhores resultados na correlação da série não estacionária, mas em outras avaliações seus resultados foram inferiores.

Entradas selecionadas pelo método Backward

Do mesmo modo que no experimento com método forward, optou-se por manter as entradas endógenas dos experimentos anteriores e usar o método backward para selecionar as variáveis exógenas.

Os resultados não foram tão diferentes dos resultados utilizando o método forward. Até mesmo a correlação geral da série não estacionária teve uma avaliação equivalente

Entradas selecionadas pelo método Forward MMQ

Preparar as variáveis exógenas antes de introduzi-las ao modelo com método forward com mínimos quadrados, além de reduzir o tempo de processamento do modelo consideravelmente, produz bons resultados.

As avaliações com o método forward com mínimos quadrados foi o melhor dentre os experimentos feitos. Não utilizar o método forward passando pelo modelo e tratar as variáveis com método de mínimos quadrados por fora da RNA, desconsideram os componentes aleatórios que podem piorar os resultados.

Mesmo as avaliações com o método Forward MMQ que não foram as melhores dentre os experimentos, chegaram próximas ou até superiores às avaliações do naïve.

Entradas selecionadas pelo método Backward MMQ

O método de seleção das entradas com backward e mínimos quadrados usa do mesmo que foi feito no experimento com Forward MMQ. A diferença está somente no algoritmo de seleção.

Os resultados não foram ruins, podendo ser comparáveis aos resultados obtidos com o experimento que utilizou os atrasos mais correlacionados.

Tabela 6.7: Comparação de resultados

verde = melhor resultado
 vermelho = pior resultado

Série	Avaliação	Divisão	Naive	Fourier	Entradas com atrasos e correlações				Seleção de variáveis			
					Entradas correlacionadas	Primeiros atrasos	Correlacionadas e primeiros atrasos	Forward	Backward	Forward MMQ	Backward MMQ	
Estacionária	RMSF (WM)	Geral	7592,5027	2985,379	4413,7231	5108,8748	5987,9079	5181,3515	5251,9943	2903,007	4881,927	
		Previsão	5614,7795	6888,096	4495,2735	8015,9793	7225,2624	4109,0728	4634,6182	2663,628	3964,045	
	Correlação	Geral	-0,0536	0,8236	0,5451	0,0412	0,3003	0,2106	0,2103	0,8368	0,3758	
Não estacionária	RMSF (MW)	Geral	14808,1139	8095,084	10550,9371	10667,0772	-0,4452	11180,236	11095,647	0,6885	11045,86	
		Previsão	12828,2739	6596,105	12779,2423	16474,3131	10961,9692	11732,373	12309,791	9672,44	12377,71	
	MAPE (%)	Geral	4,8557	2,8603	2,9545	3,5153	2,8810	3,246	3,0005	2,7471	3,213	
	Previsão	3,6232	1,9133	3,0209	5,4586	2,6173	3,146	3,0781	2,8958	3,2338		

Capítulo 7

Conclusão

Nesse trabalho, foi realizada uma análise do problema de previsão de demanda de energia elétrica do Brasil. Essa análise de comportamento futuro é fundamental para que as geradoras, distribuidoras e comerciantes de energia possam gerenciar as incertezas do mercado.

Os fatores de risco, associados ao mercado de energia elétrica, atuam de forma decisiva na qualidade da previsão de demanda. A variabilidade desse mercado tem grande importância no resultado final.

Considerando a variação do mercado de energia elétrica, realizou-se o estudo das variáveis que influenciam na demanda de energia. Foram buscadas as variáveis que pudessem estar relacionadas ao histórico de demanda e que tivessem potencial de melhorar os resultados.

O comportamento da série de demanda, sensível ao mercado de energia elétrica, é observado nos resultados obtidos com os métodos que consideraram as variáveis exógenas. É necessário armazenar essas informações para utilizar em observações futuras.

O diferencial desse trabalho, quanto às informações coletadas e enviadas para a base de dados estruturada, está na quantidade de variáveis exógenas. O repositório de dados foi carregado com 32 variáveis, o que torna a base diversificada.

Outro diferencial desse trabalho é a forma como os dados são guardados e gerenciados. Um sistema próprio foi criado para auxiliar na administração da informação. Assim, a utilização das variáveis pode extrapolar a especificidade da série de demanda de energia elétrica. Os relatórios gerados podem ser utilizados para outras observações em outros trabalhos, dependendo da necessidade.

Nesse trabalho, a informação foi gerenciada e a quantidade de variáveis direcionou a previsão de demanda sobre as mudanças que o mercado de energia elétrica enfrenta. A limitação dessa abordagem está na dificuldade de mapear as modificações substanciais nesse mercado. Quando regras mudam abruptamente, como ocorreu com o racionamento de energia em 2001, acompanhar o comportamento da demanda torna-se complexo.

As projeções de demanda de energia elétrica no mercado brasileiro devem ser assertivas. Os valores obtidos utilizando o filtro HP identificaram a existência do comportamento de tendência e comportamento cíclico para a série histórica do SIN. Nesse trabalho, a extração desses componentes aumentou a possibilidade de acerto nessas projeções, além de melhorar a qualificação da influência das variáveis endógenas.

A qualificação da influência dos atrasos da série histórica da demanda de energia, obtida utilizando autocorrelação, cumpriu com o esperado na manipulação de variáveis de entrada endógenas. Os resultados do modelo RNA, utilizando esses atrasos como entrada, mostram uma previsão que acompanha as mudanças na série histórica.

Claramente no trabalho, a previsão da demanda de energia elétrica com o modelo RNA tem uma aproximação mais assertiva quando as entradas são as variáveis endógenas em conjunto às variáveis exógenas. Com a entrada de informações externas no modelo, pode-se mapear o comportamento do ambiente. Esse ambiente, assim como a previsão da própria demanda, é conseqüentemente previsto no processo.

Mesmo havendo dependência do comportamento do ambiente pela série histórica de demanda, foi verificado que nem todas as variáveis exógenas combinadas e usadas como

entrada do modelo trazem benefícios. Nos experimentos, foi verificado que melhores resultados são garantidos com a qualificação das variáveis, para posteriormente fazer uma seleção.

A utilização do algoritmo RRelief-F para qualificação das variáveis, sendo uma técnica recente que cumpre com o esperado, agrega valor aos resultados e reforça a ideia que os algoritmos da família Relief podem ser aplicados em problemas diversificados.

Fazer um ranking de dados de entrada para o modelo RNA foi a primeira etapa da seleção de variáveis. Em seguida, a etapa de filtragem dessas informações, fazendo experimentos com método forward e backward, garantiu que os resultados fossem melhorados.

O modelo RNA é eficiente para previsão de demanda de energia elétrica considerando o cenário brasileiro. O modelo proposto no trabalho, combinado às informações subjetivas obtidas, gerou previsões consistentes que podem auxiliar na tomada de decisão das empresas prestadoras de serviço de energia.

Não apenas o modelo RNA, como também o modelo por análise da série de Fourier, têm precisão em suas previsões de demanda de energia elétrica. As empresas necessitam dos resultados desse tipo de modelo.

O modelo com análise por série de Fourier se baseia unicamente na informação histórica. Essa independência de informações externas para realizar a previsão é uma vantagem, pois simplifica o modelo.

Havendo ou não permissão de inclusão de variáveis exógenas para previsão, a estabilidade dos modelos está ligada à forma como eles são empregados na resolução do problema. Os modelos em questão, utilizados no trabalho, foram adequadamente configurados para o cenário de demanda de energia elétrica brasileiro que se desejava prever.

Dentre os modelos, o recomendável para as empresas prestadoras de serviço de energia elétrica é aquele que obteve os melhores resultados. O modelo recomendado

seria o RNA, que utiliza uma seleção de entradas forward com mínimos quadrados, mas não podem ser descartadas as possibilidades com a análise por série de Fourier e os outros experimentos. Os modelos ainda podem ser aprimorados em trabalhos futuros.

7.1 Trabalhos Futuros

Para trabalhos futuros, o modelo por série de Fourier e RNA podem ser modificados para gerar melhores resultados. A série histórica de demanda de energia elétrica e as suas variáveis de influência também podem ser alterados com o mesmo objetivo. Nesse sentido, é importante verificar e estender a base de dados para mantê-la sempre atualizada e detalhada.

Espera-se deixar a previsão detalhada utilizando uma série mensal ou até mesmo diária. Com o uso de uma série com essas características, poderia ser testado o método da série de Fourier com a atualização dos coeficientes. Esse teste depende do formato da série, já que a demanda anual não se adaptou ao modelo com essa abordagem.

Trabalhar com métodos que vão além do forward e do backward é algo pretendido para próximos trabalhos que usem modelos RNA com seleção de variáveis. Um método candidato usado na literatura é o stepwise, por ser uma combinação desses dois métodos de seleção.

Pretende-se também fazer experimentos construindo um modelo híbrido entre os modelos que obtiveram melhores resultados: série de Fourier e RNA Forward MMQ.

Referências Bibliográficas

- Alvarez, A. and Luque, B. (2003). *Rede Neural de Kohonen e Outras Técnicas para Treinamento Não-Supervisionado*. Universidade Estadual de Campinas - Faculdade de Engenharia Elétrica e Computação.
- Alves, M. F. (2013). *Previsão de demanda de cargas elétricas por seleção de variáveis stepwise e redes neurais artificiais*. Universidade Estadual Paulista - Campus Ilha Solteira.
- Azevedo, E. M., Menezes, M. and Araújo, A. L. D. G., Ramos, I., and Reis, S. (2005). *O novo modelo do setor elétrico - avaliação de riscos e impacto financeiro para as distribuidoras*. Seminário Nacional de Produção e Transmissão de Energia Elétrica.
- Batista, A. L. F. (2009). Modelos de série temporais e redes neurais artificiais na previsão de vazão. *Lavras, Minas Gerais - Universidade Federal de Lavras - Programa de Pós Graduação em Engenharia de Sistemas*.
- Box, G. and Jenkins, G. (1970). *Time Series Analysis, Forecasting and Control*. HoldenDay - San Francisco.
- Box, G. and Pierce, D. A. (1970). *Distribution of the residual autocorrelations in autoregressive-integrated moving-average time series models*. Journal of the American Statistical Association, 65, 1509 - 1526.

- Braga, A. P., Lurdemir, T. B., and Carvalho, A. C. P. L. F. (2000). *Redes Neurais Artificiais: Teoria e Aplicações*. LTC - Livros Técnicos e Científicos Editora S.A., Belo Horizonte, Brasil.
- Brentan, B. M. (2015). Previsão de demanda em tempo real usando Série de Fourier Adaptativa. *XIV SEMINARIO IBEROAMERICANO DE REDES DE AGUA Y DRENAJE*.
- Brentan, B. M., Luvizotto, E., Sebastián, J. I., and García, R. P. (2015). *Previsão de demanda em tempo real usando Série de Fourier Adaptativa* . XIV Seminario Iberoamericano de redes de água; FEC-Unicamp, Campinas, São Paulo, Brasil.
- Brockwell, P. and Davis, R. A. (1996). *Introduction to time series and forecasting*. Springer Verlag, New York.
- Cardon, A. and Müller, D. N. (1994). *Introdução às redes neurais artificiais* . Universidade Federal do Rio Grande do Sul - Instituto de Informática - Curso de Pós-Graduação em Ciência da Computação.
- Castro, C. L. d. (2004). *Sistema inteligente para previsão on-line do índice IBOVESPA baseado em redes neurais artificiais*. The University of Amsterdam. Eighth edition. 136p.
- Castro, L. N. d. and Zuben, F. J. V. (2015). *Redes Neurais Artificiais - O Neurônio Genérico em RNAs* . Universidade Estadual de Campinas - Departamento de Engenharia de Computação e Automação Industrial.
- CEMIG (2015). Companhia Energética de Minas Gerais S.A - História da eletricidade no Brasil .
- Chatfield, C. (1996). *The Analysis of Time Series: An introduction*. Chapman and Hall, fifth edition, NY.

- DASH, H. L. M. (1997). *Feature selection for classification*. p. 131-156. Intelligent Data Analysis 1.
- Davila, V. H. L. (2015). Introdução às Séries Temporais . *Universidade Estadual de Campinas, Campinas - São Paulo - Brasil*.
- Demsar, J. (2010). *Algorithms for subsetting attribute values with Relief*. 78:421-428 . Journal of Machine Learning Research.
- Diniz, G. L. (2008). *Análise harmônica do regime de precipitação em duas localidades da baixada cuiabana*. Universidade Federal de Mato Grosso - Instituto de ciências exaas tas e da terra - programa de pós graduação em física e meio ambiente.
- Diniz, H., Andrade, L. C. M., Carvalho, A. C. P. L. F., and Andrade, N. G. (1998). *Previsão de séries temporais utilizando redes neurais artificiais e modelos de Box e Jenkins*. Simpósio brasileiro de redes neurais - Belo Horizonte, Minas Gerais: UFMG.
- Eberhart, R. and Dovins, R. (1990). *R. Neural Networks PC Tools - A Practical Guide* . Academic Press, San Diego,.
- Ehlers, R. S. (2009). *Análise de séries temporais*. Pontifícia Universidade Católica.
- Ferneda, E. (2005). Redes neurais e sua aplicação em sistemas de recuperação de informação. *Documentação da faculdade de Filosofia Ciências e Letras de Ribeirão Preto - USP - Ci. Inf., Brasília, v. 35, n. 1, p. 25-30,*.
- Ferreira, R. V. (2006). *Previsão de Demanda: Um Estudo de Caso para o Sistema Interligado Nacional* . Universidade Federal de Minas Gerais - Escola de Engenharia - Programa de Pós-Graduação em Engenharia Elétrica - Belo Horizonte - 2006.
- Junior, F. d. J. M. and Caten, C. S. (2004). *Estudo sobre o efeito da Autocorrelação de Modelos AR(1) no Controle Estatístico de Processo*. XXIV Encontro Nac. de Eng. de Produção - Florianópolis, SC, Brasil.

- Kira, K. and Rendell, L. A. (1992). *'The feature selection problem: traditional methods and new algorithm.* In: Proceedings of AAAI'92.
- Kononenko, I. (1994). *Estimating Attributes: Analysis and Extensions of RELIEF.* University of Ljubljana, Faculty of Electrical Engineering e Computer Science,.
- Krose, B. and Smagt, P. V. D. (1996). *An introduction to Neural Networks.* Universidade Federal de Minas Gerais - Dissertação - Engenharia Elétrica.
- Lages, W. F. (2004). *Tópicos Especiais em Automação e Controle II Descrição de Sinais Aleatórios.* Universidade Federal do Rio Grande do Sul - Escola de Engenharia - Departamento de Engenharia Elétrica - Programa de Pós-Graduação em Engenharia Elétrica.
- MAKRIDAKIS, S., WHEELWRIGHT, S. C., and MCGEE, V. (1983). *E. Forecasting: methods and application. 2nd ed.* New York: John Wiley and Sons.
- Medeiros, T. H. d. (2004). *Treínamento de redes neurais artificiais com otimização multi-objetivo e regularização Baesiana:Um estudo comparativo.* Universidade Federal de Minas Gerais - Pós Graduação em Engenharia Elétrica.
- Mendonça, M. B. d. (2011). *Técnicas de Prospecção e Análise de Cenários Futuros nos Governos e Administração - Pública do Brasil: Revisão da Produção Científica Brasileira de 2001 a 2010.* V Encontro de Estudos em Estratégia - Porto Alegre / RS 2011.
- Miles, I. (2004). *Background analysis: trend extrapolation; analysis of framework; megatrend analysis; pp. 1-5.* Foresight Methodologies. 1st ed. United Nations Industrial Development Organization; Vienna.
- Morettin, P. A. and Toloí, C. M. C. (1981). *Modelos para Previsão de Séries Temporais.* Rio de Janeiro. IMPA.

- Oliveira, G. A. D. (2015). *Funções periódicas*. Escola Brasil - www.infoescola.com.
- Pereira, A. P. (2005). Série de fourier na morfometria geométrica. *Universidade Católica de Brasília - Departamento de Matemática*.
- Ralston, A. (1978). *A first Course in Numerical Analysis*. McGraw Hill Book Company.
- Robnik, M. and Kononenko, I. (2003). *Theoretical and Empirical Analysis of Relief and RRelief*. University of Ljubljana, Faculty of Electrical Engineering e Computer Science,.
- Ruggiero, M. A. G. and Lopes, V. L. R. (1996). *Cálculo numérico: aspectos teóricos e computacionais*. São Paulo: Pearson Makron Books.2. ed.
- Russel, S. and Norvig, P. (1995). Artificial Intelligence - A Modern Approach . *Prentice-Hall, New Jersey, 1995, p.563-597*.
- Santos, F. J. (2004). *Introdução às Séries de Fourier*. Matemática - Pontifícia Universidade Católica de Minas Gerais.
- Simpson, P. K. (1990). *Artificial Neural Systems: Foundations, Paradigms, applications, and implementations*. Pergamon Press.
- Soares, P. L. B. and Silva, J. P. d. (2011). Aplicação de Redes Neurais Artificiais em Conjunto com o Método Vetorial da Propagação de Feixes na Análise de um Acoplador Direcional Baseado em Fibra Ótica. *Revista Brasileira de Computação Aplicada (ISSN 2176-6649), Passo Fundo, v.3, n.2, p.58-72*.
- Teixeira, R. d. A. (2001). *Treinamento de redes neurais artificiais através de otimização multi-objetivo: Uma nova abordagem para o equilíbrio entre a polarização e a variância*. Universidade Federal de Minas Gerais - Pós Graduação em Engenharia Elétrica.

Tonsig, S. L. (2000). *Redes Neurais Artificiais Multicamadas e o Algoritmo Backpropagation*. Universidade de São Carlos - Gerência de Sistemas de Informação Pontifícia Universidade Católica.

Velasco, M. M. B. R. (2007). *Redes Neurais Artificiais*. Pontifícia Universidade Católica.