UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Ilha Solteira

PREVISÃO DE CARGAS ELÉTRICAS ATRAVÉS DE UM MODELO HÍBRIDO DE REGRESSÃO COM REDES NEURAIS

Thays Aparecida de Abreu Silva

Ilha Solteira – SP

Fevereiro-2012

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Ilha Solteira

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

"Previsão de Cargas Elétricas através de um Modelo Híbrido de Regressão com Redes Neurais"

THAYS APARECIDA DE ABREU SILVA

Orientador (a): Prof. Dra. Anna Diva Plasencia Lotufo

Co-0rientador (a): Prof. Dra. Mara Lucia Martins Lopes

Dissertação apresentada á Faculdade de Engenharia - UNESP - Campus de Ilha Solteira, como parte dos requisitos par obtenção do título de Mestre em Engenharia Elétrica.
Área de Conhecimento: Automação.

Ilha Solteira – SP

Fevereiro-2012

FICHA CATALOGRÁFICA

Elaborada pela Seção Técnica de Aquisição e Tratamento da Informação
Serviço Técnico de Biblioteca e Documentação da UNESP - Ilha Solteira.

Silva, Thays Aparecida de Abreu.

S586p Previsão de cargas elétricas através de um modelo híbrido de regressão com redes neurais / Thays Aparecida de Abreu Silva. -- Ilha Solteira : [s.n.], 2012

62 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2012

Orientadora: Anna Diva Plasencia Lotufo
Co-orientadora: Mara Lucia Martins Lopes
Inclui bibliografia

1. Previsão com metodologia de Box-Jenkins. 2. Modelos ARIMA de Box e Jenkins. 3. Previsão de cargas elétricas. 4. Redes neurais (Computação). 5. Redes neurais artificiais.

UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE ILHA SOLTEIRA
FACULDADE DE ENGENHARIA DE ILHA SOLTEIRA

CERTIFICADO DE APROVAÇÃO

TÍTULO: Previsão de Cargas Elétricas através de um Modelo Híbrido de Regressão com Redes Neurais

AUTORA: THAYS APARECIDA DE ABREU SILVA
ORIENTADORA: Profa. Dra. ANNA DIVA PLASENCIA LOTUFO
CO-ORIENTADORA: Profa. Dra. MARA LÚCIA MARTINS LOPES

Aprovada como parte das exigências para obtenção do Título de Mestre em Engenharia Elétrica , Área: AUTOMAÇÃO, pela Comissão Examinadora:

Profa. Dra. MARA LÚCIA MARTINS LOPES
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira

Prof. Dr. FRANCISCO VILLARREAL ALVARADO
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira

Profa. Dra. LUCIANA CAMBRAIA LEITE
Departamento de Engenharia Elétrica / Universidade Federal de Mato Grosso do Sul

Data da realização: 24 de fevereiro de 2012.

Dedicatória

A Deus

A minha avó Ana Rosa

e minha mãe Bel

Agradecimentos

Primeiramente agradeço a Deus que é minha fortaleza. Por ter me concedido saúde, força e inteligência para realização deste trabalho. Por sempre me levantar em minhas quedas.

A minha avó Ana Rosa e minha mãe Bel, que são duas guerreiras e me criaram com muito amor e sempre acreditaram que sou capaz de tudo que esta ao alcance humano. A minha irmã Mariana, meu padrasto Rogério e ao meu primo David pelo apoio e carinho.

Ao meu namorado Carlos, pelo amor, respeito, compreensão e por sempre me incentivar na realização de um sonho.

Aos meus colegas de laboratório Ana Claudia, Cícero, Dani Justo, Paulinha Sakai, Zé Guilherme, pelos conselhos, otimismo e companheirismo. Em especial ao meu amigo Klayton, pela sua humildade, generosidade e por ter me doado um pouco de seu tempo e seu conhecimento. Pessoas que adoro conviver.

Aos colegas do LaPSEE pela amizade, cafezinhos, em especial ao Lucas Teles e André Bíscaro pelas ajudas nas disciplinas, programas e etc. Sem duvidas este é meu segundo laboratório.

Ao meu amigo Miguel Paredes, que além de companheiro para todas as horas, foi essencial na realização deste trabalho. Nunca me esquecerei dessa amizade.

As minhas amigas de ontem, hoje e sempre Ariadne Galbiatti, Dani Arsenio, Maíra Arruda e Thalita Sanches, pois não importa o tempo em que ficamos sem nos falar, sem nos ver, sei que tenho esses ombros amigos. Gostaria de agradecer também a Mara Crespi por todo apoio moral e financeiro nos momentos difíceis.

As minhas companheiras de republica pela união e respeito que temos umas com as outras, em especial a Raiane, pela companheira que se mostrou nesta etapa da minha vida.

A minha orientadora Anna Diva por ter aceitado me orientar, por ter acreditado em meu trabalho e pela companheira que se mostrou durante esses dois anos.

Em especial a minha co-orientadora Mara que foi fundamental neste projeto, pelas idéias, pela força que me deu, por sempre me incentivar e me fazer acreditar que sou capaz.

Meus agradecimentos a CAPES, por ter me proporcionado uma bolsa de estudos e com isso pude me dedicar integralmente a este trabalho.

Agradeço a todas as pessoas que de alguma forma colaboraram na realização deste trabalho. Meu muito obrigado.

Resumo

Atualmente os sistemas elétricos de potência crescem em tamanho e complexidade e se faz necessário criar alternativas para minimizar o custo total de geração e operação. A previsão de cargas é uma tarefa importante para o planejamento e operação dos sistemas elétricos, pois dela dependem outras tarefas como despacho econômico, fluxo de potência, análise de estabilidade, entre outras. Para tanto esta tarefa deve ser precisa para que o sistema opere de forma segura e confiável. A precisão da previsão é de grande importância já que é através dela que é estabelecida quando e quanto de capacidade de geração e transmissão deve-se dispor para atender a carga prevista sem interrupções no fornecimento. O objetivo deste trabalho é desenvolver um modelo híbrido utilizando os modelos ARIMA de Box & Jenkins e Redes Neurais Artificiais com treinamento realizado pelo algoritmo de Levenberg-Marquartd. Este modelo será utilizado com a finalidade de melhorar a precisão dos resultados com relação à previsão de cargas elétricas a curto prazo. Os resultados obtidos através da metodologia proposta, modelo híbrido de regressão com redes neurais artificiais, foram comparados com demais trabalhos da literatura. É importante destacar que os resultados utilizados na comparação usam o mesmo banco de dados históricos (demanda de carga elétrica) de uma companhia do setor elétrico brasileiro, bem como o mesmo período de janelamento.

Palavras-chave: Modelos ARIMA de Box & Jenkins. Previsão de cargas elétricas. Redes neurais artificiais.

Abstract

Nowadays the electric power systems are increasing and becoming complexes and therefore it is necessary to provide alternatives to minimize the generation and operation costs. Load forecasting is a very important task for planning and operation of electric power systems of which other tasks are dependent, as for example, economic dispatch, power flow, and stability analysis, among others. Therefore, this task (load forecasting) must be precise for a secure and reliable operation of the power system. Forecasting precision is very important to set when and how much generation and transmission capacity is necessary to attend the load without interruptions. The objective of this work is to develop a hybrid model using ARIMA of Box & Jenkins and Neural Networks trained by Levenberg-Marquardt algorithm. This model is used aiming to improve the precision of the short term electrical load forecasting. The results obtained were compared with others available on the literature. It is emphasized that the data used is the same (from a Brazilian electric company) as well as the window period.

Keywords: ARIMA Models by Box & Jenkins. Term load forecast. Artificial neural networks.

Lista de Figuras

Figura 1 - Filtro linear, com entrada a_t, função de transferência ψ(B) e saída Z(t) 27

Figura 2 - Estágios do ciclo de Box & Jenkins 31

Figura 3 - Diagrama de blocos do software R 39

Figura 4 - Estrutura da rede neural 40

Figura 5 - Diagrama de blocos da metodologia proposta 42

Figura 6 - Vetor Erro (Aplicação 1) 45

Figura 7 - Vetor Erro (Aplicação 2) 46

Figura 8 - Vetor Erro (Aplicação 3) 46

Figura 9 - Resultado da previsão de 24h para aplicação 1 47

Figura 10 - Resultado da previsão de 24h para aplicação 2 47

Figura 11 - Resultado da previsão de 24h para aplicação 3 48

Figura 12 - Neurônio biológico 57

Figura 13 - Modelo MCP (McCulloch-Pitts) não-linear de um neurônio artificial 58

Figura 14 - Função degrau bipolar 59

Figura 15 - Função tangente hiperbólica 59

Figura 16 - Função Logística 60

Figura 17 - Função Linear 60

Figura 18 - Disposição das redes neurais artificiais 61

Lista de Tabelas

Tabela 1 - Períodos utilizados 43

Tabela 2 - Parâmetros da rede neural 44

Tabela 3 - MAPE e Erro Máximo para aplicação 1 49

Tabela 4 - MAPE e Erro Máximo para aplicação 2 49

Tabela 5 - MAPE e Erro Máximo para aplicação 3 50

Lista de Abreviaturas e Siglas

AIC	Critério de Informação de Akaike
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
BIC	Critério de Informação de Bayesiano
FAC	Função de Autocorrelação
FACP	Função de Autocorrelação Parcial
MA	Moving Average
MAPE	Erro Percentual Absoluto Médio
MCP	McCulloch-Pitts
MLP	Multilayer Perceptron
RBF	Radial Basis Function
RNA	Rede Neural Artificial
RNAs	Redes Neurais Artificiais
SARIMA	Seasonal Auto Regressive Integrated Moving Average
VLSI	Very-Large-Scale-Integration

Lista de Símbolos

X	Vetor contendo as entradas da rede neural
W	Vetor contendo os pesos dos neurônios
$x_1, x_2, \dots, x_n$	Entradas da rede neural
$w_1, w_2, \dots, w_n$	Pesos dos neurônios
y_i	Saídas da rede neural
θ	Limiar ou peso bias
s_i	Saída intermediaria do neurônio
f_d	Função degrau bipolar
f_t	Função tangente hiperbólica
f_l	Função logística
f_{li}	Função linear
H	Matrix hessiana
J	Matrix jacobiana
μ_K	Constante de Levenberg
I	Matriz identidade
Z_t	Previsão desejada da carga elétrica
a_t	Resíduos (ruído branco)
Es	Esperança
Var	Variância
ω	Termo constante
$\psi(B)$	Função de transferência do filtro linear
$\phi_1, \dots, \phi_p$	Coeficientes auto-regressivos
p	Ordem do modelo auto-regressivo

$\theta_1, \ldots, \theta_q$	Coeficientes de média móvel
q	Ordem do modelo de média móvel
$\phi(B)$	Operador diferença do modelo auto-regressivo
$\theta(B)$	Operador diferença do modelo de média móvel
M_t	Série temporal diferenciada
d	Número de diferenças para tornar a serie estacionaria
Δ	Operador diferença
ρ_j	Autocorrelação teórica de atraso j
r_j	Autocorrelação amostral
N	Número de observações da série Z_t
ρ_{jj}	Autocorrelação parcial teórica
r_{jj}	Autocorrelação parcial amostral
σ_a^2	Variância dos resíduos
$\hat{\sigma}_a^2$	Variância estimada dos resíduos
$\hat{\phi}_1, \ldots, \hat{\phi}_p$	Estimação dos coeficientes auto-regressivos
$\hat{\theta}_1, \ldots, \hat{\theta}_q$	Estimação dos coeficientes de média móvel
L	Função de verossimilhança
$\hat{a}_t$	Resíduos estimados
$\hat{r}_j$	Autocorrelações dos resíduos estimados
Z_t^{SA}	Serie sazonal ajustada
T_t	Componente de tendência
S_t	Componente sazonal
$\hat{S}_t$	Estimação da componente sazonal
$\Phi_1, \ldots, \Phi_P$	Coeficientes auto-regressivos sazonais

P	Ordem do modelo auto-regressivo sazonal
$\Theta_1, \dots, \Theta_Q$	Coeficientes de médias móveis sazonais
Q	Ordem do modelo de média móvel sazonal
D	Número de diferenças sazonais
$\Phi(B^m)$	Operador diferença do modelo auto-regressivo sazonal
$\Theta(B^m)$	Operador diferença do modelo de média móvel sazonal
Δ_m	Operador diferença sazonal
m	Período da sazonalidade
b	Dimensão do vetor X
$E(h-v)$	Valor do erro v horas anterior à hora corrente h
$E(h+1)$	Valor do erro correspondente à hora subsequente a hora corrente h
te	Vetor tempo referente aos dados históricos
E_t	Vetor erro
$L(h)$	Valor da carga real referênte a hora h
$\underline{L}(h)$	Valor da carga estimada pelo modelo híbrido referente a hora h
NT	Número total de horas

Sumário

1 Introdução 17

1.1 Contribuições do trabalho 19

1.2 Organização do texto 19

2 Previsão de Cargas Elétricas 20

2.1 Introdução 20

2.2 Classificação da previsão e aplicação 20

2.3 Fatores que influenciam a Carga Elétrica 21

2.4 Metodologias para Previsão de Carga 21

3 Redes Neurais Artificiais 23

3.1 Benefícios das Redes Neurais 23

3.2 Redes Perceptron Multicamadas 24

3.2.1 Treinamento de Lenvenberg-Marquartd 24

4 ARIMA de Box & Jenkins 27

4.1 **Introdução** 27

4.2 Modelos estacionários 27

4.2.1 Modelos auto-regressivos 28

4.2.2 Modelos de médias móveis 29

4.2.3 Modelos auto- regressivos e de médias móveis 30

4.3 Modelos não-estacionários – Modelos Auto- Regressivos Integrados de Médias Móveis (ARIMA) 30

4.4 Metodologia de Box & Jenkins 31
4.4.1 Identificação 32
4.4.2 Estimação 35
4.4.3 Verificação 36
4.5 Modelos Sazonais 37

5 Modelo Híbrido de Regressão (ARIMA) e Rede Neural Artificial 39
5.1 Introdução 39
5.2 Descrição do Modelo Híbrido 39

6 Aplicações e Resultados 43

7 Conclusões e Sugestões para Futuros Trabalhos 51
7.1 Conclusões 51
7.2 Sugestões para trabalhos futuros 51

Referências 53

Apêndice A - Redes Neurais Artificiais 56
A.1 Reflexão Histórica 56
A.2 Modelo do neurônio biológico 57
A.3 Modelo do neurônio artificial 58
A.4 Estrutura da rede neural 60
A.5 Treinamento da rede neural 62

1 Introdução

O conhecimento futuro da carga exerce um papel importante no planejamento da expansão, fluxo de potência, operação econômica, análise e controle de segurança de Sistemas de Energia Elétrica. Na literatura destacam-se as seguintes técnicas para previsão de carga: técnica de regressão linear simples ou múltipla, alisamento exponencial, estimação de estado, filtro de Kalman, ARIMA (Auto Regressive Integrated Moving Average) de Box e Jenkins (1976). Esses métodos necessitam da modelagem prévia da carga para aplicação posterior. Os fatores que ajudam a modelar a carga podem ser condições meteorológicas como: nebulosidade, velocidade do vento, variações bruscas de temperatura; fatores diversos como efeito de dias atípicos (feriados, greves, etc.) (O'DONOVAN, 1983).

Os métodos baseados na inteligência artificial vêm demonstrando grande eficiência, quando aplicados a problemas da vida real. Atualmente, o uso das redes neurais artificiais (RNAs) é um método alternativo para resolução do problema de previsão de cargas elétricas (HAYKIN, 1999).

Dentro da literatura especializada podem-se encontrar inúmeros trabalhos que utilizam redes neurais artificiais e os modelos ARIMA de Box & Jenkins na previsão de séries temporais. Uma idéia interessante seria tentar unir as vantagens que cada técnica possui de forma a obter um modelo mais robusto que proporcione previsões mais precisas, conhecido como modelo híbrido. Existem poucos trabalhos utilizando tal metodologia, modelo híbrido, principalmente quando aplicado ao problema de previsão de cargas elétricas a curto prazo. A maioria dos trabalhos que abordam o modelo híbrido, ARIMA com RNAs, são voltados para o tratamento de séries temporais aplicados a uma variedade de problemas em diversas áreas da vida real (FARUK, 2009; KHASHEI; BIJARI, 2010, 2011).

Um exemplo de modelo híbrido bastante utilizado na literatura é o modelo no qual separa a série temporal em uma parte linear e outra não linear, em que o ARIMA é responsável pela parte linear da série e a rede neural pela parte não linear dos resíduos do ARIMA (ABURTO; WEBER, 2007; KHASHEI et al., 2009; VALENZUELA et al., 2008; ZANG, 2003). Alguns destes trabalhos como o de Valenzuela (2008) e Khashei (2009) ainda combinam ou regras nebulosas (fuzzy) ou algoritmos genéticos para otimizar os resultados.

No trabalho de Faruk (2009) é feita a previsão da qualidade da água da Málasia, levando em consideração a temperatura, o boro e o oxigênio dissolvido. Devido à capacidade de reconhecer os padrões de séries temporais e características não-lineares, o modelo híbrido oferece uma precisão muito melhor ao longo dos modelos ARIMA e RNA individualmente. No trabalho de Koutroumanidis et al. (2009), é apresentada uma descrição da contribuição de fontes de energia renováveis para a produção de eletricidade. Os modelos ARIMA e RNA são unidos, sendo o ARIMA responsável por gerar os parâmetros da carga e as redes neurais são usadas para prever os preços futuro de venda de lenha produzido pela floresta estatal na Grécia. O uso do modelo híbrido obteve ótimos resultados, possibilitando aos fabricantes prosseguirem com um planejamento mais racional para produção e mercado de lenha. Já nos trabalhos de Khashei e Bijari (2010), (2011), são considerados os dados históricos dos registros das manchas solares por ano, o número de linces que é capturado por ano no rio Mackenzie (Canadá) e a taxa de cambio libra (Britânica) / dólar (Estados Unidos), também com o objetivo de obter uma previsão precisa utilizando o hibridismo entre ARIMA e RNA, sendo que os modelos ARIMA são utilizados na primeira fase para gerar os dados necessários a partir das series históricas, em seguida na segunda fase uma rede neural é usada para modelar os dados gerados pelo modelo ARIMA, e para prever o futuro da serie temporal.

Neste trabalho será desenvolvido um modelo híbrido previsor de cargas elétricas a curto prazo, utilizando os modelos ARIMA de Box & Jenkins e RNAs, sendo utilizada a rede Perceptron Multicamadas (PMC), com algoritmo de treinamento de Levenberg-Marquartd, no qual o modelo ARIMA é responsável por gerar melhor modelo e a RNA por realizar a previsão do erro futuro para substituir no modelo gerado pelo ARIMA. A previsão de carga elétrica é então executada pela substituição do modelo previsto no melhor modelo ARIMA obtido. É importante enfatizar que a metodologia utilizada neste trabalho é inovadora, e apesar de existirem métodos que trabalham com modelos ARIMA e RNAs, estes aplicam estas técnicas de forma completamente diferente do que está sendo proposto nesta dissertação.

O modelo proposto foi testado nos períodos de 30, 60 e 90 dias respectivamente, todos com o objetivo de prever o dia 1 de agosto de 1998. Os resultados foram comparados com outros trabalhos que também utilizaram os dados históricos da mesma companhia do setor elétrico brasileiro e o modelo híbrido obteve um ganho de desempenho.

1.1 Contribuições do trabalho

A contribuição deste trabalho consiste no desenvolvimento de um sistema previsor de cargas elétricas de forma confiável, através do hibridismo entre os modelos ARIMA e RNAs, que de um modo geral é capaz de realizar previsões a curto prazo, que neste trabalho especificamente será de 24 horas.

1.2 Organização do texto

O trabalho está dividido em sete capítulos, sendo organizados da seguinte forma: previsão de cargas elétricas, redes neurais artificiais, ARIMA de Box & Jenkins, metodologia proposta, resultados e finalizando com as conclusões.

O capítulo 2 trata da previsão de cargas elétricas descrevendo a classificação da previsão, os fatores que influenciam a modelagem da carga e as técnicas existentes nos dias de hoje.

O capítulo 3 é apresentada a rede Perceptron Multicamadas e o treinamento de Levenberg-Marquartd, que serão utilizados neste trabalho.

No capítulo 4 são apresentadas as classes dos modelos ARIMA e ARIMA sazonais conhecidos como SARIMA. Também são explicadas cada fase da metodologia de Box & Jenkins: identificação, estimação e verificação.

No capítulo 5 é descrita a metodologia desenvolvida, modelo híbrido utilizando o ARIMA e Perceptron Multicamadas, a fim de fazer previsões de cargas elétricas a curto prazo, com o objetivo de obter previsões mais precisas, ou seja,com menor erro.

O capítulo 6 apresenta os resultados obtidos através do modelo proposto, sendo estes comparados com outras metodologias encontradas na metodologia especializada.

No capitulo 7 são desenvolvidas algumas conclusões gerais e propostas futuras a respeito do trabalho.

Este texto, também, contém um apêndice, sendo que este apresenta os conceitos básicos das redes neurais artificiais.

2 Previsão de Cargas Elétricas

2.1 Introdução

O passo principal necessário para obter um sistema previsor de cargas elétricas é entender o comportamento da carga e utilizar os fatores que as influenciam como um instrumento de previsão.

As cargas elétricas são compostas por milhares de aparelhos e dispositivos que, quando consideradas em conjunto, apresentam tendências que são influenciadas por fatores que ajudam a modelar a carga. Como exemplo, os fatores mais importantes para previsão de cargas elétricas incluem: condições meteorológicas como nebulosidade, velocidade e direção do vento, variações brusca de temperatura, umidade; fatores diversos como efeitos de dias atípicos (feriados) e dias especiais (sábados e domingos), o dia a dia da semana e os efeitos sazonais, onde o consumo varia em diferentes estações do ano. Tais fatores tentam representar da melhor maneira o comportamento da carga para então efetuar-se a previsão (LOPES, 2000).
O método previsional necessita, como todos os demais métodos de previsão, basicamente, de duas condições fundamentais que são (ALMEIDA et al., 1991):

- Informações sobre o passado quantificadas em forma de dados;
- Assumir que o comportamento ocorrido no passado de certa forma irá se repetir no futuro.

2.2 Classificação da previsão e aplicação

A carga elétrica pode ser prevista a longo, médio, curto e curtíssimo prazo (LOTUFO; MINUSSI, 1999; OTHMAN et al., 2009). A previsão a longo prazo é realizada em um período de 10 a 20 anos, esse tipo de previsão é realizada principalmente na operação de planejamento de sistemas de potência. A previsão a médio prazo é aquela que abrange um período de poucas semanas ou até mesmo um ano, essa é utilizada para o planejamento de suprimento de combustível, programas de manutenção e planejamento de produção a médio prazo. A previsão de curto prazo é realizada no intervalo de tempo de meia hora a poucas horas, ou até mesmo de 1 dia a 1semana, ela ocorre no planejamento e operações econômica e de segurança diária dos sistemas de potência. E por fim a de curtíssimo prazo, que ocorre entre poucos se-

gundos e pouco menos de quinze minutos, esse tipo de previsão é essencialmente usada para o planejamento de operação on-line e controle.

Neste trabalho somente será enfocado o problema de previsão a curto prazo e a previsão realizada é efetuada apenas para prever 24 horas à frente.

2.3 Fatores que influenciam a Carga Elétrica

Existem variáveis e fatores que influenciam no comportamento da carga, que atuam de forma significativa na precisão da previsão. Dentre eles são (O'DONOVAN, 1983):

- Fatores climáticos: são temperatura, umidade relativa do ar, o nível de luminosidade e a velocidade do vento. Esses fatores têm influência, especialmente, em áreas residenciais. Geralmente a variável mais utilizada é a temperatura.
- Fatores cíclicos ou temporais: são fatores que se repetem como dias, semanas, meses, ano, além de feriados. Na periodicidade diário-semanal, temos uma baixa demanda para eletricidade nos finais de semana e uma alta para os dias úteis. Na periodicidade anual temos a mudança de estação ao longo do ano. Já a ocorrência de feriados diminui os valores de carga em relação aos dias comuns.
- Fatores atípicos ou aleatórios: nesta categoria podemos citar as greves ou mesmo uma transmissão de programa de televisão de interesse geral. Dificilmente, neste caso, poderíamos prever seus efeitos sobre a carga.

2.4 Metodologias para Previsão de Carga

Na literatura, encontram-se várias técnicas para realizar a previsão de carga (O´DONOVAN, 1983), dentre elas, regressão linear múltipla ou simples, alisamento exponencial, estimação de estado, filtro de Kalman, modelos ARIMA de Box & Jenkins, entre outros. Estes métodos necessitam da modelagem prévia da carga para posteriormente serem aplicados. Para modelar a carga é imperativo que se conheça algumas informações, tais como: efeitos atmosféricos (dias nublados, velocidade do vento, variações repentinas de temperatura) e efeitos de dias não convencionais (feriados, greves, etc.). Quanto mais informações forem disponibilizadas, maiores serão as chances de obter soluções de qualidade. Após ter sido

realizado a modelagem da carga, usando estas informações, inicia-se a execução do algoritmo, a fim de obter os resultados (previsão de carga). Outras metodologias para previsão de cargas foram propostas nas últimas duas décadas, como o uso de redes neurais artificiais, lógica nebulosa (fuzzy), sistemas híbridos, entre outros (HAYKIN, 1999). Neste trabalho será abordada a previsão de cargas utilizando modelos ARIMA de Box & Jenkins e RNAs.

3 Redes Neurais Artificiais

3.1 Benefícios das Redes Neurais

É de fácil aceitação que uma rede neural extrai seu poder computacional através de sua estrutura maciça paralelamente distribuída. As redes neurais possuem a capacidade de aprender e, portanto de generalizar. A generalização ocorre na rede neural visando produzir saídas adequadas para entradas que não estavam presentes durante o treinamento, possibilitando a solução de problemas complexos de difícil tratamento.

Entre as características e propriedades observadas nas redes neurais destacam-se (KAYKIN, 1999):

- Não- linearidade: é uma propriedade importante, em especial quando o mecanismo físico que gera o sinal de entrada for não linear;
- Mapeamento de entrada e saída: o treinamento é repetido até que a rede atinja um estado estável sem modificações significativa nos pesos sinápticos;
- Adaptabilidade: diz respeito à capacidade de adaptação dos pesos sinápticos às modificações do ambiente, possibilitando re-treinamento para lidar com modificações que podem ocorrer em tempo real;
- Resposta a evidências: no caso de classificações de padrões, é possível obter respostas confiáveis por meio de rejeição de padrões ambíguos, o que melhora o desempenho da rede (rede de treinamento não supervisionado);
- Informação contextual: pré-definido pela estrutura e estado de ativação e uma rede;
- Tolerância a falhas: tendo em vista o nível elevado de neurônios artificiais interconectados, o sistema torna-se tolerante a falhas caso a estrutura interna seja corrompida;
- Implementação em VLSI (very-large-scale-integration): Tal tecnologia possibilita a captura de comportamentos complexos de forma altamente hierárquica.

3.2 Redes Perceptron Multicamadas

A rede neural Perceptron Multicamadas (PMC) é uma importante classe de redes neurais e também uma das mais versáteis quanto à aplicabilidade, pois podem ser aplicadas em diversos tipos de problemas. São caracterizadas pela presença de pelo menos uma camada intermediária de neurônios situada entre a camada de entrada e a camada de saída, assim, tais redes têm no mínimo duas camadas (HAYKIN, 1999).

A rede PMC possui arquitetura do tipo feedforward de camadas múltiplas e seu treinamento é efetivado de forma supervisionada.

As redes PMC funcionam da seguinte forma: os sinais são apresentados em sua camada de entrada, em seguida os neurônios pertencentes às camadas intermediárias extraem a maior parte das informações que são codificadas em seus pesos sinápticos e limiares, enfim os neurônios da camada de saída recebem os sinais vindos da última camada intermediária e produz uma resposta padrão que será a saída da rede neural (SILVA et al., 2010).

3.2.1 Treinamento de Lenvenberg-Marquartd

A rede neural possui a capacidade de aprender, para tanto deve ser treinada, ajustando assim seus pesos, para tal, necessita-se de algoritmo de treinamento. Neste trabalho, será abordado o algoritmo de Levenberg-Marquardt que é um algoritmo que utiliza gradiente de segunda ordem, que se baseia, para a aceleração do treinamento, na determinação das derivadas de segunda ordem do erro quadrático em relação aos pesos, diferindo do algoritmo retropropagação (backpropagation) tradicional que considera as derivadas de primeira ordem. Esse algoritmo é considerado o método mais rápido para treinamento de redes feedforward.

O treinamento de Levenberg-Marquardt (BARBOSA et al., 2005) faz a aproximação para a matriz mostrada em (1), determinada em função da matriz Jacobiana, que contém as primeiras derivadas dos pesos em função dos pesos sinápticos, expressa em (2):

$$H = \frac{\partial^2 E_R(W)}{\partial W^2} \qquad 1$$

sendo $\boldsymbol{E_R}$ ($\boldsymbol{W}$) definido conforme a expressão (2):

$$E_R(W) = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_{ei})^2 \qquad 2$$

$$J = \frac{\partial e(W)}{\partial W} \qquad 3$$

sendo $e(W)$ definido conforme a expressão (4):

$$e(W) = \sum_{i=1}^{n}(y_i - y_{ei}) \qquad 4$$

sendo:

y_i = saída fornecida pela rede

y_{ei} = valor exato correspondente a saída da rede

A determinação da matriz Jacobiana é mais simples que a determinação da matriz Hessiana. Como, para uma rede neural, o desempenho de treinamento é expresso em função da soma dos erros quadráticos, a matriz Hessiana pode ser expressa pela expressão (5) (BARBOSA et al., 2005):

$$H = J^{\mathrm{T}}(W)\ J(W) \qquad 5$$

O método de Newton atualiza os pesos segundo (6):

$$W(k+1) = W(k) - H^{-1} g_k \qquad 6$$

em que g_k pode ser escrito conforme (7):

$$g_k = 2J^{\mathrm{T}}(W)\ e(W) \qquad 7$$

O algoritmo de Levenberg-Marquardt procede à atualização dos pesos baseado na expressão do método de Newton (6), realizando as modificações para a determinação da matriz Hessiana, mostrada em (8):

$$W(k+1) = W(k) - [J^{T}(W)\,J(W) + \mu_{k}\mathrm{I}]^{-1} J^{T}(W)\,e(W) \qquad 8$$

sendo:

I	=	matriz identidade
μ_k	=	constante do método de Levenberg-Marquardt

O parâmetro μ_k funciona como um fator de estabilização do treinamento, ajustando a aproximação de forma a utilizar a rápida convergência do método de Newton, evitando passos longos que possam levar a um erro de convergência (BARBOSA et al., 2005).

O método de Levenberg-Marquartd apresenta convergência em menos iterações, mas requer mais cálculos por iteração devido aos cálculos das matrizes inversas, com isso requer grande esforço computacional. Esse algoritmo de treinamento é o mais rápido para redes neurais, quando se trabalha com um número moderado de parâmetros na rede, quando esse número é elevado, a utilização desse algoritmo não é adequada (RANGANATHAN, 2004).

4 ARIMA de Box & Jenkins

4.1 Introdução

A hipótese de erros não-correlacionados introduz sérias limitações na validade de vários modelos, que descrevem os comportamentos de séries econômicas, sociais, entre outras, em que os erros observados são auto-correlacionados e influenciam na evolução do processo, logo, para estes casos os modelos ARIMA são de grande utilidade. Por mais de meio século os modelos auto-regressivos integrados de médias móveis, ARIMA(p, d, q), dominaram muitas áreas de previsão de series temporais.

Os métodos ARIMA dependem apenas de seus dados para a especificação do modelo, sendo a metodologia de Box & Jenkins a mais utilizada.

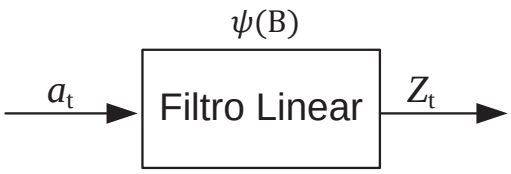
O principal enfoque de Box & Jenkins é a realização de previsão. Essa metodologia permite que valores futuros de uma série (neste caso previsão de cargas elétricas) sejam previstos utilizando apenas seus valores presente e passados, através da correlação temporal que existe entre os valores exibidos pela série. Os modelos possíveis dentro da metodologia de Box & Jenkins podem ser considerados como modelos estacionários ou não-estacionários.

4.2 Modelos estacionários

Os modelos estacionários são apropriados para descrever séries estacionárias, isto é, séries que se desenvolvem no tempo ao redor de uma média constante. Estes modelos são casos particulares de um modelo de filtro linear.

Desta forma, os modelos supõem que a série temporal seja gerada através de um filtro linear ou um sistema linear (MORETTIN; TOLOI, 2006), que está definido na Figura 1.

Figura 1 - Filtro linear, com entrada a_t, função de transferência ψ(B) e saída Z_t



Fonte: Morettin e Toloi (2006).

O modelo também pode ser definido pela Equação (9), sendo Z_t um processo linear discreto em que:

$$\begin{aligned} \mathrm{Es}(a_t) &= 0, \forall\, t, \\ Var(a_t) &= \text{series aleatórias} \\ \mathrm{Es}(a_t a_s) &= 0, s \neq t. \end{aligned}$$

$$Z_t = \omega + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \omega + \psi(B) a_t, \qquad 9$$

sendo

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots \qquad 10$$

no qual:

$\mathrm{Z_t}$	=	carga elétrica
a_t	=	resíduos (ruído branco)
ω	=	termo constante
$\psi(B)$	=	função de transferência do filtro linear

Chamando $\tilde{\mathrm{Z}}_\mathrm{t} = \mathrm{Z_t} - \omega$, irá obter-se:

$$\tilde{Z}_t = \psi\,(B) a_t\,. \qquad 11$$

Pode-se escrever $\tilde{\mathrm{Z}}_\mathrm{t}$ como uma soma ponderada de valores passados $\tilde{\mathrm{Z}}_{\mathrm{t-1}}$, $\tilde{\mathrm{Z}}_{\mathrm{t-2}}$, . . . mais o ruído branco a_t (MORETTIN; TOLOI, 2006):

$$\tilde{Z}_t = \pi_1\, \tilde{Z}_{t-1} + \pi_2\, \tilde{Z}_{t-2} + \ldots + a_t = \sum_{j=1}^{\infty} \pi_j\, \tilde{Z}_{t-j} + a_t\,. \qquad 12$$

4.2.1 Modelos auto-regressivos

Consideremos a Equação (12), se $\pi_j = 0, j > p$, será obtido um modelo auto-regressivo de ordem p, denotado por AR(p), que é baseado nos dados anteriores da série e representado pela Equação (13) (BOX et al., 1994; HAMILTON, 1994).

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \ldots + \phi_p \tilde{Z}_{t-p} + a_t \ , \qquad 13$$

no qual:

$\phi_1, \ldots, \phi_p$ = coeficientes auto-regressivos;

p = ordem do modelo auto-regressivo.

Pode-se também, escrever (13) como um operador linear:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ \ldots - \ \phi_p B^p \ , \qquad 14$$

ou

$$\phi(B)\tilde{Z}_t = \ a_t \, . \qquad 15$$

4.2.2 Modelos de médias móveis

Consideremos a Equação (9), se $\psi_j = 0, j > q$; será obtido um processo de médias móveis de ordem q, denotada por MA(q), que são baseados nos erros anteriores da série, e definidos pela Equação (16) (BOX et al., 1994; HAMILTON, 1994).

$$Z_t = \mu + a_t - \ \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \qquad 16$$

em que:

$\theta_1, \ldots, \theta_q$ = coeficientes de média móvel;

q = ordem do modelo.

sendo a Equação (17) um operador linear da Equação (16):

$$\theta(B) = 1 - \ \theta_1 B - \theta_2 B^2 - \ \ldots - \ \theta_q B^q, \qquad 17$$

ou

$$\tilde{Z}_t = \ \theta(B) a_t \, . \qquad 18$$

4.2.3 Modelos auto- regressivos e de médias móveis

Se o enfoque for um modelo com um número não muito grande de parâmetros, a inclusão de termos auto-regressivos e de médias móveis é a solução adequada, surgem então os modelos ARMA(p, q) cuja equação será (BOX et al., 1994; HAMILTON, 1994):

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \ldots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \,. \qquad 19$$

ou

$$\phi(B)\tilde{Z}_t = \theta(B) a_t. \qquad 20$$

4.3 Modelos não-estacionários – Modelos Auto- Regressivos Integrados de Médias Móveis (ARIMA)

A maioria das séries encontradas na pratica são não estacionarias, ou seja, não se desenvolvem no tempo através de uma média constante. Para torná-las estacionarias é necessária a aplicação de diferenças. O número de diferenças para tornar uma série estacionaria é denominado ordem de integração.

Se $M_t = \Delta^d Z_t$ é estacionária, podemos representar M_t por um modelo ARMA(p, q), ou seja,

$$\phi(B) M_t = \theta(B) a_t \qquad 21$$

Se M_t for uma diferença de Z_t, então Z_t é uma integral de M_t, daí dizermos que Z_t segue um modelo auto-regressivo, integrado, de médias móveis, ou modelo ARIMA, denotado por ARIMA($p,\ d,\ q$) e representado pela Equação (22) (MORETTIN; TOLOI, 2006).

$$\phi(B) \Delta^d Z_t = \theta(B) a_t \qquad 22$$

em que d é o número de diferenças necessárias para tornar a série estacionária.

Serão citados abaixo alguns casos particulares do modelo ARIMA (MORETTIN; TOLOI, 2006).

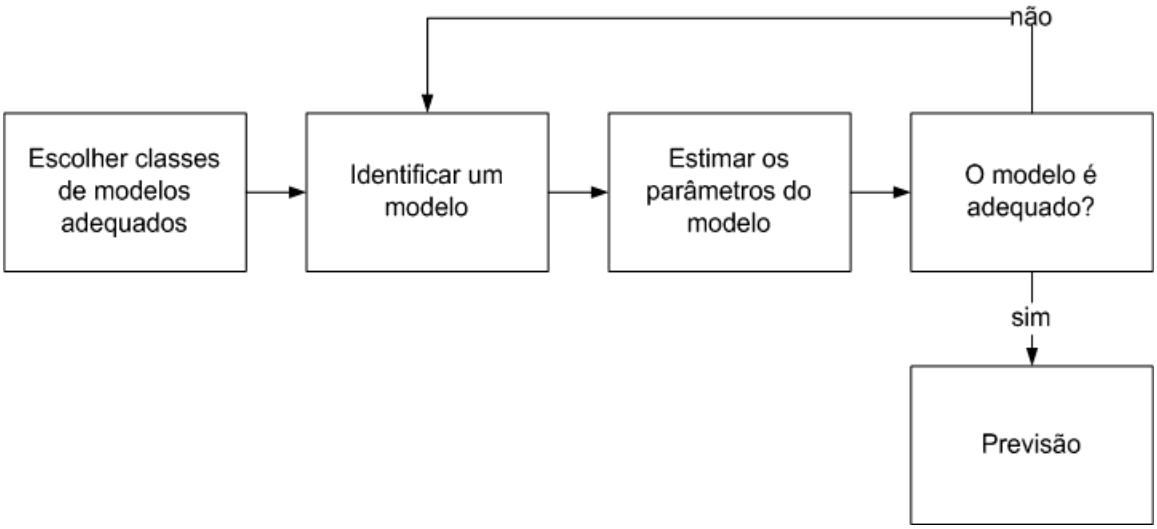
- ARIMA$(0,1,1)$: $\Delta Z_t = (1 - \theta B)a_t$;
- ARIMA$(1,1,1)$: $(1 - \phi B)\Delta Z_t = (1 - \theta B)a_t$;
- ARIMA$(p,0,0)$ =AR(p);
- ARIMA$(0,0,q)$ =MA(q);
- ARIMA$(p,0,q)$ =ARMA(p,q).

4.4 Metodologia de Box & Jenkins

Uma metodologia muito utilizada na análise de modelos paramétricos (que envolve um número finito de parâmetros) é a abordagem de Box & Jenkins (1970). Tal metodologia consiste em ajustar modelos ARIMA$(p,\ d,\ q)$ a um conjunto de dados (MORETTIN; TOLOI, 2006).

A construção do modelo é abordada em um ciclo iterativo: identificação, estimação e verificação, mostrada na Figura 2.

Figura 2 - Estágios do ciclo de Box & Jenkins



Fonte: Elaboração da própria autora

- **Identificação**. Consiste em descobrir qual dentre as várias versões do modelo ARIMA descreve o comportamento da série.
- **Estimação**. Consiste em estimar os parâmetros do modelo identificado;
- **Verificação**. Tem por finalidade avaliar se o processo de estimação foi bem sucedido.

Quando se obtém um modelo satisfatório, já é possível fazer a previsão.

4.4.1 Identificação

Trataremos agora da fase de identificação. O objetivo nesta fase, é determinar os valores de p, d e q do modelo ARIMA($p,\ d,\ q$) além de estimativas preliminares dos parâmetros da fase de estimação. Esta escolha é feita principalmente com base nas autocorrelações e autocorrelações parciais estimadas.

O procedimento de identificação consiste em (MORETTIN; TOLOI, 2006):

- tomar diferenças da série quantas vezes necessárias para obter-se uma série estacionária, de tal forma que o processo $\Delta^d Z_t$ seja reduzido a um ARMA(p, q). O número de diferenças, d necessárias para que o processo seja estacionário, é alcançado quando a FAC amostral de $W_t = \Delta^d Z_t$ decresce rapidamente para zero.
- identificar o processo resultantes ARMA(p, q), analisando as autocorrelações e auto correlações parciais estimadas.

Logo a realização do processo de identificação necessita de outros instrumentos, sendo estes, a função de autocorrelação e a função de autocorrelação parcial.

- Função de autocorrelação (FAC)

Uma função de autocorrelação é a correlação entre o comportamento anterior da carga (neste caso) e o fator que se deseja influenciar no valor futuro da carga. O coeficiente de autocorrelação de ordem j, ou seja, a autocorrelação entre Z_te Z_{t-j} é obtido pela Equação (23).

$$\rho_j = \frac{Cov(Z_t, Z_{t-j})}{V(Z_t)} = \frac{\gamma_j}{\gamma_0} \qquad 23$$

A sequencia de pares(j, ρ_j), j $= 1, 2, \ldots$, é denominada função de autocorrelação.

Geralmente, o coeficiente de autocorrelação ρ_j envolve parâmetros desconhecidos, assim na prática é necessário trabalhar com coeficiente de autocorrelação "amostral" r_j, definido pela Equação (24) (CRYEN; CHAN, 2008).

$$r_j = \frac{\sum_{t=j+1}^{N}(Z_t - \overline{Z})(Z_{t-j} - \overline{Z})}{\sum_{t=1}^{N}(Z_t - \bar{Z})^2} \quad , \qquad j = 1,2, \ldots \tag{24}$$

sendo:

ρ_j	=	autocorrelação teórica;
γ_j	=	covariância de Z_t
γ_0	=	variância de Z_t
r_j	=	autocorrelação amostral;
N	=	número de observações da série Z_t.

É importante ressaltar que as autocorrelações amostrais são somente estimativas de autocorrelações teóricas ρ_j, logo tendem ao mesmo padrão, assim pode-se concluir muitas propriedades do processo estocástico subjacente a partir de um estudo da função de autocorrelação amostral (O'DONOVAN, 1983).

Mesmo sendo de grande importância, nem sempre a função de autocorrelação amostral permite especificar o modelo apropriado, assim será necessária outra característica da série temporal sendo esta a função de autocorrelação parcial amostral expressa r_{jj}.

- Função de autocorrelação parcial (FACP)

A FACP amostral é também, por sua vez, uma estimativa da autocorrelação parcial teórica ρ_{jj}. Calculada dos valores da série temporal observada. A autocorrelação parcial teórica é definida como sendo a autocorrelação entre quaisquer duas variáveis Z_te Z_{t+j}, separadas j atrasos de tempo sendo que as variáveis Z_{t+1}, Z_{t+2}, ..., Z_{t+j-1} são eliminadas. As autocorrelações parciais amostrais possuem a seguinte forma (O'DONOVAN, 1983):

$$r_{11} = r_1$$

$$r_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

Sendo que a partir das equações r_{33} tornam-se cada vez mais complicadas.

Da mesma forma ocorrida com a FAC amostral, a FACP amostral tende ao mesmo padrão que a FACP teórica, de maneira que se pode utilizar a FACP amostral para identificar o modelo apropriado no processo estocástico.

A identificação é a fase mais crítica da metodologia de Box & Jenkins, vários pesquisadores usando a mesma série podem identificar modelos diferentes. Isto porque, trabalhando com a FAC e a FACP amostrais, muitas vezes fica difícil decidir se elas estão decrescendo ou se são truncadas.

Muitos pesquisadores preferem utilizar outros procedimentos de identificação, que não depende de quem está analisando a série de tempo. Esse procedimento faz uso de critérios de seleção de modelo construídos com base na variância estimada de a_t, no tamanho da amostra e nos valores de p e q.

Os critérios mais utilizados são o Critério de Informação de Akaike (AIC) e o Critério de Informação Bayesiano (BIC), denotados pelas Equações (25) e (26) respectivamente (FAVA, 1999).

$$AIC = ln\hat{\sigma}_a^2 + \frac{2(p+q)}{N} \qquad 25$$

$$BIC = ln\hat{\sigma}_a^2 + \frac{(p+q)ln(N)}{N} \qquad 26$$

Assim são estimados os modelos correspondentes a vários pares (p, q) e escolhidos a especificação que apresentar o menor valor para AIC ou para BIC.

A presença de p e q nas fórmulas dos critérios AIC e BIC tem por objetivo “penalizar” os modelos com muitos parâmetros, pois modelos com poucos parâmetros estimam-se com mais facilidade (FAVA, 1999).

Uma crítica que se faz a esses critérios é que eles podem conduzir a modelos com valores de p e/ou q maiores do que o necessário (FAVA, 1999).

Os critérios AIC e BIC devem ser usados como procedimentos complementares e não alternativos baseando-se na FAC e na FACP (FAVA, 1999).

4.4.2 Estimação

Identificado um modelo provisório para série temporal, o próximo passo será estimar seus parâmetros ϕ se houver um componente auto- regressivo, θ se houver o filtro de médias móvel e a variância do ruído branco σ_a^2 do modelo (WEI, 2006):

$$M_t = \phi_1 M_{t-1} + \cdots + \phi_p M_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \qquad 27$$

em que: $M_t = \Delta^d Z_t = (1-B)^d Z_t$.

A estimação pode ser feita por mínimos quadrados ou por máxima verossimilhança.

A estimação por mínimos quadrado requer a minimização da Equação (28).

$$S(\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q) = \sum_{t=1}^{N} \hat{a}_t^2 \qquad 28$$

sendo:

$$\hat{a}_t = \theta^{-1}(B)\, \theta(B) M_t;$$

A soma dos quadrados depende das observações passadas de M_t e de a_t assim como os valores dos parâmetros desconhecidos $S(\hat{\phi}, \hat{\theta})$, para o qual a soma dos quadrados é mínima (FAVA, 1999).

Para qualquer conjunto de observações de séries temporais, a função de verossimilhança L é definida como a densidade da probabilidade conjunta de se obter os dados realmente observados, sendo esta representada pela Equação (29). Em relação aos modelos ARIMA, L será um a função de ϕ, θ e σ_a^2 para as N observações de $Z_1, \dots Z_N$ (CRYEN; CHAN, 2008). Os estimadores de máxima verossimilhança são, então, definidos como aqueles valores dos parâmetros para que os dados realmente observados são mais prováveis, isto é, os valores que maximizam a função de verossimilhança dada pela Equação (29).

$$L(\phi, \theta, \sigma_a^2/M) = [2\pi\sigma_a^2]^{-\frac{N}{2}} |\Omega|^{-\frac{1}{2}} exp\left[-\frac{1}{2} M' \Omega^{-1} M/\sigma_a^2\right] \qquad 4.11\ 29$$

A solução é obtida numericamente, sendo que a maximização de L é complexa pela existência da matriz Ω.

É fácil notar que o processo de estimação demanda um trabalho muito ardo para qualquer que seja o método adotado, ou seja, é necessário o uso do computador. Felizmente há muitos pacotes econometricos para microcomputadores que estimam os modelos ARIMA.

4.4.3 Verificação

Após estimar o modelo, o próximo passo da metodologia de Box & Jenkins é verificar se o mesmo representa, ou não, de forma adequada os dados. Em caso positivo, já é possível fazer a previsão, se não, outra especificação deve ser escolhida para modelar a série, o que implica em refazer as etapas de identificação e estimação.

Existem vários testes de verificação para um dado modelo ajustado a uma série. Em geral, tais testes são baseados nas autocorrelações estimadas dos resíduos.

No teste de autocorrelação residual, se o modelo for adequado os erros devem ser não-correlacionados, os resíduos do modelo estimado, $\hat{a}_t$, deverão estar próximos do ruído branco a_t, sendo $\hat{a}_t$ uma aproximação de a_t. Se $\hat{r}_j$ indicarem as autocorrelações dos resíduos $\hat{a}_t$, assim seus coeficientes de autocorrelação devem ser estatisticamente iguais a zero. As autocorrelações $\hat{r}_j$ são calculadas pela Equação (30) (FAVA, 1999).

$$\hat{r}_j = \frac{\sum_{t=j+1}^{N} \hat{a}_t \hat{a}_{t-j}}{\sum_{t=1}^{N} \hat{a}_t^2}. \qquad 30$$

Contudo, conforme ressaltou Durbin (1970), para valores pequenos de j, a variância de $\hat{r}_j$ pode ser bem menor do que $1/\sqrt{N}$. Para valores moderados ou grandes de j, a distribuição é válida e pode- se realizar testes de hipóteses e construir intervalos de confiança para avaliar a significância de cada $\hat{r}_j$.

Para o teste conjunto utiliza-se a estatística de Box e Pierce (1970), que apesar de não detectar quebras especificas no comportamento do ruído branco, pode indicar se esses valores são muito altos.

4.5 Modelos Sazonais

Os modelos ARIMA exploram a correlação de valores de Z_t observados em instantes de tempo consecutivos, enquanto os modelos ARIMA que possuem sazonalidade apresentam correlação entre os instantes de tempo distantes entre si, por m ou múltiplos de m, em que m é o número de observações contidas em um ano (m = 12 para dados mensais e m = 4 para dados trimestrais) (FAVA, 1999).

Para acomodar esse tipo de série, surgem os modelos ARIMA sazonais, também conhecidos por modelos SARIMA, os quais são, na verdade, extensões dos modelos ARIMA.

Quando um modelo apresenta sazonalidade (BROCKWELL, 1990), o objetivo é ajustar uma série para componente sazonal, ou seja, estimar S_t e subtrair a série estimada Z_t no modelo:

$$Z_t = T_t + S_t + a_t, \qquad t = 1,2, \dots, N \tag{31}$$

Em que T_t e S_t apresentam a tendência e a sazonalidade, respectivamente. Logo, um procedimento de ajustamento sazonal consiste em:

i) Obter estimativas $\widehat{S}_t$ de S_t;
ii) Calcular

$$Z_t^{SA} = Z_t - \widehat{S}_t \tag{32}$$

Se o modelo for multiplicativo, da forma (33).

$$Z_t = T_t S_t a_t \tag{33}$$

a série sazonal será ajustada conforme a Equação (34).

$$Z_t^{SA} = Z_t / \widehat{S}_t. \tag{34}$$

Existem vários procedimentos para estimar S_t, onde os mais usuais são o método de regressão e o método de médias móveis.

O método de regressão é apropriado para sereis que apresentam sazonalidade determinística, ou seja, que podem ser previstas perfeitamente a partir de meses anteriores. Já o método de médias móveis é apropriado para séries que apresentam sazonalidade estocástica, ou seja, para uma serie temporal cuja componente sazonal varia com o tempo (BROCKWELL, 1990).

Obtendo os modelos ARIMA para a tendência e sazonalidade, teremos um modelo conjunto chamado modelo sazonal multiplicativo de ordem $(p, d, q)(P, D, Q)_m$, conhecido como SARIMA$(p, d, q)(P, D, Q)_m$, que será denotado pela Equação (35).

$$\phi(B)\Phi(B^m)\Delta^d\Delta_m^D Z_t = \theta(B)\Theta(B^m)a_t. \qquad 35$$

em que:

- $\Phi(B^m) = 1 - \Phi_1 B^m - 1 - \Phi_2 B^{2m} - \cdots - 1 - \Phi_P B^{Pm}$, é o operador autoregressivo sazonal de ordem P;
- $\Theta(B^m) = 1 - \Theta_1 B^m - 1 - \Theta_2 B^{m2} - \cdots - 1 - \Theta_Q B^{mQ}$, é o operador de médias móveis sazonal de ordem Q;
- $\Delta_m = 1 - B^m$ é o operador diferença sazonal, tal que

$$\Delta_m Z_t = Z_t - Z_{t-m} \qquad 36$$

no qual:

Δ_m^D	=	$(1 - B^m)^D$;
D	=	número de diferenças sazonais;
m	=	período da sazonalidade

Para identificação, verificação e estimação de modelos sazonais, não há em principio nenhuma dificuldade adicional. A diferença é que temos que diferenciar a série com respeito em Δ e Δ_m a fim de produzir estacionariedade. Com isso obtemos os valores de d e D, que na maioria das vezes assumem valores no máximo igual a 2 (MORETTIN; TOLOI, 2006).

Depois, inspecionamos as FAC e FACP amostrais da serie adequada diferençada, nos ”lags” 1, 2, 3,.., para obter os valores de p e q e os “lags” 12, 24, 36,..., para obter os valores P e Q, selecionando o modelo tentativo.

As fases de estimação e verificação dos modelos SARIMA$(p,\ d,\ q)(P, D, Q)_m$ são análogos aos do modelo ARIMA$(p,\ d,\ q)$.

5 Modelo Híbrido de Regressão (ARIMA) e Rede Neural

5.1 Introdução

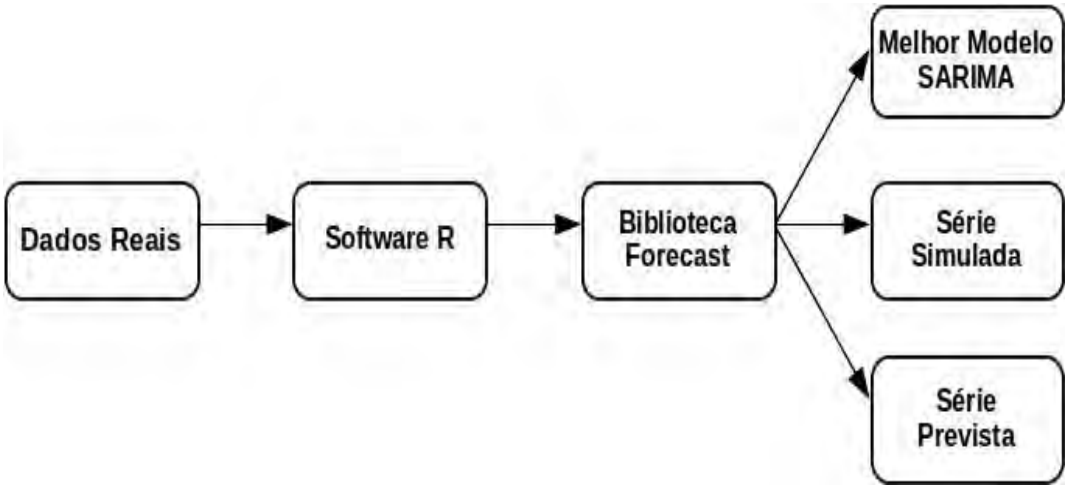
A metodologia proposta neste trabalho é composta pelos modelos ARIMA de Box & Jenkins e pelas RNAs, sendo a PMC a rede utilizada com algoritmo de treinamento de Levenberg-Marquartd, para fins de previsão de cargas elétricas a curto prazo.

5.2 Descrição do Modelo Híbrido

O modelo híbrido consiste em aplicar os dados históricos (reais) no software R utilizando a livraria forecast (CRYEN; CHAN, 2008; HYNDMAN; KHANDAKAR, 2008), com o objetivo de obter o melhor modelo ARIMA para série temporal, que em questão é carga elétrica. Os modelos obtidos apresentam sazonalidade, pois apresentam correlações em instantes de tempo distantes entre si por 24 horas. Assim, a correlação entre as cargas no período da noite devem ser altas e provavelmente maiores do que a correlação das cargas da manhã do mesmo dia. Tais modelos são conhecidos como modelos SARIMA, os quais são na verdade extensões dos modelos ARIMA.

Na Figura 3, apresenta-se detalhadamente o que se obtém através do software R.

Figura 3 - Diagrama de blocos do software R



Fonte: Elaboração da própria autora

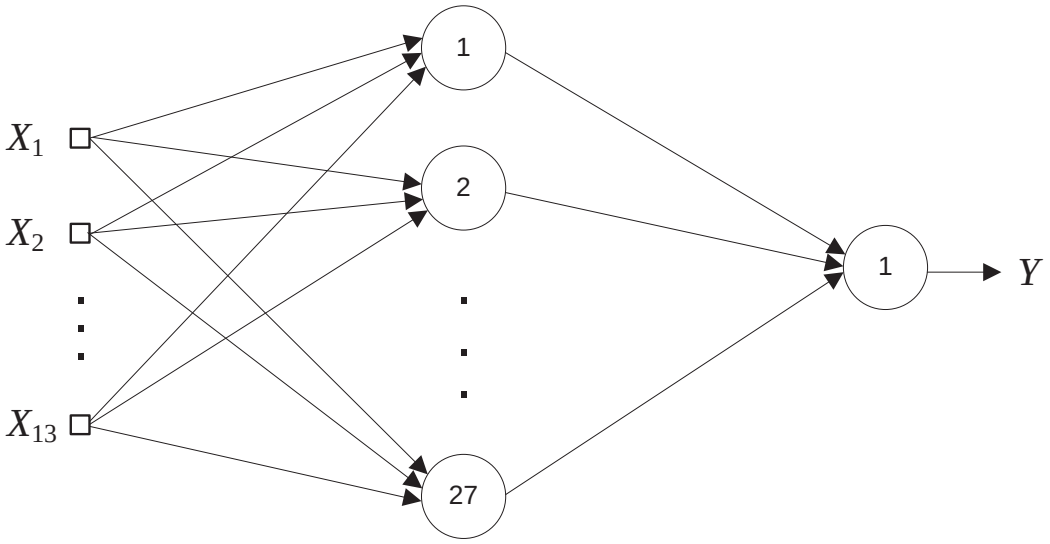
É importante enfatizar que a utilização do software R torna os procedimentos mais simples, visto que é capaz de realizar todas as fases da metodologia de Box & Jenkins com apenas a entrada dos dados da série temporal, que neste caso são cargas elétricas.

Através da série simulada do modelo SARIMA, já é possível encontrar o erro que existe entre tal modelo e os dados reais:

$$Erro = dados\ reais - dados\ simulados \qquad 37$$

O erro encontrado em (37) será utilizado como quatro entradas da rede neural, que terá a estrutura apresentada na Figura 4 para as três aplicações:

Figura 4 - Estrutura da rede neural



Fonte: Elaboração da própria autora

O conjunto que compõem o vetor de entrada e a saída, respectivamente, é definido da seguinte forma:

$$X_h = [\ te\ \ E(h-3)\ \ E(h-2)\ \ E(h-1)\ \ E(h)]^T\ ,\ \ X \in\ R^b \qquad 38$$

$$Y = [E(h+1)]\ ,\ \ Y \in\ R^1 \qquad 39$$

sendo:

b	=	dimensão do vetor X;
$E(h-v)$	=	valor do erro v horas anterior à hora corrente h;
$E(h+1)$	=	valor do erro correspondente à hora subseqüente a hora corrente h;
te	=	vetor de tempo referente aos dados históricos (feriado, dia da semana, horas) codificados em números binários (-1, +1).

Em relação à Figura (4), a rede neural possui 3 camadas, sendo que a camada de entrada é composta por 13 neurônios, a camada intermediária por 27 e a camada de saída por 1 neurônio.

Na camada de entrada, o vetor te representa 9 entradas da rede neural que é composto por dados binários correspondentes aos feriados, dias da semana e horas do dia, respectivamente, já as quatro ultimas entradas correspondem aos erros obtidos em (37). A saída da rede neural irá fazer a previsão do erro, ou seja, o erro futuro das 24 horas que se deseja prever, que será denotado por $Erro\ Futuro$.

Juntando o erro, Equação (37), com o $Erro\ Futuro$ previsto através da rede neural será obtido o vetor erro, que será chamado de E_t e representado pela Equação (40).

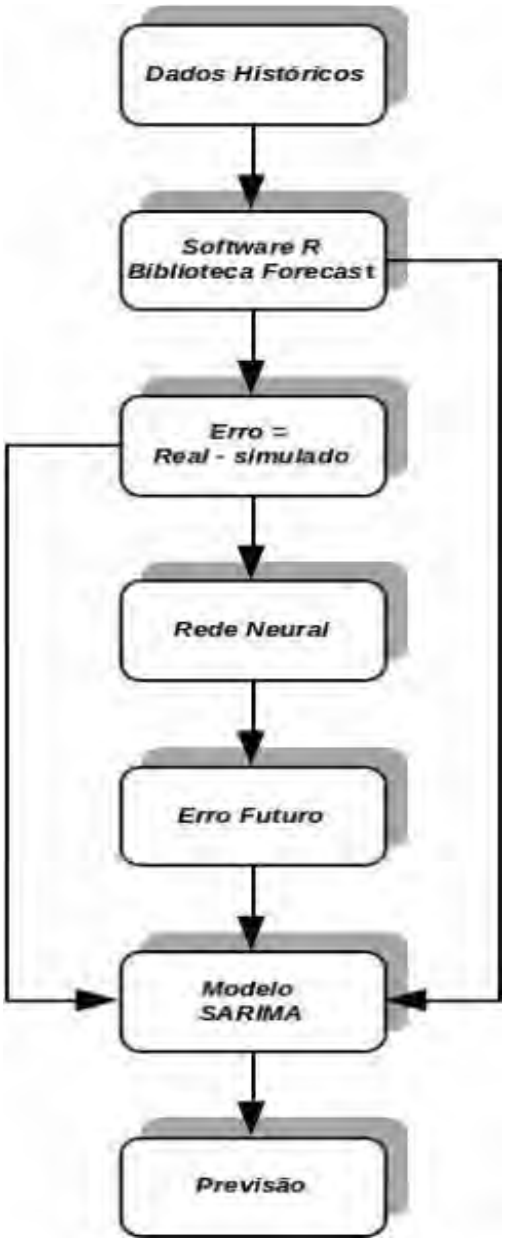
$$E_t = Erro + Erro\ Futuro \qquad 40$$

O erro E_t será substituído nos resíduos a_t, da equação SARIMA gerada pelo software R, para finalmente obter a previsão desejada da carga Z_t.

É importante enfatizar que as fases de treinamento e teste da rede neural foram realizadas no software MATLAB utilizando as ferramentas toolbox neural network (ORTIGUEIRA, 2000). Utilizou-se também o software MATLAB para realizar a previsão da carga elétrica, sendo esta encontrada depois que o vetor erro, E_t, é substituído na equação SARIMA, Equações (42), (44) e (46).

A Figura 5, explica através de um diagrama de blocos o modelo proposto.

Figura 5 - Diagrama de blocos da metodologia proposta



Fonte: Elaboração da própria autora

6 Aplicações e Resultados

Este capítulo apresenta alguns resultados obtidos através do modelo proposto sendo que o mesmo visa resolver o problema de previsão de cargas elétricas. Para testar a eficiência da metodologia híbrida, os resultados obtidos serão comparados com várias arquiteturas de redes neurais do artigo (NOSE FILHO et al., 2009) e também com os modelos SARIMA obtidos para cada aplicação. Os períodos adotados são descritos na Tabela 1.

Tabela 1 - Períodos utilizados

Aplicação	Período	Número de Vetores	Dia Previsto
1	31 dias - 01/07/98 a 31/07/98	744	01/08/1998
2	61 dias - 01/06/98 a 31/07/98	1464	01/08/1998
3	84 dias - 01/05/98 a 31/07/98	2208	01/08/1998

Fonte: Elaboração da própria autora

Os modelos obtidos através do software R para cada aplicação foram:

- Aplicação 1: SARIMA$(2,0,0)(2,0,2)_{24}$, que é representado pela Equação (42):

$$\phi(B)\Phi(B^{24})Z_t = \Theta(B^{24})a_t \qquad 41$$

substituindo os valores dos parâmetros e coeficientes:

$$(1 - 1{,}3641.B + 0{,}3879.B^2)(1 - 0{,}2462.B^{24} - 0{,}7523.B^{48})Z_t = (1 + 0{,}0971.B^{24} - 0{,}7523.B^{48})a_t \qquad 42$$

- Aplicação 2: SARIMA$(2,0,2)(2,0,2)_{24}$, sendo representada pela Equação (44):

$$\phi(B)\Phi(B^{24})Z_t = \theta(B)\Theta(B^{24})a_t \qquad 43$$

substituindo os valores dos parâmetros e coeficientes:

$$(1 - 1{,}6914.B + 0{,}7166.B^2)(1 - 0{,}5885.B^{24} - 0{,}4105.B^{48})Z_t = (1 - 0{,}3069.B - 0{,}2185.B^{48})(1 - 0{,}4444.B^{24} - 0{,}4060.B^{48})a_t \quad 44$$

- Aplicação 3: SARIMA$(3,0,1)(2,0,2)_{24}$, sendo representada pela Equação (46):

$$\phi(B)\Phi(B^{24})Z_t = \theta(B)\Theta(B^{24})a_t \quad 45$$

substituindo os valores dos parâmetros e coeficientes:

$$(1 - 0{,}3999.B - 0{,}8949.B^2 + 0{,}3808.B^3)(1 - 0{,}7955.B^{24} - 0{,}2037.B^{48})Z_t = (1 - 0{,}9779.B)(1 - 0{,}5533.B^{24} - 0{,}2837.B^{48})a_t \quad 46$$

Uma vez gerado os modelos SARIMA através do software R, obtêm-se os valores de erros, Equação (37), que serão utilizados no treinamento e diagnóstico da rede neural (Perceptron Multicamadas com treinamento de Levenberg-Marquardt). A especificação dos melhores parâmetros utilizados pela rede neural para cada aplicação é descrita na Tabela 2.

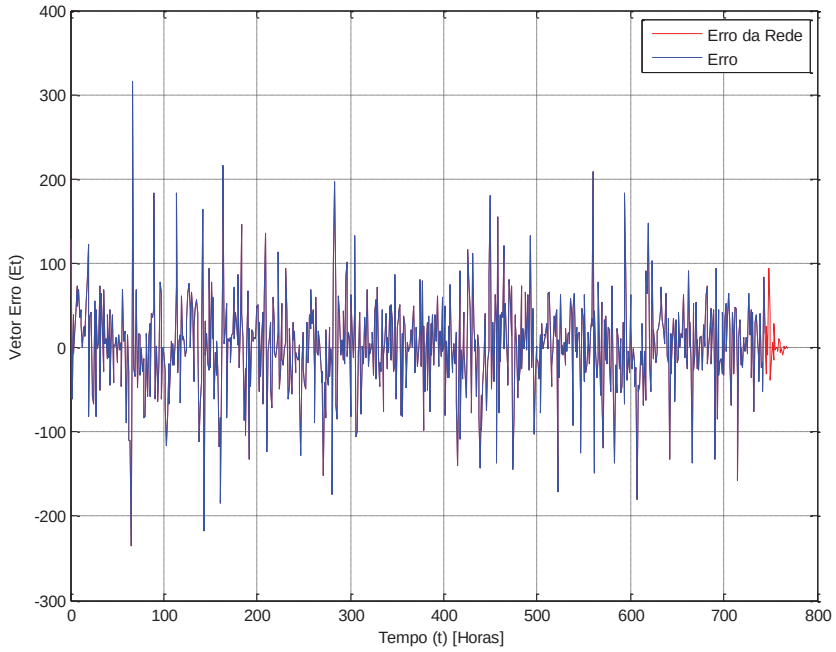
Tabela 2 - Parâmetros da rede neural

Parâmetros	Aplicação 1	Aplicação 2	Aplicação 3
Número de camadas	3	3	3
Número de neurônios por camada	13-27-1	13-27-1	13-27-1
Meta	0,001	0,001	0,001
Número Máximo de Iterações	39	45	67
Mínimo Gradiente	10^{-10}	10^{-10}	10^{-10}
Coeficiente de Levenberg-Marquardt	0,12	0,15	0,20
Coeficiente Levenberg-Marquardt Decremento	0,12	0,15	0,20
Coeficiente Levenberg-Marquardt Incremento	10	10	10
Coeficiente de Levenberg-Marquardt Máximo	10^{10}	10^{10}	10^{10}

Fonte: Elaboração da própria autora

As Figuras 6, 7 e 8 ilustram o vetor erro, E_t, que serão utilizados novamente nos modelos SARIMA com a finalidade de obter a previsão da carga elétrica a curto prazo. Observa-se que o vetor E_t contém todos os valores de erros obtidos através do modelo SARIMA e da carga elétrica mais os erros previstos pela rede neural. Os valores de E_t utilizados nas aplicações 1, 2 e 3 são mostrados pelas Figuras 6, 7 e 8, respectivamente.

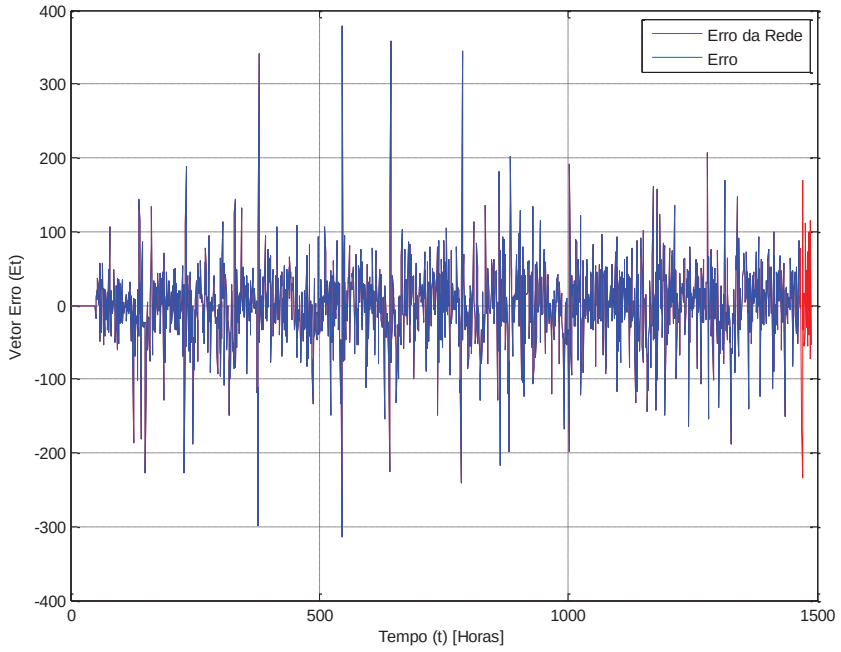
Figura 6 - Vetor Erro (Aplicação 1)



Fonte: Elaboração da própria autora

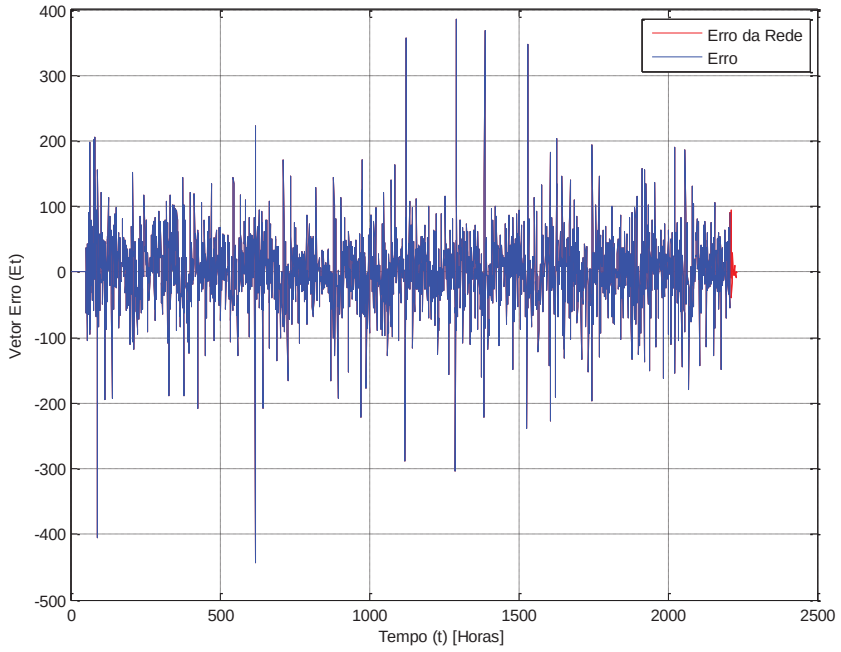
Uma vez substituído o vetor erro, E_t, em a_t nos modelos SARIMA, sendo estes obtidos para cada aplicação, Equações (42), (44) e (46) têm-se então os valores de Z_t, que são os resultados obtidos através de previsão de cargas a curto prazo.

Figura 7 - Vetor Erro (Aplicação 2)



Fonte: Elaboração da própria autora

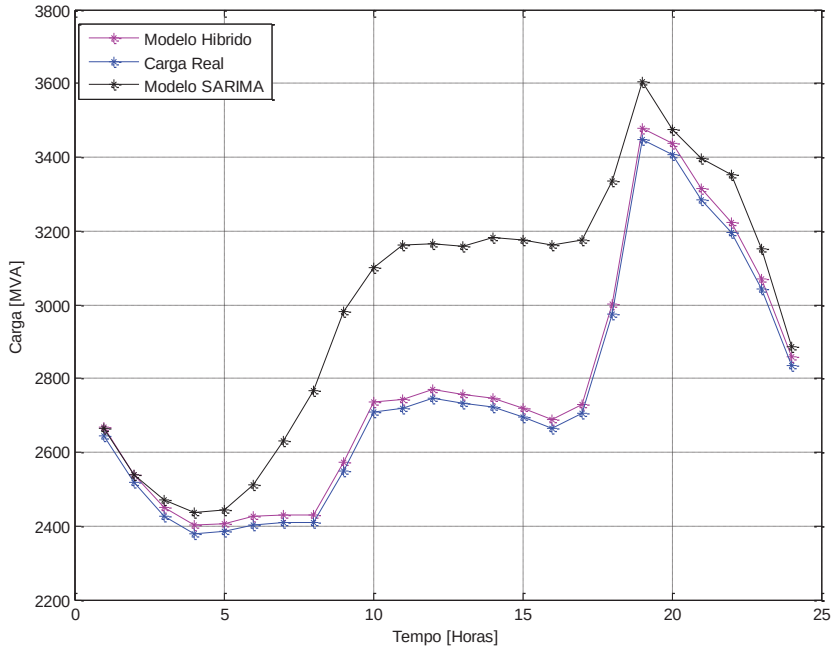
Figura 8 - Vetor Erro (Aplicação 3)



Fonte: Elaboração da própria autora

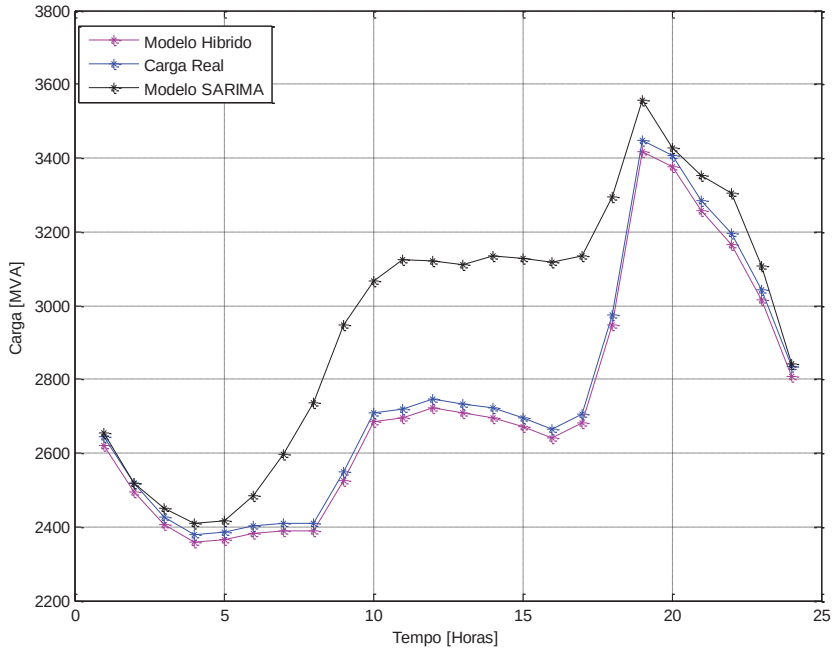
As Figuras 9, 10 e 11 apresentam os resultados encontrados para previsão do dia 1 de agosto de 1998.

Figura 9 - Resultado da previsão de 24h para aplicação 1



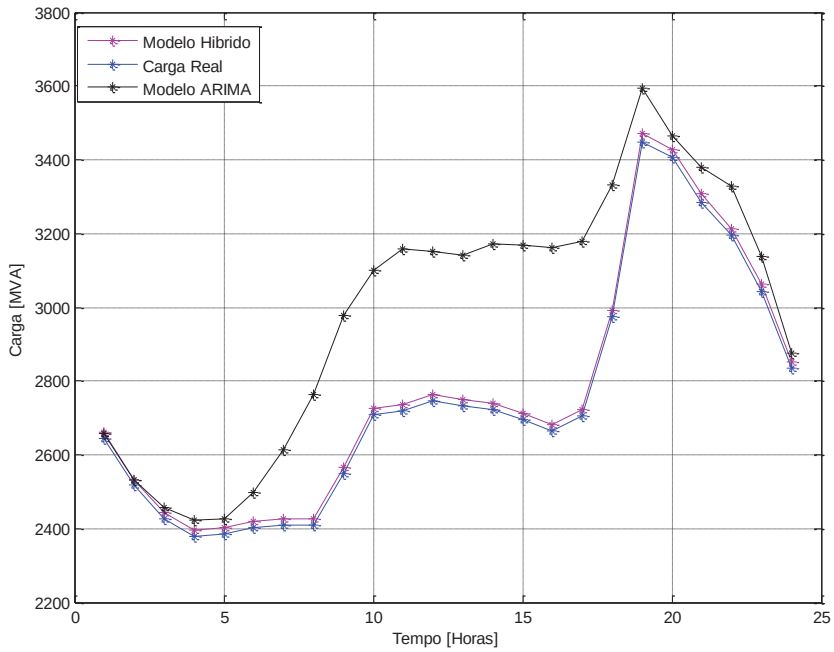
Fonte: Elaboração da própria autora

Figura 10 - Resultado da previsão de 24h para aplicação 2



Fonte: Elaboração da própria autora

Figura 11 - Resultado da previsão de 24h para aplicação 3



Fonte: Elaboração da própria autora

Na avaliação da precisão serão calculados o erro percentual absoluto médio (MAPE) e o erro máximo da previsão, comparando-se os valores reais da carga com os valores estimados pelo modelo híbrido da seguinte forma:

$$MAPE = \frac{1}{NT}\sum_{h=1}^{N}\left\{\frac{\left|L(h) - \underline{L}(h)\right|}{L(h)}\right\}.100\% \qquad 47$$

$$Erro\ máximo\ (\%) = máx\left\{\frac{\left|L(h) - \underline{L}(h)\right|}{L(h)}\right\}.100\% \qquad 48$$

sendo:

$L(h) =$ valor da carga real referênte a hora h;

$\underline{L}(h) =$ valor da carga estimada pelo modelo híbrido referente a hora h;

$NT =$ número total de horas.

Para analisar os resultados obtidos, verificamos o MAPE e o maior erro obtido durante a previsão das cargas nas 24 horas do modelo proposto, comparando-o com as diferentes arquiteturas de redes neurais do trabalho de (NOSE FILHO et al., 2009) e também com o modelo SARIMA obtido. Os resultados são mostrados nas Tabelas 3, 4 e 5, para as aplicações 1, 2 e 3 respectivamente.

Tabela 3 - MAPE e Erro Máximo para aplicação 1

Modelos	MAPE (%)	Erro máximo (%)
Modelo Híbrido	0,908	1,016
SARIMA	9,096	18,623
RNA *feedforward* com algoritmo gradiente descendente com momento	1,998	4,845
RNA *feedforward* com algoritmo Levenberg-Marquardt	1,178	3,560
ANFIS	1,090	2,861
GRNN	1,003	2,410

Fonte: Elaboração da própria autora

Tabela 4 - MAPE e Erro Máximo para aplicação 2

Modelos	MAPE (%)	Erro máximo (%)
Modelo Híbrido	0,892	0,954
SARIMA	7,774	16,955
RNA *feedforward* com algoritmo gradiente descendente com momento	2,093	5,022
RNA *feedforward* com algoritmo Levenberg-Marquardt	1,490	4,267
ANFIS	1,403	3,848
GRNN	1,118	3,650

Fonte: Elaboração da própria autora

Tabela 5 - MAPE e Erro Máximo para aplicação 3

Modelos	MAPE (%)	Erro máximo (%)
Modelo Híbrido	0,634	0,875
SARIMA	8,747	18,630
RNA *feedforward* com algoritmo gradiente descendente com momento	2,313	7,088
RNA *feedforward* com algoritmo Levenberg-Marquardt	1,698	4,687
ANFIS	0,973	3,757
GRNN	1,288	4,246

Fonte: Elaboração da própria autora

7 Conclusões e Sugestões para Futuros Trabalhos

7.1 Conclusões

As técnicas híbridas entre ARIMA e RNAs vêem tendo destaque na previsão de séries temporais e resultados bastante satisfatórios estão sendo encontrados. Este trabalho teve como objetivo o desenvolvimento de uma metodologia híbrida para previsão de cargas elétricas a curto prazo, sendo utilizados os modelos ARIMA de Box & Jenkins e as RNAs, com o treinamento baseado no algoritmo de Levenberg-Marquartd.

No modelo proposto, os modelos ARIMA foram responsáveis por gerar os parâmetros da carga (o melhor modelo) e realizar a previsão desejada. Tal previsão é feita depois que a RNA realizou a previsão do erro, sendo este juntamente com o erro encontrado através da diferença entre os dados reais e os dados simulados, substituídos no modelo ARIMA encontrado, para enfim realizar a previsão.

Foram realizados três testes para a previsão de cargas elétricas considerando-se uma base de dados históricos de uma companhia do setor elétrico brasileiro. A previsão de carga a curto-prazo foi realizada considerando-se 24 horas à frente, em que o modelo proposto obteve um excelente desempenho em todos os testes, com erros médios abaixo de 1%. Assim, pode-se concluir que o modelo foi capaz de realizar previsões para períodos distintos que também envolveram dias atípicos e finais de semana com confiança e alto desempenho. É importante observar que ao passo em que aumentamos o número de padrões de entrada, o erro da previsão diminui e com isso melhores resultados são obtidos.

A previsão de cargas a curto prazo aqui mencionadas, apontaram uma eficiência no método proposto, e os resultados foram de importância no contexto de previsão de cargas elétricas encontrados na literatura.

7.2 Sugestões para trabalhos futuros

Para trabalhos futuros, algumas propostas podem ser feitas, como:

- aperfeiçoar os resultados obtidos acrescentando outros fatores que influenciam o comportamento da carga que não somente os da carga horária da série temporal usada anteriormente;

- utilizar outro algoritmo treinamento para a rede PMC a fim de produzir melhores resultados;
- utilizar o modelo proposto para realizar previsão de médio ou até mesmo longo prazo;
- utilizar outra rede neural sendo que esta poderá ser uma da família ART;
- utilizar para executar outras previsões como de afluências, preços, ventos, etc.

Referências

ABURTO, L.; WEBER, R. Improved supply chain management based on hybrid demand forecasts. **Applied Soft Computing**, Santiago, v. 7, n. 1, p. 136-144, 2007.

ACKLEY, D. H.; HINTON, G. E.; SEJNOWSKY, T. J. A Learning algorithm for Boltzmann machines. **Cognitive Sciences**, Hoboken, v. 9, n. 1, p.147-169, 1985.

ALMEIDA, C.; FISHWICH, P. A.; TANG, Z. Time series forecasting using neural network vs. Box-Jenkins Methodology. **Simulation Councies**, [S.l.], p. 303- 310, November 1991.

BARBOSA, A. H.; FREITAS, M. S. R.; NEVES, F. A. Confiabilidade estrutural utilizando o método de Monte Carlo e redes neurais. **Escola de Minas**, Ouro Preto, v. 58, n. 3, p. 247-255, 2005.

BOX, G.E.; JENKINS, G.M**. Times series analysis**: forecasting and control. San Francisco: Holden- Day, 1976. 575p.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. **Times series analysis**: forecast and control. 3. ed. Englewood Cliffs: Prentice Hall, 1994. 598 p.

BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations is autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, New York, v. 65, n. 332, p. 1509-1526, 1970.

BROCKWELL, P. J.; DAVIS, R. A. **Times series**: theory and methods. 2. ed. New York: Springer, 1990. 588p.

CRYEN, J. D.; CHAN, K. S. **Time series analysis**: with applications in R. 2. ed. New York: Springer, 2008. 501p.

O'DONOVAN, T. M. **Short term forecasting**: an introduction to the Box-Jenkins approach, New York: John Wiley & Sons, 1983. 256p.

DURBIN, J. Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. **Econometrica**, Chicago, v. 38, n. 30, p. 410-412, 1970.

FARUK, D.O. A hybrid neural network and ARIMA model for water quality time series prediction. **Engineering Applications of Artificial Intelligence**, Aydin, v. 23, n. 4, p. 586-594, 2009.

FAVA, V. L. Metodologia de Box-Jenkins para modelos univariados. **Manual de econometria:** nível intermediário. São Paulo: ATLAS, 1999, v. 1, p. 205-232.

GROSSBERG S. Competitive learning: from interactive activation to adaptive resonance. **Cognitive Science**, Boston, v.11, n. 1, p.23-63, 1987.

HAMILTON, J. D. **Time series analysis**. New Jersey: Princeton, 1994. 799p.

HAYKIN, S. **Neural networks**: a comprehensive foundation. 2. ed. New Jersey: Prentice-Hall, 1999. 842p.

HYNDMAN, R. J.; YEASMIN, K. Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, Alexandria, v. 27, n. 3, 22 p. 2008.

KHASHEI, M.; BIJARI, M; ARDALI, G. A. R. A. Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks. **Neurocomputing**, Isfahan, n. 4-6, v. 72, p. 956-967, 2009.

KHASHEI, M.; BIJARI, M. An artificial neural network (p, d, q) model for time series forecasting. **Expert Systems with Applications**, Isfahan, n. 1, v. 37, p. 479-489, 2010.

KHASHEI, M.; BIJARI, M. A novel hybridization artificial neural networks and ARIMA models for time series forecasting, **Applied Soft Computing**, Isfahan, n. 2, v. 11 p. 2664-2675, 2011.

KOHONEN, T. Self-organized formation of topologically correct of feature maps. **Biological Cybernetics**, Heidelberg, v. 22, p. 159-168, 1982.

KOUTROUMANIDIS, T.; IOANNOU, K.; ARABATZIS, G. Predicting fuel wood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA – ANN model. **Energy Police**, Orestiada, n. 9, v. 37, p. 3627-3624, 2009.

KROSE, B.; SMAGT, P. V. **An introduction to neural networks**. Amsterdam: University of Amsterdam, 1996. 135p.

LOPES, M. L. M. **Desenvolvimento de um sistema previsor de cargas elétricas via redes neurais**. 2000. 100 f. Dissertação (Mestrado) – Faculdade de Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho", Ilha Solteira, 2000.

LOPES, M. L. M. **Desenvolvimento de redes neurais para previsão de cargas elétricas de sistemas de energia elétrica**. 2005. 169 f. Tese (Doutorado) – Faculdade de Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho", Ilha Solteira, 2005.

LOTUFO, A. D. P.; MINUSSI, C. R. Electric power systems load forecasting: a survey. In: INTERNATIONAL CONFERENCE ON ELECTRIC POWER ENGINEERING - POWER-TECH, 1999, Budapest. **Proceedings…** Budapest: [s.n.], 1999. p. 36-41.

MCCULLOCH, W. S.; PITTS, W. H. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, Chicago, n. 4, v. 5, p. 115-133, 1943.

MENDES FILHO, E. F.; CARVALHO, A. C. P. L. **Tutorial introdutório sobre redes neurais artificiais**. São Paulo: Departamento de Ciências de Computação e Estatística – USP, 1997. Apostila.

MINUSSI, C. R.; LOTUFO, A. D. P. Redes neurais - introdução e principais conceitos. Ilha Solteira: Departamento de engenharia elétrica - Unesp, 2008. Apostila.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Edgard Blucher, 2006. 544 p.

NOSE-FILHO, K. et al. Previsão de cargas elétricas utilizando diferentes arquiteturas de redes neurais artificiais com o auxilio das toolboxes do matlab. In: LATIN-AMERICAN CONGRESS ON ELETRICITY GENERATION AND TRANSMISSION, 8., 2009, Ubatuba, SP. **Proceedings…** Ubatuba: [s.n.], 2009.

ORTIGUEIRA, M. D. **Manual de introdução ao MATLAB**. Lisboa: Faculdade de Ciências e Tecnologia da Universidade Nova Lisboa, 2000. 55p. Apostila.

OTHMAN, M. M. et al. The applications of Box-Jenkins models in short term load forecasting. **International Journal of Power, Energy and Artificial Intelligence**, Olney, v. 2, n. 1, p. 104-110, 2009.

PIERCE, D. A. Seasonal adjustment when both deterministic and stochastic seasonality are present. In: ZELLNER, A**. Seasonal Analysis of Economic Times Series**. Washington: NBER, 1979. p. 242-280.

RANGANATHAN, A. **The Levenberg- Marquartd algorithm**. [S.l: s.n.], 2004. 5p.

ROSEMBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, Washington, v. 65, n. 6, p. 386-408, 1958.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by Backpropagation errors. **Nature**, London, v. 323, n. 9, p. 533-536, 1986.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010. 399 p.

VALENZUELA, O. et al. Hybridization of intelligent techniques and ARIMA models for time series prediction. **Fuzzy Sets and Systems**, Granada, v.159, n.7, p. 821-845, 2008.

WEI, W. W. S. **Time series analysis:** univariate and multivariate methods. 2. ed. New York: Pearson, 2006. 634 p.

WERBOS, P. J. Beyond regression: **new tools for prediction and analysis in the behavioral sciences**. 1974. 102 f. Dissertação (Mestrado) - Harvard University, Harvard, 1974.

WIDROW, B.; LEHR, M. A. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. **Proceedings of the IEEE**, Piscataway, v. 78, n. 9, p. 1415-1442 1990.

ZHANG, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, Atlanta, v. 50, p. 159-175, 2003.

Apêndice A - Redes Neurais Artificiais

A.1 Reflexão Histórica

As redes neurais artificiais (RNAs) (HAYKIN, 1999) são modelos computacionais inspirados no sistema nervoso dos seres vivos e tem como principal característica simular o seu funcionamento.

As características mais importantes envolvidas com as aplicações das RNAs são a adaptação por experiência, capacidade de aprendizado, habilidade de generalização, organização de dados, tolerância a falhas e o armazenamento distribuído (SILVA et al., 2010).

Embora os primeiros trabalhos em redes neurais artificiais tenham sido publicados há mais de 50 anos, tal tema começou a ser fortemente pesquisado a partir de 1990, tendo até hoje um potencial de pesquisa atrativo, pois as RNAs são capazes de resolver problemas complexos como previsão de séries temporais, aproximações de funções e identificação de padrões, sem a necessidade da modelagem matemática do problema a ser resolvido.

A era moderna das redes neurais começou em 1943, com o trabalho pioneiro de McCulloch e Pitts (1943), que descreveu o primeiro neurônio artificial. Em 1958 Rosenblatt, apresentou seu trabalho sobre o perceptron, uma das primeiras RNAs, com apenas um neurônio e aprendizagem supervisionada. Em 1960, Widrow e Hoff introduziram o algoritmo do mínimo quadrado (LMS - Least Mean-Square) e o usaram para formular o Adaline (Adaptive linear element). A diferença entre o perceptron e o Adaline esta no procedimento de aprendizagem. Após dois anos Widrow propôs uma das primeiras RNAs de estrutura múltipla (Madaline, Multipleadaline). Nos anos 70 surgiram vários trabalhos a respeito dos mapas auto-organizaveis utilizando aprendizagem competitiva. Em 1976, surgem os primeiros trabalhos de Grossberg *(*1976a, 1976b) a respeito da teoria da ressonância adaptativa (ART - Adaptative Ressonance Theory). Nos anos 80 as redes neurais com realimentação atraiu muita atenção, e tornou-se conhecida como redes de Hopfield e os mapas auto-organizáveis de Kohonen (1982), utilizando uma estrutura de rede unidimensional ou bidimensional. Em 1985 surgiu a máquina de Boltzmann, sendo esta desenvolvida por Ackley et al. (1986), que foi a primeira realização bem sucedida de uma rede neural de múltiplas camadas. Em 1986, Rumelhart et al.(1986) apresentam o algoritmo retropropagação (backpropagation), primeiramente proposto por Werbos (1974). O algoritmo de treinamento retropropagação (backpropagation) é um dos mais utilizados até hoje. Grossberg (1987) apresentou sua primeira rede baseada na teoria da ressonância adaptativa em 1987, tal rede é conhecida como ART1.

Várias outras redes foram desenvolvidas nos últimos 20 anos, e o uso de técnicas híbridas vem sendo cada vez mais estudadas.

A.2 Modelo do neurônio biológico

Um neurônio artificial é baseado no funcionamento do neurônio biológico e uma rede neural se compõe por uma quantidade pré definida de neurônios artificiais que possibilita a solução de problemas de um número considerável de áreas distintas.

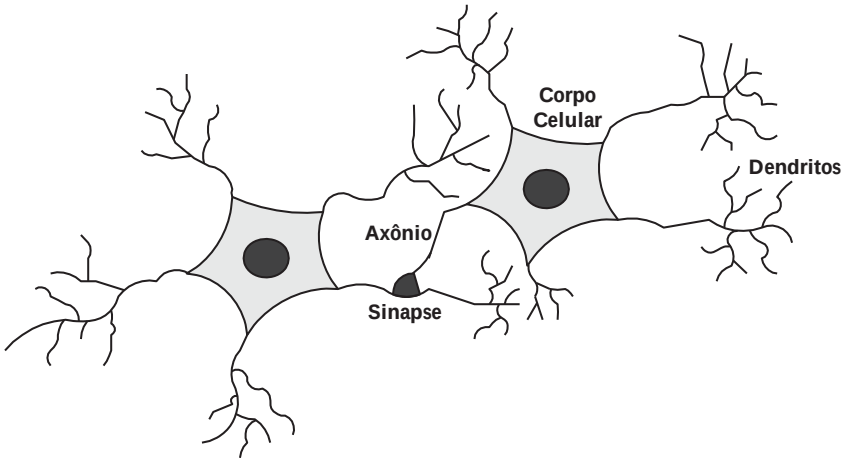
O cérebro humano pode ser considerado um processador complexo formado por bilhões de neurônios responsáveis pelo controle motor, da percepção, bem como do reconhecimento de padrões. A ligação desses neurônios se dá através de sinapses, as quais transmitem estímulos e que formam uma grande rede que estende o resultado pelo corpo humano.

O neurônio biológico é composto por:

- Dentritos: têm a função de receber estímulos de outros neurônios;
- Corpo celular (soma): coleta e combina informações recebidas;
- Axônio: composto por uma fibra tubular, que transmite estímulos a outras células;
- Sinapses: conecta o axônio aos dentritos de outros neurônios, propagando pulsos nervosos e excitando ou inibindo tais pulsos.

Na Figura 12 é mostrada a ilustração de um neurônio biológico.

Figura 12 - Neurônio biológico

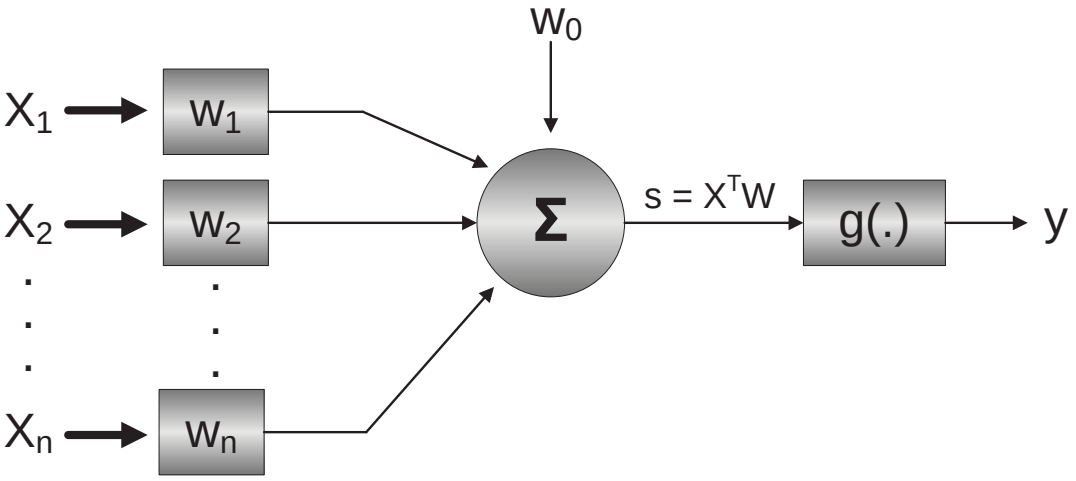


Fonte: Mendes Filho e Carvalho (1997)

A.3 Modelo do neurônio artificial

A Rede Neural Artificial é um processador paralelo composto por neurônios que possui a capacidade de armazenar conhecimento. Semelhante ao cérebro humano, a rede neural possui um neurônio artificial, que busca imitar as características do neurônio biológico. A Figura 13 ilustra o modelo de um neurônio artificial proposto por McCulloch e Pitts (1943) com n entradas.

Figura 13 - Modelo MCP (McCulloch-Pitts) não-linear de um neurônio artificial



Fonte: Elaboração da própria autora

O processo do treinamento de um neurônio artificial se desenvolve da seguinte forma (LOPES, 2005):

- Um conjunto de entradas X, $x_1, x_2, \ldots, x_n$, são aplicadas na camada de entrada do neurônio, onde elas representadas os sinais dentro de um neurônio biológico;
- Cada sinal é ponderado a um peso associado W, $w_1, w_2, \ldots, w_n$, que indica sua influência na saída da unidade;
- As entradas ponderadas são aplicadas ao bloco somatório, onde se adiciona também uma entrada independente, $x_0 = \pm 1$, denominada bias;
- Se a soma ponderada das entradas for maior que o valor de w_0 (limiar) o neurônio é ativado provocando um pulso na saída, como mostra a Equação (3.1). Caso contrário o neurônio não é ativado (WIDROW; LEHR, 1990).

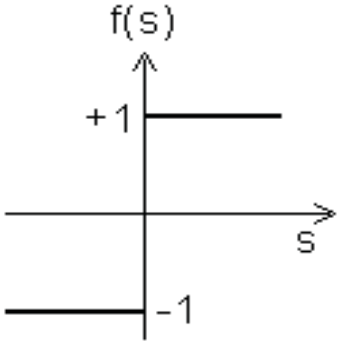
A saída intermediária do neurônio é, então, dada por:

$$s_i = \sum_{j}^{n} w_{ij} x_j \qquad 47$$

A ativação do neurônio artificial é feita através de uma função de ativação. As funções de ativação mais conhecidas e utilizadas são mostradas nas Figuras 14, 15, 16 e 17:

- Função degrau bipolar → $f_d(s) = \begin{cases} +1, & se\ s \geq 0 \\ -1, & se\ s \leq 0 \end{cases}$

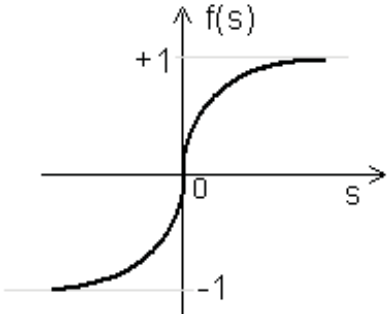
Figura 14 - Função degrau bipolar



Fonte: Minussi (2008)

- Função tangente hiperbólica → $f_t(s) = \frac{1-e^{-\lambda s}}{1+e^{-\lambda s}}$

Figura 15 - Função tangente hiperbólica

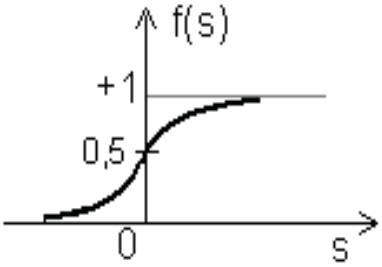


Fonte: Minussi (2008)

sendo λ a inclinação da curva.

- Função Logística → $f_l(s) = \frac{1}{1+e^{-\lambda s}}$

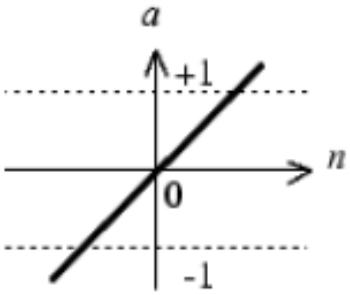
Figura 16 - Função Logística



Fonte: Minussi (2008)

- Função Linear → $f_{li}(n) = n$ para todo n

Figura 17 - Função Linear



Fonte: Minussi (2008)

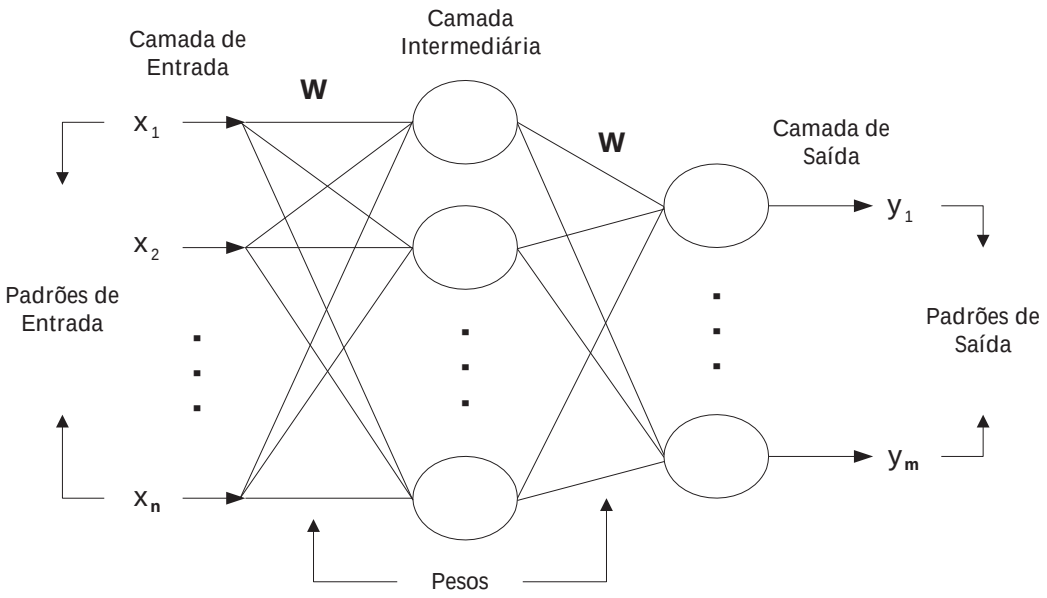
As funções tipo relé são apropriadas para sistemas binários, enquanto que as funções sigmoidais podem ser empregadas tanto para sistemas contínuos como binários (LOPES, 2000).

A.4 Estrutura da rede neural

As redes neurais artificiais são definidas como conjuntos estruturados de unidades de processamento denominadas neurônios, interligadas entre si que forma uma disposição estrutural de camadas e conexões entre as camadas.

Basicamente as redes neurais são formadas por três camadas: camada de entrada, camada intermediaria (oculta, escondida ou invisível) e camada de saída (HAYKIN, 1999) como ilustra a Figura 18.

Figura 18 - Disposição das redes neurais artificiais



Fonte: Lopes (2005)

- Camada de entrada: é responsável por receber informações dos dados;
- Camada intermediaria: é responsável por extrair as características associadas ao processo ou sistema a ser inferido. Quase todo o processamento interno da rede é realizado nessa camada. O número de camadas escondidas e a quantidade de neurônios nessa camada dependem principalmente da complexidade do problema a ser mapeado pela rede;
- Camada de saída: é também constituída de neurônios, sendo esta responsável pela apresentação dos resultados obtidos pela rede.

As principais arquiteturas das RNAs em relação à propagação dos dados podem ser classificadas em: redes feedforward, redes recorrentes e redes reticuladas.

- Redes feedforward: também são conhecidas como redes não recorrentes, os dados fluem das unidades de entrada para as unidades de saída, ou seja, a saída é exclusivamente determinada em função das entradas e dos valores dos pesos. O processamento dos dados pode se estender sobre múltiplas unidades (camadas). As principais redes cuja arquitetura é do tipo feedforward são o Perceptron multicamadas (Multilayer Perceptron – MLP) e as redes de base radial (radial basis function - RBF) (KROSE; SMAGT, 1996);

- Redes recorrentes: contém conexões feedback, e desenvolvem uma memória ao longo prazo nos neurônios internos. Devido à realimentação, as redes com essa arquitetura produzem saídas levando-se em consideração os valores das saídas anteriores recentes. A principal rede que possui realimentação é denominada rede de Hopfiled (KROSE; SMAGT, 1996);
- Redes Reticuladas: Tem como principal característica considerar a disposição espacial dos neurônios visando à extração de características, isto é, a localização espacial dos neurônios está relacionada com o processo de ajuste de seus pesos e limiares. A principal rede desse tipo de arquitetura é o mapa auto-organizável de Kohonen (SILVA et al. , 2010).

A.5 Treinamento da rede neural

A propriedade mais importante das redes neurais é a habilidade de aprender e com isso melhorar seu desempenho. A rede é treinada de maneira que um conjunto de entrada produza o conjunto de saída desejada ou que pelo menos seja consistente (MENDES FILHO; CARVALHO, 1997).

A aprendizagem das RNAs ocorre na fase de treinamento. Os tipos de treinamento mais utilizados são os supervisionados e não supervisionado (SILVA et al, 2010).

- Treinamento supervisionado: consiste em um método de aprendizagem em que as combinações dos padrões de entrada e saída são arbitrados por um tutor ou professor na fase de aprendizado, como é o caso da rede Perceptron Multicamadas;
- Treinamento não supervisionado: não possui tutor, ou seja, a rede é auto organizada, sendo capaz de descobrir estaticamente, padrões relevantes aos dados de entrada, o que ocorre na rede ART;
- Treinamento com reforço: o conjunto de treinamento é formado somente por entradas, no entanto, há um elemento externo que, em vez de retornar o erro de saída da rede, retorna um sinal de satisfatório ou não em relação à última ação da rede. Caso a ação não seja satisfatória, ela será reforçada, tendo menor chance de ocorrer futuramente. Caso contrário, se a ação resultou em uma melhora no desempenho, ou seja, foi satisfatória, ela será reforçada aumentando a probabilidade de ocorrência da mesma no futuro, servindo de exemplo a rede de Kohonen.