**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO**

LUCAS ASSIS DE MORAES

**Desenvolvimento de uma abordagem *fuzzy* para
estimação de demanda de potência em um sistema de
distribuição de energia elétrica**

São Carlos
2014

LUCAS ASSIS DE MORAES

**Desenvolvimento de uma abordagem *fuzzy* para
estimação de demanda de potência em um sistema de
distribuição de energia elétrica**

Dissertação de Mestrado apresentada ao
Departamento de Engenharia Elétrica e de
Computação da Universidade de São Paulo
como parte dos requisitos para a obtenção do
título de Mestre em Ciências.

Área de Concentração: Sistemas Dinâmicos.

ORIENTADOR: Prof. Dr. Rogério Andrade
Flauzino

São Carlos
2014

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

M827d Moraes, Lucas Assis de
Desenvolvimento de uma abordagem fuzzy para
estimação de demanda de potência em um sistema de
distribuição de energia elétrica / Lucas Assis de
Moraes; orientador Rogério Andrade Flauzino. São
Carlos, 2014.

Dissertação (Mestrado) - Programa de Pós-Graduação
em Engenharia Elétrica e Área de Concentração em
Sistemas Dinâmicos -- Escola de Engenharia de São
Carlos da Universidade de São Paulo, 2014.

1. Sistemas inteligentes. 2. Distribuição de
energia elétrica. 3. Sistemas de inferência fuzzy. 4.
Estimação de séries temporais. 5. Modelagem e simulação
de sistemas dinâmicos. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **LUCAS ASSIS DE MORAES**.

Título da dissertação: "Desenvolvimento de uma abordagem fuzzy para estimação de demanda de potência em um sistema de distribuição de energia elétrica".

Data da defesa: 01/08/2014

Comissão Julgadora:

Resultado:

Prof. Dr. **Rogério Andrade Flauzino (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

Aprovado

Prof. Dr. **Adrião Duarte Dória Neto**
(Universidade Federal do Rio Grande do Norte/UFRN)

APROVADO

Prof. Associado **Alexandre Cláudio Botazzo Delbem**
(Instituto de Ciências Matemáticas e Computação/ICMC)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Titular **Denis Vinicius Coury**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Paulo César Lima Segantine**

[Faint, illegible text and signature at the bottom of the page]

Aos meus familiares.

AGRADECIMENTOS

À minha família, simplesmente por tudo.

Aos meus amigos, antigos ou recentes, de perto ou de longe, com os quais sei que posso contar.

Ao Professor Rogério Andrade Flauzino pela orientação.

À CAPES e à FAPESP pelo auxílio.

A todos os funcionários do Departamento de Engenharia Elétrica e de Computação pela ajuda, atenção e apoio que me foram dados.

RESUMO

MORAES, L. A. **Desenvolvimento de uma abordagem *fuzzy* para estimação de demanda de potência em um sistema de distribuição de energia elétrica.** 2014. 74 f. Dissertação (Mestrado) – Departamento de Engenharia Elétrica e de Computação, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2014.

Este trabalho tem por objetivo desenvolver uma abordagem *fuzzy* focando na estimação de curto prazo da demanda de potência ativa de um alimentador de sistema de distribuição de energia elétrica. A motivação para este trabalho encontra-se na redução do erro de estimação para que o sistema de distribuição como um todo seja corretamente operado. O destaque da abordagem desenvolvida é a metodologia de seleção de entradas para o sistema de estimação, que o treina fornecendo-lhe informações não redundantes e não desnecessárias sobre o comportamento da série temporal. Os resultados, obtidos com treinamento e teste de um sistema de inferência *fuzzy* multicamadas, mostram que as estimações realizadas selecionando as entradas do sistema de forma criteriosa apresentam menor erro que quando não há critério de seleção. Conclui-se então que a metodologia foi funcional e eficiente para o caso estudado, o que faz com que este trabalho resulte em válidas contribuições nas áreas de sistemas inteligentes, de sistemas dinâmicos e inclusive na forma metodológica de especificação de modelos de estimação de séries temporais.

Palavras-chave: Sistemas inteligentes. Distribuição de energia elétrica. Sistemas de inferência *fuzzy*. Estimação de séries temporais. Modelagem e simulação de sistemas dinâmicos.

ABSTRACT

MORAES, L. A. **Development of a fuzzy approach for power demand forecast in an electrical energy distribution system.** 2014. 74 f. Dissertação (Mestrado) – Departamento de Engenharia Elétrica e de Computação, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2014.

This work aims to develop a fuzzy approach focusing on the short-term active power demand forecast in a feeder of an electrical energy distribution system. This work motivation lies on the reduction of the forecast error so that the whole distribution system can be correctly operated. The highlight of the developed approach is the methodology to select the inputs for the estimation system, which trains it giving to it non-redundant and non-unnecessary information about the time series behavior. The results, obtained by training and testing a multilayer fuzzy inference system, show that the estimations made by following a criterion to select the inputs have smaller error than when there is no selection criterion at all. It is therefore concluded that the methodology was functional and efficient for the case under study, what makes this work result in valid contributions for the fields of intelligent systems, dynamic systems and in the methodological way to specify models to estimate time series.

Keywords: Intelligent systems. Electrical energy distribution. Fuzzy inference systems. Time series estimation. Dynamic systems modeling and simulation.

LISTA DE FIGURAS

Figura 2-1 – Exemplo de fuzzificação dos dados de temperatura (adaptado de [30]).	35
Figura 2-2 – Exemplo de agregação de duas regras <i>fuzzy</i> . Em cinza está a região <i>fuzzy</i> de saída obtida (adaptado de [30]).	37
Figura 3-1 – Representação das entradas do SIF multicamadas [8].	42
Figura 3-2 – SIF multicamadas [8].	43
Figura 3-3 – Algoritmo de extração das regras <i>fuzzy</i> [8].	44
Figura 3-4 – Algoritmo de ajuste paramétrico [8].	47
Figura 4-1 – Representação dos vetores da equação (4-1).	50
Figura 4-2 – Algoritmo de seleção de entradas.	51
Figura 4-3 – Série temporal normalizada de demanda de potência ativa.	52
Figura 4-4 – Demanda típica ao longo de um dia.	52
Figura 4-5 – Dados de treinamento do SIF.	53
Figura 4-6 – Autocorrelação dos dados de treinamento.	54
Figura 4-7 – Correlação entre o vetor <i>dnt</i> e os dados de treinamento.	54
Figura 4-8 – Média dos MAPEs dos treinamentos seguindo a metodologia desenvolvida (última coluna da Tabela 4-1).	56
Figura 4-9 – Média dos MAPEs dos testes seguindo a metodologia desenvolvida (última coluna da Tabela 4-2).	58
Figura 4-10 – Número de vezes que cada atraso, com exceção aos referentes à primeira e à segunda entradas, foi selecionado.	59
Figura 4-11 – Funções de pertinência otimizadas para a primeira entrada do melhor SIF treinado seguindo a metodologia desenvolvida.	60
Figura 4-12 – Funções de pertinência otimizadas para a segunda entrada do melhor SIF treinado seguindo a metodologia desenvolvida.	60
Figura 4-13 – Funções de pertinência otimizadas para a saída do melhor SIF treinado seguindo a metodologia desenvolvida.	61
Figura 4-14 – Base de regras otimizada para o melhor SIF treinado seguindo a metodologia desenvolvida.	61
Figura 4-15 – Histograma do erro da estimação realizada seguindo a metodologia desenvolvida.	61

Figura 4-16 – Propagação do erro da estimação realizada seguindo a metodologia desenvolvida.	62
Figura 4-17 – Média dos MAPEs dos treinamentos que não seguem a metodologia desenvolvida.	63
Figura 4-18 – Média dos MAPEs dos testes que não seguem a metodologia desenvolvida. ..	63
Figura 4-19 – Funções de pertinência otimizadas para a primeira entrada do melhor SIF treinado que não segue a metodologia desenvolvida.	64
Figura 4-20 – Funções de pertinência otimizadas para a segunda entrada do melhor SIF treinado que não segue a metodologia desenvolvida.	64
Figura 4-21 – Funções de pertinência otimizadas para a terceira entrada do melhor SIF treinado que não segue a metodologia desenvolvida.	65
Figura 4-22 – Funções de pertinência otimizadas para a quarta entrada do melhor SIF treinado que não segue a metodologia desenvolvida.	65
Figura 4-23 – Funções de pertinência otimizadas para a saída do melhor SIF treinado que não segue a metodologia desenvolvida.	65
Figura 4-24 – Base de regras otimizada para o melhor SIF treinado que não segue a metodologia desenvolvida.	66
Figura 4-25 – Histograma do erro da estimação realizada que não segue a metodologia desenvolvida.	66
Figura 4-26 – Propagação do erro da estimação realizada que não segue a metodologia desenvolvida.	67

LISTA DE TABELAS

Tabela 2-1 – Conjunto de regras <i>fuzzy</i> (adaptado de [30]).	36
Tabela 4-1 – MAPEs dos treinamentos seguindo a metodologia desenvolvida.	55
Tabela 4-2 – MAPEs dos testes seguindo a metodologia desenvolvida.	56
Tabela 4-3 – Atrasos das entradas selecionadas seguindo a metodologia desenvolvida.	58

SUMÁRIO

1	INTRODUÇÃO À DISSERTAÇÃO DE MESTRADO	21
1.1	Introdução ao tema	21
1.2	Justificativas, motivações e aplicabilidades	21
1.3	Objetivos	22
1.4	Contribuições da dissertação de mestrado	23
1.5	Organização da dissertação.....	25
2	TÉCNICAS PARA ESTIMAÇÃO DE SÉRIES TEMPORAIS E SUAS APLICAÇÕES NO SETOR ELÉTRICO	27
2.1	Introdução.....	27
2.2	Métricas de cálculo de erro de estimação	27
2.3	Métodos lineares de estimação.....	29
2.3.1	Autorregressivo	29
2.3.2	Média móvel	30
2.3.3	Outros métodos lineares.....	31
2.4	Métodos não-lineares de estimação	32
2.4.1	Redes neurais artificiais	33
2.4.2	Sistemas de inferência <i>fuzzy</i>	34
2.4.3	Outros métodos não-lineares	38
2.5	Considerações sobre os métodos de estimação de séries temporais ...	39
3	O SISTEMA DE INFERÊNCIA <i>FUZZY</i> MULTICAMADAS	41
3.1	Introdução.....	41
3.2	Camada de entrada	41
3.3	Camada de inferência	42
3.4	Camada de saída	43
3.5	Ajuste estrutural.....	44
3.6	Ajuste paramétrico	45

4	ESTIMAÇÃO DE DEMANDA DE POTÊNCIA ATIVA UTILIZANDO SISTEMAS <i>FUZZY</i>	49
4.1	Introdução	49
4.2	Metodologia para seleção das entradas do sistema <i>fuzzy</i>	49
4.3	Cenário do caso de estudo	51
4.4	Resultados da aplicação da metodologia desenvolvida no cenário estudado.....	53
4.5	Comparação dos resultados	62
5	CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS.....	69
	REFERÊNCIAS	71

1 INTRODUÇÃO À DISSERTAÇÃO DE MESTRADO

1.1 Introdução ao tema

Nas últimas décadas, o setor de energia elétrica mundial experimentou inúmeros avanços decorrentes do advento de novas tecnologias e de sua incorporação aos sistemas elétricos de potência, de novas exigências do mercado consumidor principalmente no que tange à qualidade e à confiabilidade do fornecimento de energia elétrica, e do sensível aumento no emprego de formas alternativas de energia que colaboraram para o crescimento da geração distribuída. Complementando, observa-se ainda a ampliação dos sistemas elétricos interconectados, os quais, em função de sua complexidade, de sua área de abrangência e do montante de carga instalada, demandam estudos detalhados para elaboração de seu sistema de controle, de comunicação e de proteção a fim de alinhar sua operação a estes novos paradigmas impostos aos sistemas elétricos de potência [1] - [3].

Neste contexto, o sistema elétrico brasileiro é permeado por grandes desafios em sua manutenção e expansão, e a estimação de séries temporais, mais especificamente a estimação da demanda de potência, é um dos estratagemas utilizados pelos agentes do setor elétrico como suporte às tomadas de decisões que acabam por ultrapassar estes desafios. Estimar esta demanda para o curto prazo, cujo horizonte é menor que algumas semanas, ajuda na operação do sistema. Estimações de demanda de médio prazo, com horizonte entre alguns meses e cinco anos, têm função de auxílio na compra e venda de energia elétrica em leilões realizados pela Câmara de Comercialização de Energia Elétrica. Já a finalidade de estimar esta grandeza para o longo prazo, com horizonte maior que cinco anos, é planejar e projetar novas usinas e linhas de transmissão [4], [5].

Conhecendo a importância da estimação de demanda de potência para o setor elétrico brasileiro, pode-se discorrer sobre as justificativas, motivações e aplicabilidades deste trabalho, assim como seus objetivos.

1.2 Justificativas, motivações e aplicabilidades

O processo de tomada de decisão pode ser modelado na forma de um problema de otimização em que o risco, ou o custo operacional, vem a ser a variável a ser minimizada. A estimação da demanda da potência ativa tem essa finalidade quando no contexto dos sistemas

elétricos. Para tanto, é preciso estimar o comportamento futuro da demanda de potência com erro reduzido, para que as decisões tomadas tenham um risco pequeno associado. Assim, desenvolver uma ferramenta de estimação capaz de retornar um pequeno erro auxiliará os agentes do setor elétrico na tomada de decisões.

A justificativa para almejar um pequeno erro de estimação é reduzir os custos operacionais pagos pelos agentes, custos estes que são tanto maiores quanto mais incertas forem as estimações. Uma estimação de curto prazo com erro elevado pode fazer com que uma subestação não seja adequadamente operada, causando uma possível interrupção no fornecimento de energia. Para o médio prazo, grandes erros de estimação podem desencadear compra e venda de energia de forma inapropriada, resultando em multas devido à formalização de contratos inadequados. Já no longo prazo, anos de estudo, projeto e investimento podem dar muito prejuízo caso a estimação inicial esteja muito divergente da situação real [6], [7].

Agentes de áreas como a operação, o planejamento do sistema e o mercado de energia elétrica, frequentemente necessitam de estimações precisas de demanda de potência para tomar decisões cujos efeitos serão visualizados em curto, médio ou longo prazo. Percebem-se assim, as contribuições deste trabalho no que diz respeito às suas diversas áreas de aplicabilidade.

1.3 Objetivos

Embora as aplicabilidades deste trabalho possam abranger diversos horizontes de estimação e beneficiar diversos agentes do setor elétrico, o trabalho está focado na estimação de curto prazo da demanda de potência ativa de um alimentador de um sistema de distribuição de energia elétrica. Portanto, o principal objetivo deste trabalho é desenvolver uma abordagem *fuzzy* capaz de realizar esta estimação com erro reduzido. A validação desta abordagem, cujo principal destaque é a metodologia para selecionar as entradas do sistema de previsão, acontece empregando-se o Sistema de Inferência *Fuzzy* (SIF) multicamadas desenvolvido em [8].

Além disso, é objetivo deste trabalho gerar contribuições para as áreas de sistemas inteligentes, de sistemas dinâmicos e inclusive na forma metodológica de especificação de modelos de estimação de séries temporais, já que a principal contribuição do trabalho é caracterizada pela metodologia de seleção das entradas do SIF de previsão, que, mesmo sendo apenas uma etapa de todo o processo de estimação, pode derivar futuras investigações relacionadas à previsão de séries temporais utilizando sistemas inteligentes.

1.4 Contribuições da dissertação de mestrado

O desenvolvimento da abordagem *fuzzy* propiciou a realização de diversos trabalhos científicos, que foram realizados em conjunto com outros pesquisadores do Laboratório de Automação Inteligente de Processos e Sistemas. A seguir estão todos os trabalhos que são frutos do tema desta dissertação de mestrado. Todos eles foram publicados em anais de congressos.

- MORAES, L. A.; FLAUZINO, R. A.; ARAUJO, M. A.; BATISTA, O. E.: Development of a Methodology to Forecast Time Series using Few Input Variables. In: 2013 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America), 2013, Sao Paulo. 2013 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America).
- MORAES, L. A.; FLAUZINO, R. A.; ARAUJO, M. A.; BATISTA, O. E.: Development of an Approach for Choosing the Input Variables for a Forecast Fuzzy Inference System. In: World Congress in Computer Science, Computer Engineering, and Applied Computing, 2013, Las Vegas. Expert Systems for Intelligent Automation Purposes in Power Systems, 2013.
- MORAES, L. A.; FLAUZINO, R. A.; ARAUJO, M. A.; BATISTA, O. E.: A Fuzzy Methodology to Improve Time Series Forecast of Power Demand in Distribution Systems. In: 2013 IEEE Power & Energy Society General Meeting, 2013, Vancouver. 2013 IEEE Power & Energy Society General Meeting.
- MORAES, L. A.; FLAUZINO, R. A.; ARAUJO, M. A.; BATISTA, O. E.: Sistema Fuzzy com Entradas não Triviais aplicado à Estimaco de Demanda de Potncia. In: X Latin-American Congress on Electricity Power Generation, Transmission and Distribution (CLAGTEE), 2013, Via del Mar. Proceedings of CLAGTEE, 2013.

Tambm foram realizados trabalhos cujo tema principal no  o mesmo desta dissertaco, mas que colaboraram para o aprendizado de algumas das ferramentas utilizadas nesta. So eles:

- MORAES, L. A.; FLAUZINO, R. A.; BATISTA, O. E.; SILVA, I. N.: Nova Abordagem na Determinaco da Gravidade de Infraoes de Trnsito Atravs da

Lógica Fuzzy. In: Congresso Brasileiro de Automática, 2012, Campina Grande. Anais do XIX Congresso Brasileiro de Automática, CBA 2012, 2012. p. 965.

- BATISTA, O. E.; FLAUZINO, R. A.; MORAES, L. A.; ARAUJO, M. A.: The Faults Variability in Distribution Systems with Distributed Generation and Robustness of Smart Grids. In: 2013 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America), 2013, Sao Paulo. 2013 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America).
- ARAUJO, M. A.; FLAUZINO, R. A.; BATISTA, O. E.; MORAES, L. A.: Protection of the Distribution Lines against Lightning Overvoltages by Surge Arrester in the Context of Distributed Generation and Smart Grids. In: 2013 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America), 2013, Sao Paulo. 2013 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America).
- BATISTA, O. E.; FLAUZINO, R. A.; MORAES, L. A.; ARAUJO, M. A.: Applying a Fuzzy System for Robust Fault Location in a Distribution System with Distributed Generation. In: World Congress in Computer Science, Computer Engineering, and Applied Computing, 2013, Las Vegas. Expert Systems for Intelligent Automation Purposes in Power Systems, 2013.
- ARAUJO, M. A.; FLAUZINO, R. A.; BATISTA, O. E.; MORAES, L. A.: Application of Fuzzy System for Analysis of Lightning Overvoltages and Protection of Distribution Systems by Surge Arresters. In: World Congress in Computer Science, Computer Engineering, and Applied Computing, 2013, Las Vegas. Expert Systems for Intelligent Automation Purposes in Power Systems, 2013.
- BATISTA, O. E.; FLAUZINO, R. A.; MORAES, L. A.; ARAUJO, M. A.: Redes Neurais Artificiais aplicadas à Localização Robusta de Faltas em Sistemas de Distribuição com Geração Distribuída. In: X Latin-American Congress on Electricity Power Generation, Transmission and Distribution (CLAGTEE), 2013, Viña del Mar. Proceedings of CLAGTEE, 2013.
- ARAUJO, M. A.; FLAUZINO, R. A.; BATISTA, O. E.; MORAES, L. A.: Application of Surge Arrester for Overvoltage Protection Caused by Direct Lightning on Distribution Lines. In: X Latin-American Congress on Electricity Power Generation, Transmission and Distribution (CLAGTEE), 2013, Viña del Mar. Proceedings of CLAGTEE, 2013.

- ARAUJO, M. A.; FLAUZINO, R. A.; BATISTA, O. E.; MORAES, L. A.: Applying a Fuzzy System for Lightning Location and Protection of Distribution Systems against Overvoltage by Surge Arresters. In: Simpósio Brasileiro de Automação Inteligente (SBAI), 2013, Fortaleza. Anais do SBAI, 2013.
- BATISTA, O. E.; FLAUZINO, R. A.; MORAES, L. A.; ARAUJO, M. A.: Abordagem Possibilística Fuzzy para Localização Robusta de Falhas em Alimentadores com Geração Distribuída. In: Simpósio Brasileiro de Automação Inteligente (SBAI), 2013, Fortaleza. Anais do SBAI, 2013.
- ARAUJO, M. A.; FLAUZINO, R. A.; BATISTA, O. E.; MORAES, L. A.; MARTINS, C. H. R.: Protection of the distribution lines with distributed generation against lightning overvoltages by surge arrester in the context of smart grids. In: 2013 World Congress on Sustainable Technologies (WCST), 2013, London. 2013 World Congress on Sustainable Technologies (WCST). p. 36.
- BATISTA, O. E.; FLAUZINO, R. A.; MORAES, L. A.; SOARES, M. M.: Gestão Energética Empresarial e suas Potencialidades no Setor Elétrico Brasileiro. In: Congresso Brasileiro de Automática, 2012, Campina Grande. Anais do XIX Congresso Brasileiro de Automática, CBA 2012, 2012. p. 1297.

1.5 Organização da dissertação

O texto da dissertação de mestrado está estruturado em cinco Capítulos. Após esse Capítulo introdutório será desenvolvido o Capítulo 2, que apresenta algumas das ferramentas utilizadas para estimação de séries temporais e suas aplicações voltadas ao setor elétrico. Ainda no Capítulo 2 serão discutidas as principais formas de se avaliar a qualidade de uma estimação.

No Capítulo 3 será descrito o SIF multicamadas que será empregado para realizar as estimações de demanda de potência ativa em um sistema de distribuição de energia elétrica, abordando desde a função de cada camada até os ajustes paramétricos e estruturais do SIF em questão.

O Capítulo 4 ilustra a metodologia proposta para identificar as entradas do sistema inteligente de estimação utilizado neste trabalho bem como o cenário do caso onde a metodologia foi aplicada. Também expõe e discute sobre os resultados obtidos.

Por fim, no Capítulo 5, estão as conclusões da pesquisa e as sugestões para trabalhos futuros no âmbito da estimação de séries temporais.

2 TÉCNICAS PARA ESTIMAÇÃO DE SÉRIES TEMPORAIS E SUAS APLICAÇÕES NO SETOR ELÉTRICO

2.1 Introdução

Existem diversas áreas de estudo onde a estimação de séries temporais é aplicada [9], [10]. Na meteorologia, por exemplo, tenta-se prever a temperatura em determinada região, a umidade do ar, o índice pluviométrico e a velocidade dos ventos. Já em economia, as variáveis de interesse são os índices de preço, a inflação, as taxas de juros, as cotações de moedas estrangeiras e das ações na bolsa de valores.

Em se tratando de sistemas dinâmicos, sempre que for possível coletar dados criando um histórico sobre o comportamento de determinada variável, originando assim uma série temporal, é possível a adoção de técnicas para estimar o comportamento futuro desta grandeza.

Estimar séries temporais é um problema importante no âmbito do setor elétrico, tendo em vista a operação, a formalização de contratos de compra e venda de energia, o planejamento e a expansão do sistema. Em função deste aspecto, muitos são os trabalhos acadêmicos que tratam sobre este assunto.

Assim sendo, neste Capítulo, serão abordadas primeiramente, na Seção 2.2, as principais métricas de cálculo de erro de estimação, que tem por função analisar a precisão de uma estimação. Depois, nas Seções 2.3 e 2.4, será explicado o funcionamento de métodos e ferramentas de estimação de séries temporais lineares e não-lineares respectivamente. Juntamente com cada técnica serão citados exemplos de aplicação relacionados ao setor elétrico, o que torna possível comparar o nível de erro que cada método de estimação pode atingir. Por fim, na Seção 2.5 há uma breve consideração sobre os métodos explicados e os exemplos de aplicações fornecidos.

2.2 Métricas de cálculo de erro de estimação

As estimações de séries temporais são feitas com base em um horizonte de previsão que, muitas vezes, contempla mais que um único instante da série. Para analisar a acurácia de uma estimação para o horizonte como um todo, é comum calcular a média dos erros de cada instante, de forma que um valor numérico indique a precisão da estimação realizada [11].

Como o erro de estimação pode ser positivo ou negativo para cada instante do horizonte, para evitar que no cálculo da média os erros se cancelem, é interessante empregar métricas que trabalham apenas com valores não-negativos para verificar a qualidade de uma estimação [12].

Os erros médios mais empregados na estimação de séries temporais são o Erro Médio Absoluto (MAE, do inglês *mean absolute error*), o Erro Médio Absoluto Percentual (MAPE, do inglês *mean absolute percentage error*), o Erro Quadrático Médio (MSE, do inglês *mean square error*) e Raiz quadrada do Erro Quadrático Médio (RMSE, do inglês *root mean square error*), apresentados respectivamente pelas equações (2-1), (2-2), (2-3) e (2-4) [11] - [14].

$$MAE = \frac{\sum_{t=1}^N |X_t - Y_t|}{N} \quad (2-1)$$

$$MAPE = \frac{\sum_{t=1}^N \left| \frac{X_t - Y_t}{X_t} \right|}{N} \quad (2-2)$$

$$MSE = \frac{\sum_{t=1}^N (X_t - Y_t)^2}{N} \quad (2-3)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (X_t - Y_t)^2}{N}} \quad (2-4)$$

sendo:

- t o instante estimado;
- X_t o valor conhecido da série temporal no instante t ;
- Y_t o valor estimado da série temporal no instante t ;
- N o número total de instantes estimados.

O MAE e o RMSE são medidos na mesma unidade da série temporal e por isto são bastante utilizados em relatórios e artigos para comunicar o resultado aos interessados, representando qual seria a magnitude de um erro típico. O MAE é, normalmente, um pouco menor que o RMSE, e também é matematicamente de mais fácil compreensão [13].

Por vezes, pratica-se para cada instante da série temporal a divisão do RMSE pelo próprio valor conhecido da série, resultando em um RMSE percentual, que passa a não ser dependente da escala da série temporal, assim como o MAPE.

Tanto o RMSE quanto o MSE, devido ao fato de elevarem a diferença entre valor estimado e conhecido ao quadrado, são bastante sensíveis a pontos extremos [14].

Já o MAPE é a métrica mais utilizada para mensurar a qualidade de uma estimação, pois ela retorna um erro adimensional, que pode ser compreendido e analisado mesmo desconhecendo-se a unidade da série temporal, não é dependente de escala e não é sensível a pontos extremos. Entretanto, estimar uma série temporal com valores próximos a zero pode resultar um MAPE elevado mesmo que a estimação tenha sido satisfatória [11].

Conhecendo-se as principais métricas de cálculo de erro de estimação é possível prosseguir para a explanação dos principais métodos de estimação de séries temporais, além de ser possível também analisar a qualidade e a precisão de exemplos de estimações de grandezas relacionadas ao setor elétrico.

2.3 Métodos lineares de estimação

Os métodos lineares de estimação de séries temporais calculam o valor do instante presente da série em questão associando linearmente, por meio de parâmetros, as informações contidas na própria série temporal, ou mesmo em outra série correlacionada com a que se deseja prever. Para calcular o valor do instante atual da série, basta calcular estes parâmetros.

Como assume-se uma dependência linear, a técnica dos mínimos quadrados é satisfatória no cálculo dos parâmetros, resultando no menor erro possível de estimação para os casos dos modelos lineares [15] - [17]. Esta técnica pode ser aplicada para encontrar os parâmetros de todos os métodos lineares explanados a seguir.

2.3.1 Autorregressivo

O primeiro dos métodos lineares de estimação de séries temporais abordado neste trabalho é o autorregressivo (AR), cujo modelo está representado por meio da equação (2-5). Este método define que o valor de cada instante de uma série temporal está distante de sua média por um ruído branco, de média zero e variância constante, e que o valor da série temporal

para o instante t depende linearmente dos valores dos N instantes imediatamente anteriores a ele [16].

$$X_t = C + \sum_{i=1}^N W_i^{AR} X_{t-i} + \varepsilon_t \quad (2-5)$$

sendo:

- t o instante a ser estimado;
- X_t o valor da série temporal no instante t ;
- W_i^{AR} o i -ésimo parâmetro do modelo;
- ε_t o ruído branco, ou erro, associado ao instante t ;
- N o número total de instantes considerados no modelo, e também sua ordem;
- C a média da série temporal.

Um exemplo de aplicação do modelo AR para estimação de séries temporais relacionadas ao setor elétrico é o estudado em [18]. Estimou-se a velocidade dos ventos para a geração eólica de energia em Hong Kong. Os MAPEs alcançados foram da ordem de 10,7%.

Outro exemplo aparece em [19]. Foi definida uma margem de propagação de erro aceitável para diversos horizontes de estimação de curtíssimo prazo relacionada à geração eólica na Turquia. O modelo AR empregado foi capaz de estabelecer um MAE de estimação 1,44% melhor que o erro considerado como referência.

2.3.2 Média móvel

O método da média móvel (MA, do inglês *moving average*) utiliza a equação (2-6) para obter os resultados de estimação. Este modelo também define que o valor de um instante da série temporal é a soma de um valor médio e um ruído branco [16].

$$X_t = C + \sum_{i=1}^M W_i^{MA} \varepsilon_{t-i} + \varepsilon_t \quad (2-6)$$

Tanto o modelo AR de ordem N quanto o modelo MA de ordem M levam em consideração os valores dos N ou M instantes anteriores ao que se deseja prever para estimá-

lo, mas o primeiro método propõe que o valor do instante que se deseja estimar depende linearmente dos últimos valores da série, enquanto que o segundo modelo diz que o valor desconhecido depende linearmente do erro relacionado aos últimos instantes da série [17].

O modelo MA, juntamente com o modelo AR, é importante na composição de outros modelos lineares, como explicado a seguir.

2.3.3 Outros métodos lineares

Relacionar linearmente o valor de um instante de uma série temporal apenas com seus anteriores, ou apenas com os erros a ele relacionados, constitui uma prática que acaba por ocasionar erros de estimação bastante elevados em determinados problemas relacionados ao setor elétrico, já que os modelos AR e MA isolados podem não englobar todas as particularidades de comportamento de determinada série temporal. Para tentar reduzir este erro, é comum combinar os métodos lineares, e inclusive aprimorá-los, para criar outros métodos lineares de estimação. Assim, por meio destes métodos compostos são obtidos resultados melhores que os obtidos pelas ferramentas individualmente [17], [20].

Uma ferramenta linear híbrida, aprimorada e bastante completa a ser mencionada é o Método Linear Autorregressivo e de Média Móvel com Entradas Exógenas (ARMAX), que incorpora os modelos AR, MA e uma série temporal (Z_t) que é correlacionada à série temporal que se deseja estimar. O modelo ARMAX considera que existe dependência linear entre o valor do instante atual da série original e os valores dos P últimos instantes desta série (Z_t), além das dependências lineares com os valores dos N últimos instantes e os M últimos ruídos brancos da própria série. A definição do modelo linear generalizado ARMAX de ordem (N, M, P) está na equação (2-7) [16].

$$X_t = C + \sum_{i=1}^N W_i^{AR} X_{t-i} + \sum_{j=1}^M W_j^{MA} \varepsilon_{t-j} + \sum_{k=1}^P W_k^{EX} Z_{t-k} + \varepsilon_t \quad (2-7)$$

Existem vários outros modelos lineares que podem ser originados a partir do ARMAX. Por exemplo, considerando-se todos os parâmetros $W_k^{EX} = 0$, obtém-se um modelo ARMA de ordem (N, M) . Ou ainda, considerando todos os parâmetros W_j^{MA} nulos, tem-se o modelo ARX de ordem (N, P) .

Em [19], foi empregado um modelo ARX, cuja série principal era a geração de energia eólica, e a série exógena era a previsão da série principal realizada inicialmente com o já citado modelo AR. Este procedimento fez com que o a propagação do MAE de estimação do modelo ARX fosse 19,7% melhor que o erro considerado aceitável.

Na referência [21] utilizou-se o modelo ARMAX para prever a demanda do sistema elétrico de Taiwan para o curto prazo usando como série temporal principal os dados do sistema e como série exógena os dados climáticos de temperatura ambiente. O MAPE atingido foi de 2,22%, um valor muito bom para o erro.

Os autores de [22] empregaram um modelo ARMA para estimar geração eólica de curtíssimo prazo. Os MAPEs encontrados chegam a 18,5%.

Já [23] empregou um modelo ARX para estimar o preço diário da energia no mercado inglês. Esta referência utilizou como série temporal principal o preço médio diário da energia elétrica, e como série exógena seu preço instantâneo, mensurado a cada meia hora. Os autores obtiveram RMSEs de aproximadamente 6 £/MWh, sendo que o preço médio da energia na Inglaterra no período estudado, entre 2005 e 2012, foi de cerca de 50 £/MWh.

2.4 Métodos não-lineares de estimação

Como verificado em alguns dos exemplos citados na Seção 2.3, métodos lineares de estimação às vezes retornam um valor de erro de estimação alto demais para ser considerado satisfatório quando no contexto dos sistemas elétricos. Isto acontece porque nestas vezes, considerar uma dependência linear pode ser simplório demais.

Para conseguir reduzir o erro de estimação de uma série temporal neste contexto, pode-se utilizar métodos não-lineares de estimação. A equação do modelo não-linear generalizado, o NARMAX, está expressa em (2-8) [24].

$$X_t = F[X_{t-1}, X_{t-2}, \dots, X_{t-N}, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-M}, Z_{t-1}, Z_{t-2}, \dots, Z_{t-P}] + \varepsilon_t \quad (2-8)$$

sendo F alguma função não-linear que associa os valores dos N últimos instantes da série temporal, os M últimos ruídos e os valores dos P últimos instantes da série exógena. As duas formas mais comumente usadas para representar F são a polinomial e a racional, como a seguir nas equações (2-9) e (2-10) respectivamente [15].

$$X_t = \sum_h c_h \prod_{i=1}^N X_{t-i} \prod_{j=0}^M \varepsilon_{t-j} \prod_{k=1}^P Z_{t-k} \quad (2-9)$$

$$X_t = \frac{\sum_{h1} c_{h1} \prod_{i=1}^{N1} X_{t-i} \prod_{j=1}^{M1} \varepsilon_{t-j} \prod_{k=1}^{P1} Z_{t-k}}{\sum_{h2} c_{h2} \prod_{i=1}^{N2} X_{t-i} \prod_{j=1}^{M2} \varepsilon_{t-j} \prod_{k=1}^{P2} Z_{t-k}} + \varepsilon_t \quad (2-10)$$

sendo:

- h o número total de coeficientes dos polinômios;
- c_h o h -ésimo coeficiente;
- 1 e 2 a discriminação entre coeficientes e variáveis presentes respectivamente no numerador e no denominador.

Considerando o modelo generalizado de [24], modelos inteligentes podem ser considerados modelos NARMAX, uma vez que associam as entradas fornecidas a eles de maneira não-linear. Estes modelos inteligentes estão detalhados a seguir.

2.4.1 Redes neurais artificiais

As Redes Neurais Artificiais (RNAs) são uma ferramenta inteligente com vasto campo de aplicações, como a classificação de padrões, aproximação de funções, controle de processos, otimização de sistemas, recuperação de dados corrompidos e inclusive estimação de séries temporais. Elas são compostas de neurônios artificiais interconectados e organizados em camadas. Uma RNA é capaz de armazenar informações sobre determinado problema de forma distribuída em matrizes de pesos sinápticos [25].

Outras características que tornam as RNAs uma boa opção de ferramenta são a adaptação por experiência, a capacidade de aprendizado e a habilidade de generalização. Em outras palavras, para resolver um problema, uma RNA necessita basicamente de uma base conhecida de dados de entrada e saída do sistema, com a qual será realizado seu treinamento. Devido a estas características, cada vez mais as RNAs vêm sendo empregadas na resolução de problemas [26].

Em [2], foi utilizada uma RNA com 25 neurônios em sua camada escondida para prever a demanda de carga, a velocidade dos ventos e a geração eólica no México em curto prazo. Foram obtidos MAPEs de 2,5%, 1,0% e 3,5% respectivamente.

Já em [3], foram usados diferentes algoritmos de treinamento em uma RNA com uma camada escondida com um número variável de neurônios para estimar a carga de uma *smart grid* no curto prazo no Canadá. Foram usados como entradas dados da carga e também diversos dados meteorológicos. Todos os resultados obtidos apresentaram MSE entre 0,051 e 0,063, valores baixos dada a escala dos dados utilizados.

Em [5] foi utilizada uma RNA para estimar a demanda de carga mensal para o médio prazo da cidade de Jeddah, na Arábia Saudita. O MAPE atingido pela RNA foi de 5,46% e o RMSE percentual foi de 1,96%.

A referência [21] comparou seus resultados obtidos com o modelo ARMAX com resultados obtidos por uma RNA. O primeiro resultado foi ligeiramente melhor, já que o MAPE obtido por meio da estimação com uma RNA com 10 neurônios na camada escondida foi de 2,31%.

Já [27] utiliza uma RNA com duas camadas escondidas com 15 neurônios cada para prever a carga de curto prazo de quatro alimentadores de uma subestação. Em alguns dos casos estudados, além dos dados históricos da carga, foram utilizados também dados sobre a temperatura ambiente. Esse trabalho conseguiu atingir MAPEs entre 1,4% e 5,7%.

Em [28] foi estimada a carga para o dia seguinte no sistema elétrico da região da Nova Inglaterra para o ano de 2012, utilizando como entradas os dados de carga de 2007 a 2011, além de informações climáticas e sobre a hora do dia. A RNA, com 42 neurônios em sua camada escondida, foi capaz de retornar MAPEs que não ultrapassaram os 5,87%.

Um último exemplo de aplicação de RNA no setor elétrico é o desenvolvido em [29], que estimou o consumo de energia anual na China. Para isto, foram usados como entrada dados normalizados de consumo anuais entre 1978 e 2007, além do número de habitantes e do índice de urbanização para cada ano. Diversas topologias foram testadas para esta estimação de longo prazo, e a melhor delas, com 8 neurônios na camada escondida, resultou MAPEs que variaram de 0,3% a 8%.

2.4.2 Sistemas de inferência *fuzzy*

Os SIFs, assim como as RNAs, também são uma ferramenta inteligente. Apesar de ambos terem várias aplicações semelhantes, trabalham de formas diferentes. A primeira diferença é que ao invés de armazenarem as informações sobre determinado problema em uma matriz, os SIFs as armazenam em um conjunto de regras linguísticas. Por este motivo, geralmente, a interpretação da solução dos problemas feita por um ser humano acontece de

forma mais natural quando o problema é resolvido por meio de um SIF do que quando o mesmo problema é resolvido por uma RNA [25].

Outra diferença entre os SIFs e as RNAs é a forma como eles tratam os dados. SIFs são capazes de manipular dados carregados com grande grau de incerteza, ou seja, podem analisar dados qualitativamente por meio de funções de pertinência. Já as RNAs utilizam em seu treinamento valores exatos, com os quais cálculos são feitos para obter as matrizes de pesos sinápticos. Elas analisam os dados quantitativamente [25].

Para o correto funcionamento de um SIF, ou seja, para que ele seja capaz de capturar um conjunto de dados de um meio externo e obter uma saída coerente, existe uma série de processos que precisam ser realizados ordenadamente. O primeiro deles é chamado de fuzzificação, ou seja, é a transformação dos dados de entrada em conjuntos *fuzzy*. Fuzzificar determinado dado significa transformar seu universo de discurso, que contém os valores numéricos destes dados, em funções de pertinência, que relacionam cada valor a um conceito linguístico. É a fuzzificação que permite que um SIF receba as informações com as quais ele trabalhará, pois é neste processo que as características quantitativas deixam de ser importantes, enquanto que as qualitativas ganham importância [30].

É também neste processo *fuzzy* que as incertezas passam a ser, de certa forma, ignoradas, pois uma faixa relativamente grande de valores pode ser toda tratada da mesma maneira segundo um conceito linguístico [31].

Para melhor explicar um SIF, que é a ferramenta de estimação empregada neste trabalho, será usado um exemplo de [30]. Na Figura 2-1 está a fuzzificação dos dados de temperatura em cinco conceitos linguísticos: muito baixa, baixa, confortável, alta e muito alta. Cada um destes conceitos representa uma função de pertinência. A quantidade destas funções, seu formato e a amplitude do universo de discurso para cada conjunto de dados são parâmetros que precisam ser definidos em um SIF.

Figura 2-1 – Exemplo de fuzzificação dos dados de temperatura (adaptado de [30]).

De posse das entradas fuzzificadas pode-se realizar o segundo processo *fuzzy*, que é denominado inferência. A inferência é responsável pela associação das funções de pertinência das entradas a regras linguísticas do tipo "se – então", cujos consequentes são as saídas do sistema. Sejam, por exemplo, duas entradas a temperatura e a variação de temperatura em uma sala, e uma saída a potência de um aquecedor presente nesta sala. Uma típica regra *fuzzy* seria: Se (temperatura é baixa) e (variação de temperatura é pouco negativa) – então (potência do aquecedor é média) [30].

De fato, expressões linguísticas como a anterior, que envolvem conceitos qualitativos, geralmente são mais inteligíveis ao ser humano que equações com termos quantitativos. Porém, para que um SIF possa atuar tão bem quanto uma ferramenta quantitativa, é preciso criar uma base de regras *fuzzy* que contenha as mais importantes ou senão todas as combinações entre as funções de pertinência das entradas. Portanto, a quantidade de regras *fuzzy* que forma a base de regras também é um parâmetro a ser ajustado em um SIF. A Tabela 2-1 exibe um conjunto de regras *fuzzy* criado para controlar o aquecedor do exemplo dado anteriormente [30].

Tabela 2-1 – Conjunto de regras *fuzzy* (adaptado de [30]).

Variação Temp. Temp.	Muito negativa	Pouco negativa	Zero	Pouco positiva	Muito positiva
Muito baixa		Potência alta	Potência alta	Potência média	Potência média
Baixa	Potência alta	Potência média	Potência média	Potência baixa	Potência baixa
Confortável		Potência média	Potência baixa	Potência baixa	
Alta		Potência baixa	Potência baixa	Potência baixa	
Muito alta	Potência média	Potência baixa	Potência baixa		

Uma vez que as entradas foram fuzzificadas e associadas à saída por meio de regras de inferência *fuzzy*, pode-se pensar que já é possível obter a saída do SIF. Porém, antes de obtê-la, é preciso realizar o processo de agregação das regras *fuzzy*. Observando a Figura 2-1 pode-se perceber que alguns valores de temperatura encaixam-se em mais de uma função de pertinência. Isto significa que pode acontecer de mais de uma regra *fuzzy* ser ativada ao mesmo tempo para um mesmo conjunto de valores de dados de entrada. A temperatura aos 22 °C, por exemplo, pode ser "confortável" com pertinência 0,4 e "alta" com pertinência 0,3, o que ativaria duas

regras com consequências diferentes caso a variação de temperatura fosse "pouco negativa", com pertinência 1,0 [30].

Quando isto acontece nenhuma das regras deve ser descartada. Pelo contrário, todas as consequências devem ser agregadas para formar uma região *fuzzy* de saída do SIF. Existem vários métodos de agregação, e pode-se considerar sua escolha outro parâmetro do SIF a ser definido. A Figura 2-2 ilustra o resultado obtido pelo método máximo-mínimo, regido pela equação (2-11), aplicado ao caso das duas regras ativadas no exemplo anterior. A equação abrange todo o universo de discurso da saída [30].

$$\mu_S = \max \left\{ \min \left[\max(\mu_{E1}^{(R1)}, \mu_{E2}^{(R1)}), \mu_S^{(R1)} \right], \min \left[\max(\mu_{E1}^{(R2)}, \mu_{E2}^{(R2)}), \mu_S^{(R2)} \right] \right\} \quad (2-11)$$

sendo:

- μ_S a pertinência da região fuzzy de saída;
- μ_{Ei}^{Rj} a pertinência da i-ésima entrada na j-ésima regra ativada;
- μ_S^{Rj} os valores da função de pertinência do consequente da j-ésima regra ativada.

Figura 2-2 – Exemplo de agregação de duas regras *fuzzy*. Em cinza está a região *fuzzy* de saída obtida (adaptado de [30]).

Depois da agregação, tem-se uma região *fuzzy* que representa a combinação de todas as regras *fuzzy* ativadas por determinado conjunto de entrada de dados, conjunto este que pertencia a um universo de discurso de valores e que foi fuzzificado para conceitos qualitativos. Porém, assim como as entradas vieram de um ambiente não-*fuzzy*, a saída será exportada a um ambiente não-*fuzzy*, e por isso ela deve ser um valor numérico, e não uma região *fuzzy*. Este valor numérico pode significar a potência do aquecedor a ser ajustada no exemplo de dado anteriormente, ou, no caso deste trabalho, significará a previsão da demanda da potência ativa

para determinado instante. Para obter este valor, executa-se o último processo de um SIF, a defuzzificação.

Dentre os métodos de defuzzificação, o mais comum é o centro de área. Calcula-se em que posição do eixo horizontal está o centro de área da região *fuzzy* obtida com a agregação e este valor será a saída numérica do SIF. Definir o método de defuzzificação também faz parte da parametrização do SIF.

Conhecendo-se a teoria que envolve os SIFs, alguns exemplos de sua aplicação na estimação de variáveis do setor elétrico podem ser explorados.

A referência [32] utilizou um SIF para estimar a demanda anual de energia elétrica para o médio prazo para diferentes classes de consumidores gregos. Foi usado um histórico de consumo de quinze anos para estimar os próximos três. Os resultados, segundo o próprio autor, são melhores que os que se obtém utilizando métodos lineares, porém são bastante semelhantes aos que são obtidos por meio de RNAs. O maior MAPE atingido foi de 2,16% para os consumidores comerciais de média tensão, e o menor foi de 0,55% para os consumidores industriais de média tensão.

Já em [33], foi usado um SIF para estimar a carga de curto prazo da cidade de Jiangning, na China. Para isso foram usados dados referentes às três últimas primaveras, e obteve-se como resposta o consumo no próximo dia primaveril. Este artifício é válido para diminuir a influência de variáveis climáticas na previsão. Os resultados obtidos foram satisfatórios, pois os erros absolutos percentuais alcançados foram inferiores a 2,07%, e a média deles, o MAPE, foi de 0,78%.

Finalmente, em [34], tenta-se estimar a demanda de curto prazo do *campus* da *Universiti Teknologi Petronas* na Malásia. Com os dados de um ano de consumo, foram estimadas as 24 semanas seguintes. O MAPE atingido foi de 5,15%, mas os erros absolutos de alguns períodos do horizonte de previsão foram bastante elevados.

2.4.3 Outros métodos não-lineares

Assim como para os métodos lineares de estimação, também é possível associar dois ou mais métodos não-lineares para se obter modelos ainda mais aprimorados de previsão de séries temporais.

Um modelo híbrido sobre o qual vale a pena comentar é aquele que combina as duas ferramentas não-lineares detalhadas neste capítulo. As chamadas ferramentas *neuro-fuzzy* aperfeiçoam a maneira qualitativa e incerta como as informações são armazenadas num SIF por

meio de um treinamento baseado no realizado com as RNAs. A mais conhecida ferramenta neuro-*fuzzy* é o Sistema de Inferência Neuro-*Fuzzy* Adaptativo (ANFIS) [25].

Esta ferramenta foi utilizada em [35] para estimar carga de curto prazo na África do Sul para diferentes tipos de dias (dias da semana, sexta-feira, fins de semana e feriados) no verão e no inverno. Como entradas do sistema foram usados dados da carga do próprio sistema, a umidade do ar e temperatura. A pior estimacão, ainda assim excelente, foi para o caso de feriados no inverno, cujo MAPE foi de 2,08% e o RMSE percentual foi de 2,59%.

O modelo ANFIS também foi utilizado em [36] para estimar o preço da energia elétrica no mercado da Califórnia para o curto prazo. Utilizando um histórico de preços e cargas de doze dias, estimou-se o preço da energia no dia seguinte, e obteve-se um MAPE de 5,38%. Segundo o autor, o MAPE foi aceitável, embora tenha havido casos isolados de períodos com erros de estimacão de até 14,54%.

Como último exemplo, pode-se citar [37], que aplicou o modelo ANFIS para prever a carga de médio prazo da cidade de Targoviste na Romênia. Este autor atingiu MAPE de 5,42% e RMSE percentual de 3,83% ao utilizar como entradas do sistema apenas dados de carga que se desejava prever.

2.5 Considerações sobre os métodos de estimacão de séries temporais

A pesquisa na literatura aponta que não há muitas estimacões de séries temporais do setor elétrico utilizando modelos lineares. Isto é um indício de que estas ferramentas muitas vezes não são capazes de atingir um valor de erro aceitável para os padrões do setor elétrico, que conforme [38], que cita o Ministério de Minas e Energia, é de 6% para o caso brasileiro.

Pode-se notar ainda que quando utiliza-se uma ferramenta não-linear, na maioria das vezes, não há critério de escolha das entradas do sistema, e assim, ocasionalmente informacões redundantes ou desnecessárias lhe são inseridas. O treinamento de um sistema que precisa tratar dados repetidos ou dados que não carregam informacões importantes certamente não será tão rápido quanto um sistema que trata apenas dados pertinentes.

A união destes dois fatores citados converge para a motivacão deste trabalho, que é a de criar um sistema eficiente de estimacão de series temporais voltado ao setor elétrico. Será empregado um modelo não-linear de estimacão, a ser detalhado no Capítulo 3, juntamente com um critério de seleccão de suas entradas, a ser explanado na Seccão 4.2.

3 O SISTEMA DE INFERÊNCIA *FUZZY* MULTICAMADAS

3.1 Introdução

Como o método de escolha de entradas desenvolvido por este trabalho será aplicado no SIF multicamadas desenvolvido em [8], vale a pena explicar primeiramente o funcionamento deste SIF para que, posteriormente, o assunto deste trabalho possa ser abordado adequadamente. Assim, neste capítulo, que será inteiramente baseado no trabalho de [8], encontra-se o detalhamento do funcionamento do SIF multicamadas.

Como visto anteriormente, um SIF realiza processos sequenciais distintos e claramente definidos. Basicamente, estes processos são a fuzzificação das entradas do sistema, a inferência das regras associadas ao mesmo, a agregação das regras e a posterior defuzzificação do resultado da agregação, que se torna a saída do SIF.

Por este motivo, a interpretação de um SIF por meio de um modelo multicamadas é compreensível. Cada camada pode realizar um ou mais processos diferentes, a fim de transformar um conjunto de entradas em uma saída desejada. Neste SIF, propõe-se utilizar três camadas: a de entrada, a de inferência e a de saída.

3.2 Camada de entrada

A camada de entrada conecta os dados de entrada, advindos do meio externo, ao SIF, e também os fuzzifica. Dessa forma, o grau de pertinência de cada entrada a um conceito linguístico, ou conjunto *fuzzy*, é definido pela camada de entrada.

Podem-se usar tantos conjuntos *fuzzy* por entrada quanto necessário para defini-la. Seja um SIF com uma entrada e com N funções de pertinência associadas a ela, a saída Y da camada de entrada do SIF será um vetor que contém as funções de pertinência $p_k(x)$ definidas para a entrada x , como na equação (3-1).

$$Y(x) = \begin{bmatrix} p_1(x) \\ p_2(x) \\ \vdots \\ p_N(x) \end{bmatrix} \quad (3-1)$$

Pode-se expandir este conceito de fuzzificação para o caso de um SIF com mais dados de entrada apenas concatenando os diversos vetores-coluna referentes a cada entrada. Para o caso de um SIF com duas entradas e com três funções de pertinência por entrada por exemplo, como na Figura 3-1, o vetor de saída da camada de entrada seria como o da equação (3-2).

Figura 3-1 – Representação das entradas do SIF multicamadas [8].

$$Y(x) = [p_1^{(1)}(x_1) \quad p_2^{(1)}(x_1) \quad p_3^{(1)}(x_1) \quad p_1^{(2)}(x_2) \quad p_2^{(2)}(x_2) \quad p_3^{(2)}(x_2)]^T \quad (3-2)$$

As funções de pertinência das entradas podem ser definidas por meio de qualquer função, não necessariamente as triangulares e trapezoidais como as da Figura 3-1. A única restrição é que elas estejam definidas no domínio fechado $[0,1]$. Porém, aconselha-se utilizar funções de pertinência simples, a fim de aumentar a velocidade de processamento e reduzir o uso de memória.

3.3 Camada de inferência

A camada de inferência das regras *fuzzy* recebe os dados de entrada já fuzzificados e os combina entre si de acordo com uma base de regras definida. Esta base de regras é onde fica armazenado o conhecimento sobre determinado problema. Por este motivo, é fundamental criar uma base de regras que seja coerente com o problema.

Além da criação da base de regras, na segunda camada do SIF multicamadas acontece a ponderação destas regras. Assim podem-se valorizar aquelas regras que mais influenciam na solução do problema. Esta ponderação acontece segundo a equação (3-3).

$$R_i(Y(x)) = w_i r_i(Y(x)) \quad (3-3)$$

sendo:

- $R_i(\cdot)$ a função que representa o valor ponderado da i -ésima regra fuzzy;
- w_i o fator de ponderação da i -ésima regra fuzzy;
- $r_i(\cdot)$ o valor fuzzy da i -ésima regra fuzzy.

3.4 Camada de saída

Na camada de saída são realizados os processos de agregação das regras *fuzzy* e de defuzzificação, que são os dois últimos processos restantes para que se obtenha a saída do SIF. Além disso, é na terceira camada do SIF multicamadas que ficam armazenadas as funções de pertinência das saídas.

Escolher corretamente qual método de agregação e de defuzzificação utilizar é importante por dois motivos. Em primeiro lugar, a região *fuzzy* de resposta deve ser capaz de inferir adequadamente o conhecimento contido na base de regras, e em segundo lugar, o SIF deve ser capaz de traduzir esta região para um valor não-*fuzzy* que corresponda à realidade do problema. Além disso, os métodos não podem comprometer a velocidade de processamento.

Como a camada de saída trata regiões *fuzzy*, sua formulação é semelhante ao da camada de entrada, que também trata este tipo de região.

Para resumir e ilustrar vê-se a seguir a Figura 3-2, que exibe um esquemático do SIF multicamadas com duas entradas e três regras ativadas.

Figura 3-2 – SIF multicamadas [8].

3.5 Ajuste estrutural

Como dito anteriormente, o número de regras *fuzzy* e seu conteúdo linguístico, que formam a base de regras do SIF, são parâmetros a ser definidos. Encontrar a melhor dentre todas as bases não é uma tarefa fácil, já que existe um enorme número de possibilidades, dependendo da quantidade de entradas e de funções de pertinência de cada uma delas.

É conveniente, então, que o SIF multicamadas realize a busca pela melhor base de regras em determinado espaço válido. Este algoritmo de busca, que é baseado em [39], visa minimizar uma função custo explicitada em termos do MSE entre a saída obtida e a desejada.

Entretanto, a superfície formada pela função custo pode conter vários mínimos locais, o que, em muitos casos, pode atrapalhar a busca pelo mínimo global. Assim sendo, faz-se necessário que uma estratégia de busca local aconteça depois de efetuada a busca global.

O algoritmo da busca pela melhor base de regras está exposto na Figura 3-3. A busca global, que é a primeira etapa deste algoritmo, consiste em sortear várias bases de regras e selecionar aquela que apresenta o menor custo. Em seguida, realiza-se uma busca local em torno da vizinhança desta base escolhida, a fim de encontrar alguma outra base com custo ainda menor. Por fim, a base de regras do SIF será aquela que apresentar o menor custo depois de realizadas todas as iterações da primeira e da segunda etapas.

```

Inicializar o sistema fuzzy
R1 ← Selecionar uma base de regras aleatória
Contador ← 0
REPITA
  R2 ← Selecionar uma base de regras aleatória
  SE custo(R2) < custo(R1)
  ENTÃO
    R1 ← R2
    Contador ← Contador + 1
  FIM SE
ATÉ Contador muito grande

Contador ← 0
REPITA
  R2 ← Inserir uma pequena perturbação em R1
  SE custo(R2) < custo(R1)
  ENTÃO
    R1 ← R2
    Contador ← Contador + 1
  FIM SE
ATÉ Contador muito grande

```

Figura 3-3 – Algoritmo de extração das regras *fuzzy* [8].

3.6 Ajuste paramétrico

Dividir um SIF em camadas, cada qual realizando um processo diferente, faz com que haja em cada camada parâmetros livres referentes aos processos por elas desempenhados. Na camada de entrada há os parâmetros das funções de pertinência das entradas, enquanto que na camada de inferência as ponderações das regras *fuzzy* são os parâmetros, e na camada de saída existem os parâmetros dos conjuntos *fuzzy* do espaço de saída.

Assim como no caso do ajuste estrutural, o ajuste paramétrico também tem por objetivo minimizar uma função custo. Considerando um conjunto de ajuste de entradas e saídas $\{x, y\}$ fixo durante todo o ajuste paramétrico, pode-se escrever esta função custo ξ responsável pela sintonia dos parâmetros livres das três camadas $\{h^{(1)}, h^{(2)}, h^{(3)}\}$ no mapeamento *fuzzy* $g: x \rightarrow y$ em função deles, como na equação (3-4).

$$\xi_g = \xi(h^{(1)}, h^{(2)}, h^{(3)}) \quad (3-4)$$

Assim, depois de minimizado ξ_g , têm-se os parâmetros que, juntamente com a base de regras encontrada por meio do algoritmo da Figura 3-3, otimizam o SIF multicamadas. Esta minimização de ξ_g , que é o passo final para o ajuste completo do SIF, pode ser considerada como um caso de otimização irrestrita e, portanto, pode ser resolvida pelo método do gradiente descendente.

Para esclarecer o processo de ajuste paramétrico, considera-se como exemplo um SIF que contém uma entrada com duas funções de pertinência gaussianas, três regras *fuzzy* e uma saída também com duas funções de pertinência gaussianas. Assim, os parâmetros $\{h^{(1)}, h^{(2)}, h^{(3)}\}$, referentes respectivamente à camada de entrada, à camada de inferência e à camada de saída do SIF são dados por:

$$h^{(1)} = [\mu_1^{(e)} \quad \sigma_1^{(e)} \quad \mu_2^{(e)} \quad \sigma_2^{(e)}]^T \quad (3-5)$$

$$h^{(2)} = [w_1 \quad w_2 \quad w_3]^T \quad (3-6)$$

$$h^{(3)} = [\mu_1^{(s)} \quad \sigma_1^{(s)} \quad \mu_2^{(s)} \quad \sigma_2^{(s)}]^T \quad (3-7)$$

sendo:

- $\mu_i^{(e)}$ a média da i-ésima função de pertinência da entrada;
- $\sigma_i^{(e)}$ a variância da i-ésima função de pertinência da entrada;
- w_i o fator de ponderação da i-ésima regra fuzzy;
- $\mu_i^{(s)}$ a média da i-ésima função de pertinência da saída;
- $\sigma_i^{(s)}$ a variância da i-ésima função de pertinência da saída.

Como cada camada do SIF tem processos distintos, e como o ajuste dos parâmetros de um destes processos pouco influencia os outros, é possível realizar o ajuste paramétrico do SIF camada a camada. Dessa forma, para exemplificar, na equação (3-8) aplica-se o método do gradiente descendente na camada de inferência.

$$h^{(2)}(k) = h^{(2)}(k-1) + \eta \cdot \nabla \xi \left(h^{(2)}(k-1) \right) \quad (3-8)$$

sendo η a taxa de aprendizagem e:

$$\nabla \xi(h^{(\cdot)}) = \left[\frac{\partial \xi}{\partial h_1^{(\cdot)}} \quad \frac{\partial \xi}{\partial h_2^{(\cdot)}} \quad \dots \quad \frac{\partial \xi}{\partial h_n^{(\cdot)}} \right]^T \quad (3-9)$$

sendo:

- $h_i^{(\cdot)}$ o i-ésimo elemento do vetor $h^{(\cdot)}$;
- n o número de elementos do vetor $h^{(\cdot)}$.

Como a expressão analítica do gradiente é difícil de ser obtida, utiliza-se sua aproximação numérica por meio do método de equações às diferenças, que está formulado na equação (3-10) [40].

$$\frac{\partial \xi}{\partial h_i^{(\cdot)}} \cong \frac{\xi(h^{(\cdot)} + \rho \cdot e_i) - \xi(h^{(\cdot)})}{\rho} \quad (3-10)$$

sendo ρ um valor muito pequeno e $e_i = [0 \quad 0 \quad \dots \quad 1 \quad \dots \quad 0]^T$ com o elemento não nulo ocupando a posição i .

Realizando este processo de otimização em todas as três camadas do SIF pode-se minimizar a função custo ξ_g , que é um MSE análogo ao definido pela equação (2-3). A Figura 3-4 ilustra este processo de otimização aplicado à camada de inferência. O algoritmo exposto é similar para as demais camadas.

```

Inicializar os parâmetros livres do sistema fuzzy multicamadas
Codificar os parâmetros em  $h^{(1)}$ ,  $h^{(2)}$  e  $h^{(3)}$ 
Deixar  $h^{(1)}$  e  $h^{(3)}$  fixos
Época  $\leftarrow 0$ 
REPITA
  PARA todos os padrões de sintonia FAÇA
    Propagar padrão camada a camada e obter saída  $y_k$ 
    Guardar o erro quadrático do respectivo padrão
    Calcular  $\nabla \xi(h^{(2)})$ 
    Atualizar  $h^{(2)}$ 
  FIM PARA
  Obter o erro quadrático médio  $\xi_g$ 
  Época  $\leftarrow$  Época + 1
ATÉ ( $\xi_g < \varepsilon$ ) OU (Contador muito grande)

```

Figura 3-4 – Algoritmo de ajuste paramétrico [8].

Para finalizar os ajustes dos parâmetros livres das camadas do SIF, ainda faz-se necessário definir os parâmetros referentes aos operadores *fuzzy*, às entradas, à base de regras e à saída. Define-se, portanto, o operador "ou probabilístico" para realizar a agregação e o método do centro de área [41] para realizar a defuzzificação. Todas as funções de pertinência são do tipo gaussiana assimétrica, e o número de regras existente na base de regras será um parâmetro escolhido no decorrer das simulações, como explicado na Seção 4.4.

4 ESTIMAÇÃO DE DEMANDA DE POTÊNCIA ATIVA UTILIZANDO SISTEMAS *FUZZY*

4.1 Introdução

Sabe-se que os métodos inteligentes de estimação realizam cálculos e relações mais complexas com as informações contidas em uma série temporal que uma simples dependência linear, e assim conseguem atender às exigências do setor elétrico no que diz respeito às estimativas de séries temporais, tendo em vista as referências apresentadas no Capítulo 2. Porém, inspirando-se nos métodos lineares de estimação, é comum que sejam utilizadas como entradas dos métodos inteligentes os valores de alguns instantes da própria série imediatamente anteriores ao instante que se deseja estimar. Esta prática de escolha de entradas, embora produza resultados confiáveis, pode não ser a que produz os melhores resultados, com os mínimos erros.

A fim de extrair melhor as informações contidas na série temporal para estimar seus valores futuros, propõe-se neste trabalho um novo método de escolha de entradas para ferramentas inteligentes. Estas entradas serão fornecidas ao SIF multicamadas explanado no Capítulo 3 para estimar a demanda de potência ativa em um alimentador de uma subestação.

4.2 Metodologia para seleção das entradas do sistema *fuzzy*

A fim de encontrar um conjunto adequado de entradas para que um SIF estime determinada série temporal com pequeno erro, a metodologia desenvolvida neste trabalho propõe que sejam escolhidas as entradas do sistema uma após outra, seguindo critérios de seleção. Estes critérios visam selecionar entradas que não contenham informações redundantes entre si sobre o comportamento da série temporal, e que também não contenham informações descorrelacionadas com o instante que se deseja prever. Dessa forma, o treinamento do SIF acontece mais rapidamente.

Assim, a primeira entrada do SIF de estimação é selecionada por meio do cálculo da autocorrelação dos dados de treinamento da série temporal. Aquele instante que apresentar maior autocorrelação com o valor atual, que é aquele que se deseja prever, terá seu valor escolhido como a primeira entrada do SIF. No MATLAB, *software* utilizado para implementar as simulações deste trabalho, a função que realiza este cálculo denomina-se “autocorr()”.

Já a segunda variável a compor as entradas do SIF deve ser aquela que melhor se correlacionar com o vetor normal à projeção do vetor correspondente à primeira entrada no vetor de demanda a ser estimado. O cálculo do vetor normal, $d_n(t)$, à projeção do vetor correspondente à primeira entrada, $d(t - t_1)$, no vetor de demanda a ser estimado $d(t)$ pode ser feito segundo a equação (4-1).

$$d_n(t) = d(t - t_1) - \frac{\langle d(t) | d(t - t_1) \rangle}{\|d(t)\|^2} d(t) \quad (4-1)$$

sendo $\langle \cdot | \cdot \rangle$ o produto escalar entre os argumentos, $\|\cdot\|^2$ a norma euclidiana e t_1 o número de atrasos da primeira entrada selecionada. Este cálculo seleciona um segundo instante da série temporal que não sofre tanta influência da primeira entrada escolhida, de forma que haja mais diversidade de informação inserida no SIF, mas que ainda está bastante correlacionado com o instante presente, e isto melhora o treinamento do SIF e a estimação da série. Para ilustrar o que representa cada termo da equação (4-1), tem-se a Figura 4-1. A correlação do vetor $d_n(t)$ com os dados de treinamento da série temporal é feita por meio da função “crosscorr()” no MATLAB.

Figura 4-1 – Representação dos vetores da equação (4-1).

Conhecendo as duas primeiras entradas do SIF, pode-se prosseguir para o processo iterativo de escolha das próximas entradas. Todas as entradas a partir da terceira são obtidas por meio da correlação dos dados de treinamento da série temporal com o erro de estimação obtido com as entradas até então selecionadas. Este procedimento visa inserir ao SIF informações sobre quão errada está sua estimação com o conjunto disponível, para desta forma reduzir o erro de estimação. A Figura 4-2 detalha todo o processo de seleção de variáveis.

```

 $d(t) \leftarrow$  Dados de treinamento da série temporal
Calcular autocorrelação de  $d(t)$ 
 $t_1 \leftarrow$  o instante com a maior autocorrelação em  $d(t)$ 
 $Entrada_1 \leftarrow d(t - t_1)$ 
Calcular  $d_n(t)$ 
Correlacionar  $d_n(t)$  com  $d(t)$ 
 $t_2 \leftarrow$  o instante com a maior correlação entre  $d_n(t)$  e  $d(t)$ 
 $Entrada_2 \leftarrow d(t - t_2)$ 
Definir número de regras do SIF
 $i \leftarrow 3$ 
REPITA
    Treinar o SIF para realizar uma estimacão com as entradas
    já selecionadas
    Obter o erro de estimacão para cada amostra dos dados de
    teste da série temporal
    Correlacionar este erro com  $d(t)$ 
     $t_i \leftarrow$  o instante com a maior correlação entre o erro e  $d(t)$ 
     $Entrada_i \leftarrow d(t - t_i)$ 
     $i \leftarrow i + 1$ 
ATÉ  $i >$  valor máximo de entradas estipulado

```

Figura 4-2 – Algoritmo de seleção de entradas.

Vale dizer que um outro critério de parada do algoritmo seria se o erro de estimacão atingido em alguma das iterações fosse menor que um valor considerado aceitável. Porém, neste trabalho, este critério de parada não foi utilizado porque deseja-se encontrar o melhor SIF para estimar a série temporal sob estudo.

4.3 Cenário do caso de estudo

Foi utilizada para treinamento, teste e validação do SIF multicamadas de estimacão uma série temporal de demanda de potência ativa adquirida em um alimentador de uma subestação de um sistema de distribuiçāo de energia elétrica. O intervalo entre cada mediçāo foi de cinco minutos e os dados foram medidos durante aproximadamente 210 dias. O total de instantes da série temporal é 60.610. A Figura 4-3 exibe esta série com seus valores normalizados, sendo considerada a maior demanda como sendo 1 pu.

De forma a se permitir uma melhor visualizaçāo do comportamento diário da demanda, apresentam-se por meio da Figura 4-4 duas curvas distintas. A curva azul corresponde à demanda média para determinado horário ao longo de todos os 210 dias. Em verde está a curva

que indica o percentil de 90% da demanda, ou seja, em 90% dos dias a demanda no horário indicado no gráfico encontra-se abaixo da curva verde.

Figura 4-3 – Série temporal normalizada de demanda de potência ativa.

Figura 4-4 – Demanda típica ao longo de um dia.

4.4 Resultados da aplicação da metodologia desenvolvida no cenário estudado

Para aplicar a metodologia desenvolvida no SIF de estimação, utilizou-se como dados de treinamento duas semanas do total de dados disponíveis. Estas duas semanas, destacadas entre as linhas vermelhas da Figura 4-5, contém o valor mínimo e o valor máximo da demanda dentre toda a série temporal. Os dados de testes utilizados são toda a série temporal.

Figura 4-5 – Dados de treinamento do SIF.

Calculou-se então, segundo a metodologia, a autocorrelação dos dados de treinamento. Percebe-se, por meio da Figura 4-6, que o instante que apresenta a maior autocorrelação com o instante presente é o que tem atraso igual a 1 (equivalente a 0,08333 horas, ou cinco minutos), e esta autocorrelação é de 0,996. Por isso, a primeira entrada selecionada é $d(t - 1)$

Com isto, calculou-se o vetor $d_n(t)$ segundo a equação (4-1), utilizando $t_1 = 1$. Este vetor foi correlacionado com os dados de treinamento, resultando no gráfico apresentado na Figura 4-7. O instante com a maior correlação é o que tem 2,5 horas de atraso, o que significa que $t_2 = 30$. O valor deste instante, $d(t - 30)$, foi selecionado como segunda entrada do SIF.

Figura 4-6 – Autocorrelação dos dados de treinamento.

Figura 4-7 – Correlação entre o vetor $d_n(t)$ e os dados de treinamento.

Neste momento, deve-se realizar a primeira estimação para depois poder-se prosseguir com a metodologia. Porém, o SIF exige que lhe seja fornecido um número de regras para que o treinamento seja feito adequadamente. Além disso, é necessário fornecer um critério de

parada do processo iterativo. Assim, foram feitos diversos treinamentos e testes, variando-se o número de regras entre 1 e 15, e limitando o número de entradas a 4. Estas entradas foram selecionadas seguindo a metodologia proposta. Nos gráficos e tabelas a seguir encontram-se os dados resumidos de treinamento e teste de todos os SIFs criados.

Tabela 4-1 – MAPEs dos treinamentos seguindo a metodologia desenvolvida.

Número de:		Treinamento					Média
Regras	Entradas	1	2	3	4	5	
1	2	61,4211	61,4027	61,4451	61,4288	61,4150	61,4225
	3	55,8711	60,7214	60,7118	51,2099	60,7277	57,8484
	4	60,7130	60,7070	54,8596	60,7331	52,7592	57,9544
2	2	16,9159	15,6907	8,3332	11,9059	5,8412	11,7374
	3	24,3538	13,3646	12,6273	20,2860	17,0835	17,5430
	4	12,9441	17,0956	19,6791	26,0275	23,8809	19,9254
3	2	5,2091	7,5632	9,7765	8,6040	4,6628	7,1631
	3	5,0276	7,5888	7,5109	8,4220	8,8651	7,4829
	4	11,0063	14,6834	16,0652	15,5805	10,0121	13,4695
4	2	5,6648	5,4650	10,0211	9,7185	9,5619	8,0863
	3	8,9711	6,4970	6,2026	7,9539	10,9707	8,1191
	4	11,3367	11,2550	9,2208	10,1059	10,1413	10,4119
5	2	7,8488	8,6206	4,7215	4,9977	4,3889	6,1155
	3	10,2369	13,5893	6,7565	15,5243	11,3306	11,4875
	4	9,7757	11,0506	5,7116	11,5932	11,5092	9,9281
6	2	4,1717	5,5944	7,8073	4,6066	4,5906	5,3541
	3	5,1133	6,4368	4,5995	6,2373	7,3506	5,9475
	4	12,6359	8,3027	11,4245	7,3488	11,5534	10,2531
7	2	5,6541	5,5509	4,8719	4,6053	6,2424	5,3849
	3	5,7505	11,8752	8,9881	7,8685	10,2383	8,9441
	4	13,2348	9,1372	9,3576	7,2644	14,8301	10,7648
8	2	5,0839	4,2793	5,0606	4,1400	4,0687	4,5265
	3	11,8329	10,6152	5,3919	10,6069	6,7930	9,0480
	4	6,9611	6,9131	11,3439	10,4248	10,9788	9,3243
9	2	4,3877	4,9214	4,8339	4,7411	4,6063	4,6981
	3	5,5196	10,5535	6,7591	10,3802	10,6637	8,7752
	4	5,8299	7,8632	5,2663	6,8144	5,4998	6,2547
10	2	4,9970	4,8181	4,6020	4,5605	7,7092	5,3374
	3	4,8675	4,8102	4,7580	5,2978	8,5534	5,6574
	4	5,9181	5,0382	5,6800	5,5808	11,1571	6,6748
11	2	4,8742	4,7596	4,5091	4,5499	5,1031	4,7592
	3	4,6793	4,7565	4,3981	4,9102	6,1902	4,9869
	4	8,0667	10,2902	5,4062	6,7739	6,2634	7,3601
12	2	4,8928	4,7541	4,0846	4,5149	4,8438	4,6180
	3	5,9545	5,8429	5,2797	7,7155	5,1457	5,9877
	4	7,6256	11,1372	6,4887	11,9713	11,3094	9,7064

Tabela 4-1 – MAPEs dos treinamentos seguindo a metodologia desenvolvida.

Número de:		Treinamento					Média
Regras	Entradas	1	2	3	4	5	
13	2	5,3838	4,5670	4,5442	7,0210	4,5899	5,2212
	3	4,8088	4,8927	4,5639	4,7161	4,5537	4,7070
	4	4,7971	6,7525	6,4684	5,3551	6,0089	5,8764
14	2	6,5255	4,5845	5,9363	5,2068	5,2993	5,5105
	3	5,3998	4,7168	5,1014	4,8972	4,3257	4,8882
	4	5,0293	6,6976	5,6128	5,5194	5,2732	5,6265
15	2	4,7980	4,6659	4,6847	4,8375	4,7314	4,7435
	3	5,7295	4,8471	5,1444	5,1973	4,8675	5,1572
	4	5,1540	5,9601	5,9890	5,2149	5,9392	5,6514

Figura 4-8 – Média dos MAPEs dos treinamentos seguindo a metodologia desenvolvida (última coluna da Tabela 4-1).

Tabela 4-2 – MAPEs dos testes seguindo a metodologia desenvolvida.

Número de:		Teste					Média
Regras	Entradas	1	2	3	4	5	
1	2	70,3754	70,3388	70,4231	70,3908	70,3633	70,3783
	3	69,4781	73,6037	73,5838	59,2351	73,6167	69,9035
	4	73,5862	73,5737	66,5751	73,6281	62,4397	69,9606
2	2	16,8679	12,9007	8,3741	11,9966	6,8178	11,3914
	3	24,1753	12,5636	14,2855	18,7177	16,0953	17,1675
	4	9,4486	12,8067	14,4910	23,9154	23,9672	16,9258

Tabela 4-2 – MAPEs dos testes seguindo a metodologia desenvolvida.

Número de:		Teste					Média
Regras	Entradas	1	2	3	4	5	
3	2	5,9115	8,6318	10,9348	10,5676	5,3275	8,2746
	3	5,7674	7,8426	7,2747	7,6904	10,5209	7,8192
	4	11,1658	14,0829	16,7534	16,0729	9,6798	13,5510
4	2	6,0699	6,0201	9,4607	9,0151	8,8985	7,8929
	3	9,7476	7,8685	7,2645	7,3057	10,9966	8,6366
	4	10,2839	13,3442	9,1146	11,0489	11,9680	11,1519
5	2	9,3256	8,0288	5,3238	5,6027	5,0377	6,6637
	3	9,8914	13,9197	6,7974	16,9847	10,1046	11,5396
	4	10,7781	11,0514	7,5720	13,5939	14,1403	11,4271
6	2	4,6702	7,0045	7,7705	5,3256	5,0637	5,9669
	3	5,7632	7,0525	5,4281	7,1632	8,4856	6,7785
	4	12,7864	9,6669	11,6517	7,8377	12,7512	10,9388
7	2	6,7112	6,5196	5,4015	5,0162	6,4256	6,0148
	3	7,1732	12,2407	8,6356	8,9552	10,1511	9,4312
	4	16,3689	9,0348	8,9951	7,0324	16,2392	11,5341
8	2	5,8230	4,7174	5,5411	4,6964	4,6716	5,0899
	3	11,9227	11,0304	6,5381	10,7933	7,5458	9,5661
	4	7,4531	9,1977	11,9482	10,5750	10,2668	9,8882
9	2	4,8617	5,4863	5,8192	5,2778	5,6151	5,4120
	3	6,3430	10,8941	6,9277	10,2783	11,1345	9,1155
	4	6,7150	9,5642	5,5589	8,4621	6,2408	7,3082
10	2	5,9081	5,3074	5,3774	5,0839	7,8525	5,9059
	3	5,8546	5,4937	5,9264	6,1515	8,1153	6,3083
	4	6,4574	6,0738	6,7233	7,3550	11,1496	7,5518
11	2	5,7401	5,3611	4,9073	5,3369	5,8660	5,4423
	3	5,7002	4,8838	5,3532	6,1141	7,8934	5,9889
	4	8,5057	10,8179	7,7010	8,2311	6,8575	8,4226
12	2	5,5523	5,5177	4,5965	4,9701	5,5317	5,2337
	3	6,5304	6,6458	6,1708	7,4267	5,7730	6,5093
	4	8,5104	11,8560	6,7742	12,6419	11,7048	10,2975
13	2	5,5029	5,5195	5,3368	7,6089	5,4179	5,8772
	3	5,4962	6,0075	5,0227	6,0480	5,7920	5,6733
	4	6,1645	7,7426	7,8796	7,0815	8,0939	7,3924
14	2	6,8816	5,5926	6,2224	5,7469	5,8461	6,0579
	3	6,8350	5,9199	6,0463	5,8454	5,3490	5,9991
	4	7,0384	7,3897	6,7501	6,4407	6,2676	6,7773
15	2	5,6786	5,1917	5,7644	6,0505	5,3743	5,6119
	3	7,1952	6,2682	6,3289	6,4754	5,9803	6,4496
	4	6,3847	6,5295	6,9946	6,3215	7,4636	6,7388

Figura 4-9 – Média dos MAPEs dos testes seguindo a metodologia desenvolvida (última coluna da Tabela 4-2).

Tabela 4-3 – Atrasos das entradas selecionadas seguindo a metodologia desenvolvida.

		Regras																			
		1				2				3				4				5			
		t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
Teste	1	1	30	196	9	1	30	210	190	1	30	165	195	1	30	59	143	1	30	143	59
	2	1	30	196	195	1	30	183	184	1	30	162	62	1	30	177	59	1	30	59	134
	3	1	30	196	195	1	30	168	207	1	30	148	59	1	30	57	149	1	30	165	59
	4	1	30	196	195	1	30	213	180	1	30	166	184	1	30	59	60	1	30	234	211
	5	1	30	196	195	1	30	186	184	1	30	250	158	1	30	59	215	1	30	65	69
		Regras																			
		6				7				8				9				10			
		t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
Teste	1	1	30	158	178	1	30	143	177	1	30	65	59	1	30	65	146	1	30	153	149
	2	1	30	173	206	1	30	65	64	1	30	165	65	1	30	166	57	1	30	143	158
	3	1	30	62	151	1	30	234	260	1	30	166	171	1	30	166	165	1	30	159	158
	4	1	30	171	41	1	30	243	229	1	30	146	59	1	30	162	59	1	30	165	59
	5	1	30	211	156	1	30	69	57	1	30	149	139	1	30	162	59	1	30	65	59
		Regras																			
		11				12				13				14				15			
		t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
Teste	1	1	30	162	247	1	30	233	178	1	30	59	57	1	30	65	170	1	30	158	176
	2	1	30	243	242	1	30	230	206	1	30	165	166	1	30	166	149	1	30	57	243
	3	1	30	59	151	1	30	213	202	1	30	254	128	1	30	63	159	1	30	159	59
	4	1	30	247	176	1	30	250	59	1	30	65	158	1	30	167	171	1	30	166	178
	5	1	30	173	172	1	30	68	230	1	30	143	243	1	30	65	171	1	30	237	159

Figura 4-10 – Número de vezes que cada atraso, com exceção aos referentes à primeira e à segunda entradas, foi selecionado.

Dentre todos os resultados apresentados, vale a pena detalhar o comportamento do melhor SIF treinado para o caso sob estudo. Observando os gráficos das Figuras 4-8 e 4-9, nota-se que para este caso, apresentar ao SIF apenas duas entradas elegidas segundo a metodologia é a melhor opção, pois no geral, o SIF com duas entradas apresenta menor erro de estimação.

É importante ressaltar que, segundo a Tabela 4-3 e a Figura 4-10, existe uma tendência a se escolher como alguma das próximas entradas do SIF o valor do instante da série temporal com atraso igual a 59, já que esta quantidade de atrasos é a moda entre as terceiras e também entre as quartas entradas selecionadas. Assim, se o melhor SIF fosse algum com três entradas, seria muito provável que esta terceira entrada fosse o valor do instante com 59 atrasos.

Mas este não é o caso. Para esta aplicação, é suficiente definir como SIF que melhor estima a série temporal estudada aquele que apresenta menor erro de previsão com duas entradas. Tem-se então duas excelentes opções, que são o SIF do teste 5 com 8 regras, e o SIF do teste 3 com 12 regras. Os resultados apresentados a seguir referem-se ao primeiro SIF citado, que será explorado mais a fundo unicamente por ser mais simples que o segundo SIF citado.

As funções de pertinência com os parâmetros otimizados para a primeira entrada, para a segunda entrada e para a saída aparecem respectivamente nas Figuras 4-11, 4-12 e 4-13. Já a base de regras otimizada criada para este SIF está na Figura 4-14.

Figura 4-11 – Funções de pertinência otimizadas para a primeira entrada do melhor SIF treinado seguindo a metodologia desenvolvida.

Figura 4-12 – Funções de pertinência otimizadas para a segunda entrada do melhor SIF treinado seguindo a metodologia desenvolvida.

Figura 4-13 – Funções de pertinência otimizadas para a saída do melhor SIF treinado seguindo a metodologia desenvolvida.

1. Se (entrada_1 é mf_1) então (saida_1 é mf_1) (0,48262)
2. Se (entrada_1 é mf_2) e (entrada_2 é mf_2) então (saida_1 é mf_1) (0,9747)
3. Se (entrada_1 é mf_3) e (entrada_2 é mf_3) então (saida_1 é mf_3) (1,2679)
4. Se (entrada_1 é mf_3) e (entrada_2 é mf_2) então (saida_1 é mf_3) (0,60602)
5. Se (entrada_1 é mf_2) e (entrada_2 é mf_3) então (saida_1 é mf_2) (1,3808)
6. Se (entrada_1 é mf_2) e (entrada_2 é mf_3) então (saida_1 é mf_1) (1,4355)
7. Se (entrada_1 é mf_2) e (entrada_2 é mf_3) então (saida_1 é mf_1) (1,3609)
8. Se (entrada_1 é mf_4) e (entrada_2 é mf_3) então (saida_1 é mf_4) (0,84569)

Figura 4-14 – Base de regras otimizada para o melhor SIF treinado seguindo a metodologia desenvolvida.

Com esta configuração, o SIF retornou o erro de estimação disposto no histograma da Figura 4-15. Percebe-se que seu comportamento assemelha-se a uma distribuição normal com média zero, o que indica que a estimação não está enviesada. O desvio padrão do erro de previsão é de 0,0127 pu.

Figura 4-15 – Histograma do erro da estimação realizada seguindo a metodologia desenvolvida.

Por fim, como último resultado interessante, deve-se comentar sobre a propagação do erro de estimação. Tanto o treinamento quanto o teste do SIF consideraram a previsão do valor de apenas um instante da série temporal, mas esta não é uma prática comum quando se estima séries temporais do setor elétrico. É bastante mais comum estimar um horizonte mais longo, como um dia, por exemplo.

A Figura 4-16 exibe como se propagaria o erro de previsão para o caso estudado se o SIF escolhido fosse usado para estimar um dia de demanda de potência ativa. Percebe-se que este erro permanece inferior a 6% durante todo o horizonte de estimação, o que classifica a estimação como satisfatória do ponto de vista do setor elétrico. Isto aconteceu pois o desvio padrão do erro de estimação foi pequeno o suficiente.

Figura 4-16 – Propagação do erro da estimação realizada seguindo a metodologia desenvolvida.

4.5 Comparação dos resultados

Para verificar a eficiência da metodologia desenvolvida é necessário comparar os resultados obtidos por meio dela com resultados alcançados sem empregá-la. Por isto, o mesmo procedimento de treinamento foi realizado uma segunda vez, com a única diferença de que as entradas foram os valores dos instantes da série temporal imediatamente anteriores ao instante presente. Como anteriormente, a quantidade de regras variou entre 1 e 15 e o número de

entradas entre 2 e 4, e também foram feitos 5 treinamentos para cada configuração. O conjunto de treinamento é o mesmo da Figura 4-5.

As Figuras 4-17 e 4-18 exibem, respectivamente, as médias dos MAPEs de treinamento e teste para todos os SIFs criados sem os critérios de seleção de entradas.

Figura 4-17 – Média dos MAPEs dos treinamentos que não seguem a metodologia desenvolvida.

Figura 4-18 – Média dos MAPEs dos testes que não seguem a metodologia desenvolvida.

Há duas principais diferenças perceptíveis ao comparar estas Figuras com as Figuras 4-8 e 4-9. A primeira é que no geral, a média dos erros é ligeiramente inferior quando se aplica a metodologia. A segunda diferença é que existe maior discrepância entre os erros ao variar-se o número de entradas seguindo a metodologia.

Ambas as diferenças existentes na comparação das médias dos MAPEs de treinamento e teste acontecem porque no caso onde a metodologia não foi aplicada as entradas fornecidas aos SIFs contêm quase que a mesma informação, devido ao fato de serem altamente correlacionadas entre si (vide Figura 4-6). Assim, existe informação redundante inserida, o que faz com que o erro não se altere muito ao se modificar o a quantidade de entradas, e ao mesmo tempo há menos informação sobre o comportamento da série temporal no conjunto de entradas, o que provoca um erro de estimação maior.

Mesmo já se notando que a metodologia no geral melhora as estimações, deve-se investigar mais a fundo o melhor SIF que não a segue. Dentre os 5 testes realizados com cada configuração, o melhor SIF é aquele criado no quinto teste que contém 14 regras e 4 entradas.

As Figuras 4-19, 4-20, 4-21, 4-22, 4-23 e 4-24 exibem, respectivamente, as funções de pertinência da primeira à quarta entrada, da saída e a base de regras criada para este SIF.

Figura 4-19 – Funções de pertinência otimizadas para a primeira entrada do melhor SIF treinado que não segue a metodologia desenvolvida.

Figura 4-20 – Funções de pertinência otimizadas para a segunda entrada do melhor SIF treinado que não segue a metodologia desenvolvida.

Figura 4-21 – Funções de pertinência otimizadas para a terceira entrada do melhor SIF treinado que não segue a metodologia desenvolvida.

Figura 4-22 – Funções de pertinência otimizadas para a quarta entrada do melhor SIF treinado que não segue a metodologia desenvolvida.

Figura 4-23 – Funções de pertinência otimizadas para a saída do melhor SIF treinado que não segue a metodologia desenvolvida.

1. Se (entrada_1 é mf_1) ou (entrada_2 é mf_1) ou (entrada_3 é mf_1) ou (entrada_4 é mf_1) então (saida_1 é mf_1) (0,97413)
2. Se (entrada_1 é mf_2) e (entrada_2 é mf_2) e (entrada_3 é mf_2) e (entrada_4 é mf_2) então (saida_1 é mf_2) (0,62747)
3. Se (entrada_1 é mf_3) e (entrada_2 é mf_3) e (entrada_3 é mf_3) e (entrada_4 é mf_3) então (saida_1 é mf_3) (0,94169)
4. Se (entrada_1 é mf_4) e (entrada_2 é mf_3) e (entrada_3 não é mf_1) então (saida_1 é mf_4) (0,99816)
5. Se (entrada_1 é mf_3) e (entrada_2 é mf_4) e (entrada_3 é mf_2) e (entrada_4 é mf_2) então (saida_1 é mf_3) (1,0669)
6. Se (entrada_1 é mf_3) e (entrada_2 é mf_3) e (entrada_3 é mf_3) e (entrada_4 é mf_3) então (saida_1 é mf_2) (0,76436)
7. Se (entrada_1 é mf_3) e (entrada_2 é mf_1) e (entrada_3 não é mf_1) e (entrada_4 é mf_1) então (saida_1 é mf_4) (1,0994)
8. Se (entrada_1 é mf_3) e (entrada_2 é mf_2) e (entrada_3 é mf_2) e (entrada_4 é mf_3) então (saida_1 é mf_1) (1,5245)
9. Se (entrada_1 é mf_2) e (entrada_2 não é mf_1) e (entrada_3 é mf_2) e (entrada_4 é mf_3) então (saida_1 é mf_3) (0,84085)
10. Se (entrada_1 não é mf_2) e (entrada_2 é mf_4) e (entrada_3 é mf_3) e (entrada_4 é mf_2) então (saida_1 não é mf_1) (0,57909)
11. Se (entrada_1 não é mf_1) e (entrada_2 não é mf_2) e (entrada_3 é mf_4) e (entrada_4 é mf_1) então (saida_1 não é mf_1) (1,8864)
12. Se (entrada_1 é mf_2) e (entrada_2 é mf_2) e (entrada_3 é mf_2) e (entrada_4 é mf_4) então (saida_1 é mf_1) (1,2784)
13. Se (entrada_1 é mf_2) e (entrada_2 é mf_2) e (entrada_3 não é mf_1) e (entrada_4 é mf_3) então (saida_1 é mf_1) (1,992)
14. Se (entrada_1 é mf_4) e (entrada_2 é mf_2) e (entrada_3 é mf_1) e (entrada_4 é mf_2) então (saida_1 é mf_4) (1,211)

Figura 4-24 – Base de regras otimizada para o melhor SIF treinado que não segue a metodologia desenvolvida.

Já a Figura 4-25 ilustra o histograma do erro de estimação deste SIF treinado sem critério de seleção de entradas. Percebe-se que o erro ainda se assemelha a um ruído branco, porém o desvio padrão é maior que se comparado ao da Figura 4-15, sendo esse igual a 0,0140 pu.

Figura 4-25 – Histograma do erro da estimação realizada que não segue a metodologia desenvolvida.

Como último comparativo, dispõe-se do erro propagado para a estimação com horizonte de um dia, apresentado na Figura 4-26. Diferentemente do caso onde foi aplicada a metodologia (em verde, e também na Figura 4-16), o erro atingido por este SIF (em azul) foi alto demais, atingindo um MAPE de quase 20%. Percebe-se então que o pequeno aumento no desvio padrão do erro de estimação do valor de um único instante acabou por provocar um grande acúmulo no erro para maiores horizontes.

Figura 4-26 – Propagação do erro da estimativa realizada que não segue a metodologia desenvolvida.

5 CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS

Foi demonstrado que os resultados obtidos escolhendo-se as entradas por meio da metodologia desenvolvida neste trabalho foram melhores que os resultados obtidos quando as entradas são os valores dos instantes mais recentes da série temporal, que é a forma mais comum de se eleger entradas de um sistema de estimação encontrada na literatura.

Outro ponto positivo da metodologia é que o melhor SIF encontrado ao aplicá-la é mais simples que o melhor encontrado sem sua aplicação, ou seja, possui menor quantidade de regras e necessita de menos entradas. E ainda assim, os erros de estimação continuaram menores quando comparados ao SIF sem a metodologia aplicada.

Conclui-se, então, que a metodologia desenvolvida para selecionar as entradas de um SIF de estimação foi funcional e eficiente quando aplicada no caso estudado. Além disso, pode-se afirmar também que os objetivos do trabalho foram alcançados com êxito, já que foi possível estimar a demanda de potência ativa de um sistema de distribuição de energia elétrica mantendo o nível de erro dentro de valores considerados aceitáveis pelo setor elétrico.

Mas ainda existem futuras investigações que valem a pena ser efetuadas. Para este mesmo caso estudado seria interessante investigar como a metodologia se comporta quando aplicada a outros métodos de estimação de séries temporais, como RNAs, por exemplo. Ainda no mesmo caso, e também com o mesmo SIF de estimação, pode-se comparar esta metodologia de escolha de entradas com outras existentes na literatura, inclusive analisando a robustez das estimações realizadas. Por fim, pode-se aplicar esta metodologia para estimar outras séries temporais, e assim verificar sob quais configurações e situações ela continua sendo funcional e eficiente.

REFERÊNCIAS

- [1] S. C. Chan, K. M. Tsui, H. C. Wu, Y. Hou, Y.-C. Wu e F. F. Wu, “Load Price Forecasting and Managing Demand Response for Smart Grids - Methodologies and Challenges,” *IEEE Signal Processing Magazine*, vol. 9, pp. 68-85, 2012.
- [2] L. J. Ricalde, B. Cruz, G. Catzin, A. Y. Alanis e E. N. Sanchez, “Forecasting for Smart Grid Applications with Higher Order Neural Networks,” em *World Automation Congress*, Puerto Vallarta, México, 2012.
- [3] H.-t. Zhang, F.-y. Xu e L. Zhou, “Artificial Neural Network for Load Forecasting in Smart Grid,” em *Proceedings of the 9th International Conference on Machine Learning and Cybernetics*, Qingdao, China, 2010.
- [4] N. Amjady e A. Daraeepour, “Midterm Demand Prediction of Electrical Power Systems Using a New Hybrid Forecast Technique,” *IEEE Transactions on Power Systems*, vol. 26, nº 2, pp. 755-765, Maio 2011.
- [5] A. A. E. Desouky e M. M. E. Kateb, “Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA,” *IEE Proceedings - Generation, Transmission & Distribution*, vol. 147, pp. 213-217, 2000.
- [6] E. G. Romera, M. A. J. Moran e D. C. Fernandez, “Monthly Electric Energy Demand Forecasting Based on Trend Extraction,” *IEEE Transactions on Power Systems*, vol. 21, nº 4, pp. 1946-1953, 2006.
- [7] P.-F. Pai e W.-C. Hong, “Forecasting Regional Electricity Load Based on Recurrent Support Vector Machines With Genetic Algorithms,” *Electric Power Systems Research*, vol. 74, nº 3, pp. 417-425, 2005.
- [8] R. A. Flauzino, “Identificação e Localização de Faltas de Alta Impedância em Sistemas de Distribuição Baseadas em Decomposição por Componentes Ortogonais e Inferência Fuzzy,” Tese de Doutorado, Universidade de São Paulo, São Carlos, Brasil, 2007.
- [9] P. J. Brockwell e R. A. Davis, *Introduction to Time Series and Forecasting*, Nova York, Estados Unidos: Springer, 2002.
- [10] S. M. Kendall e J. K. Ord, *Time Series*, Nova York, Estados Unidos: John Wiley and Sons, 1990.
- [11] J. S. Armstrong e F. Collopy, “Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons,” *International Journal of Forecasting*, pp. 69-80, Junho 1992.
- [12] Z. Chen e Y. Yang, “Assessing Forecast Accuracy Measures,” 14 Março 2004. [Online]. Available: <http://www.stat.iastate.edu/preprint/articles/2004-10.pdf>. [Acesso em 08 Maio 2014].

- [13] “How to Compare Models,” [Online]. Available: <http://people.duke.edu/~rnau/compare.htm>. [Acesso em 15 Abril 2013].
- [14] F. d. O. Lemos, “Metodologia para Seleção de Métodos de Previsão de Demanda,” Dissertação de Mestrado, Porto Alegre, Brasil, 2006.
- [15] L. A. Aguirre, Introdução à Identificação de Sistemas - Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais, Belo Horizonte, Brasil: Editora UFMG, 2004.
- [16] G. E. P. Box, G. M. Jenkins e G. C. Reinsel, Time Series Analysis - Forecasting and Control, Englewood Cliffs, Estados Unidos: Prentice-Hall, 1994.
- [17] T. Masters, Neural, Novel and Hybrid Algorithms for Time Series Prediction, Nova York, Estados Unidos: John Wiley and Sons, 1995.
- [18] X. Zhang, J. Zhang, Y. Li e R. Zhang, “Short-Term Forecasting of Wind Speed Based on Recursive Least Squares,” em *International Conference on Electrical and Control Engineering*, Yichang, China, 2011.
- [19] S. Buhan, E. Terciyanli, M. B. Özkan, H. Tuna e H. Ünsal, “Verification of a Very Short Term Wind Power Forecasting Algorithm for Turkish Transmission Grid,” em *4th International Conference on Power Engineering, Energy and Electrical Drives*, Istambul, Turquia, 2013.
- [20] B. Kedem e K. Fokianos, Regression Models for Time Series Analysis, Hoboken, Estados Unidos: John Wiley and Sons, 2002.
- [21] H.-T. Yang e C.-M. Huang, “A New Short-Term Load Forecasting Approach Using Self-organizing Fuzzy ARMAX Models,” *IEEE Transactions on Power Systems*, vol. 13, nº 1, pp. 217-225, 1998.
- [22] Y. Cao, Y. Liu, D. Zhang, W. Wang e Z. Chen, “Wind Power Ultra-Short-Term Forecasting Method Combined with Pattern-Matching and ARMA-Model,” em *PowerTech*, Grenoble, França, 2013.
- [23] K. Maciejowska e R. Weron, “Forecasting of Daily Electricity Spot Prices by Incorporating Intra-Day Relationships: Evidence Form the UK Power Market,” em *10th International Conference on the European Energy Market*, Estocolmo, Suécia, 2013.
- [24] S. A. Billings, Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains, Sheffield, Inglaterra: John Wiley and Sons, 2013.
- [25] L. H. Tsoukalas e R. E. Uhrig, Fuzzy and Neural Approaches in Engineering, Nova York, Estados Unidos: John Wiley and Sons, 1997.
- [26] I. N. d. Silva, D. H. Spatti e R. A. Flauzino, Redes Neurais Artificiais para Engenharia e Ciências Aplicadas - Curso Prático, São Paulo, Brasil: Artliber, 2010.

- [27] J. Yasuoka, J. L. P. Brittes, H. P. Schmidt e J. A. Jardini, "Artificial Neural Network-Based Distribution Substation and Feeder Load Forecast," em *16th International Conference and Exhibition on Electricity Distribution*, Amsterdã, Holanda, 2001.
- [28] K. B. Sahay e M. M. Tripathi, "Day Ahead Hourly Load Forecast of PJM Electricity Market and ISO New England Market by using Artificial Neural Network," em *IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, Washington, Estados Unidos, 2014.
- [29] W. Jin-ming e L. Xin-heng, "The Forecast of Energy Demand on Artificial Neural Network," em *International Conference on Artificial Intelligence and Computational Intelligence*, Xangai, China, 2009.
- [30] H.-J. Zimmermann, *Fuzzy Set Theory and Its Applications*, Norwell, Estados Unidos: Kluwer Academic Publishers, 1996.
- [31] M. Brown e C. Harris, *Neurofuzzy Adaptive Modelling and Control*, Londres, Inglaterra: Prentice Hall, 1994.
- [32] C. N. Elias e N. D. Hatziargyriou, "An Annual Midterm Energy Forecasting Model Using Fuzzy Logic," *IEEE Transactions on Power Systems*, vol. 24, nº 1, pp. 469-478, 2009.
- [33] Y.-D. Gu, J.-Z. Cheng e Z.-Y. Wang, "An Fuzzy Forecasting Algorithm for Short Term Electricity Loads Based on Partial Clustering," em *International Conference on Machine Learning and Cybernetics*, Guilin, China, 2011.
- [34] M. F. I. b. Khamis, Z. b. Baharudin, N. H. b. Hamid, M. F. b. Abdullah e S. S. M. Yunus, "Electricity Forecasting for Small Scale Power System Using Fuzzy Logic," em *9th International Power & Energy Conference*, Singapura, Singapura, 2010.
- [35] W. Yuill, R. Kgokong, S. Chowdhury e S. P. Chowdhury, "Application of Adaptive Neuro Fuzzy Inference System (ANFIS) based Short Term Load Forecasting in South African Power Networks," em *45th Universities' Power Engineering Conference*, Cardiff, País de Gales, 2010.
- [36] H. Zhou, X. H. Wu e X. G. Li, "An ANFIS model of Electricity Price Forecasting Based on Subtractive Clustering," em *Power and Energy Society General Meeting*, San Diego, Estados Unidos, 2011.
- [37] O. E. Dragomir, F. Dragomir, R. Gouriveau e E. Minca, "Medium Term Load Forecasting Using ANFIS Predictor," em *18th Mediterranean Conference on Control & Automation*, Marrakech, Marrocos, 2010.
- [38] E. M. C. Franco, "Previsão Espacial de Demanda em Sistemas de Distribuição com uma Base Reduzida de Dados," Tese de Doutorado, Universidade Estadual Paulista "Júlio de Mesquita Filho", Ilha Solteira, Brasil, 2008.
- [39] J.-S. R. Jang, C.-T. Sun e E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Englewood Cliffs, Estados Unidos: Prentice Hall, 1997.

- [40] J. E. Dennis e R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Englewood Cliffs, Estados Unidos: Prentice Hall, 1983.
- [41] W. Pedrycz e F. A. C. Gomide, An Introduction to Fuzzy Sets – Analysis and Design, Cambridge, Estados Unidos: MIT Press, 1998.