

TESTES DE MODELAGENS PARA PREVISÃO DE UMA SÉRIE TEMPORAL DE DISTRIBUIÇÃO DE ENERGIA ELÉTRICA

Cesar Machado Pereira

Universidade do Estado do Rio de Janeiro – UERJ. Departamento de Engenharia Eletrônica,
UERJ. São Francisco Xavier, 524, Rio de Janeiro, 20550-900, Brasil.

cesarpmachado@gmail.com

Nival Nunes de Almeida

Universidade do Estado do Rio de Janeiro – UERJ. Departamento de Engenharia Eletrônica,
UERJ. São Francisco Xavier, 524, Rio de Janeiro, 20550-900, Brasil.

nivalnunes@yahoo.com.br

Maria Luiza Fernandes Velloso

Universidade do Estado do Rio de Janeiro – UERJ. Departamento de Engenharia Eletrônica,
UERJ. São Francisco Xavier, 524, Rio de Janeiro, 20550-900, Brasil.

mlfv@centroin.com.br

Resumo

Este trabalho testa e compara dois tipos de modelagem para previsão de uma mesma série temporal. Foi observada uma série temporal de distribuição de energia elétrica e, como estudo de caso, optou-se pela região metropolitana do Estado da Bahia. Foram testadas as combinações de três variáveis exógenas em cada modelo: a quantidade de clientes ligados na rede de distribuição de energia elétrica, a temperatura ambiente e a precipitação de chuvas. O modelo linear de previsão de séries temporais utilizado foi um SARIMAX. A modelagem de inteligência computacional utilizada para a previsão da série temporal foi um sistema de Inferência Fuzzy. Na busca de um melhor desempenho, foram feitos testes de quais variáveis exógenas melhor influenciam no comportamento da energia distribuída em cada modelo. Segundo a avaliação dos testes, o sistema Fuzzy de previsão foi o que obteve o menor erro. Porém dentre os menores erros, os resultados dos testes também indicaram diferentes variáveis exógenas para cada modelo de previsão.

Palavras-Chaves: Previsão; Séries Temporais; Energia Elétrica; SARIMAX; Sistemas Fuzzy.

Abstract

This paper tests and compares two types of modelling to predict the same time series. A time series of electric load was observed and, as a case study, we opted for the metropolitan region of Bahia State. The combination of three exogenous variables were attempted in each model. The exogenous variables are: the number of customers connected to the electricity distribution network, the temperature and the precipitation of rain. The linear model time series forecasting used was a SARIMAX. The modelling of computational intelligence used to predict the time series was a Fuzzy Inference System. According to the evaluation of the attempts, the Fuzzy forecasting system presented the lowest error. But among the smallest errors, the results of the attempts also indicated different exogenous variables for each forecast model.

Keywords: Forecast; Time Series; Electric Load; SARIMAX; Fuzzy Inference System.

1. DISTRIBUIÇÃO DE ENERGIA ELÉTRICA

A energia elétrica é essencial no dia-a-dia da sociedade, seja nas residências ou nos diversos segmentos da economia. Atividades rotineiras como assistir televisão ou o uso do ar condicionado só são possíveis através da distribuição da energia elétrica.

O segmento de distribuição de energia recebe grande quantidade de energia do sistema de transmissão e a distribui para os consumidores finais, residências, pequenos comércios e indústrias. O fornecimento de energia é um dos serviços mais desafiantes da sociedade moderna. Para que o consumidor disponha de energia no momento que aciona um interruptor ou conecta um aparelho elétrico na tomada em sua residência é preciso que todo o sistema em cadeia esteja apto a operar de forma coordenada, com centrais geradoras, linhas de transmissão, subestações, linhas e transformadores de distribuição.

A Companhia de Eletricidade do Estado da Bahia (Coelba) é responsável pela distribuição de energia elétrica no Estado da Bahia, Brasil, desde 1960. Ela é a terceira maior distribuidora do Brasil em clientes e a sexta maior em volume de energia fornecida.

1.1. OBJETIVO

A série temporal objeto de estudo é a de distribuição de energia elétrica da região Metropolitana da Coelba, usualmente chamada de carga. A série temporal tem periodicidade diária, unidade de Megawatts-hora (MWh) e foram utilizados somente os dias de semana, descartando assim os finais de semana (sábados e domingos) da série temporal. Os feriados foram mantidos na série temporal, exceto os que são nos finais de semana. A série temporal de energia elétrica com início em janeiro de 2010 e fim em outubro de 2013 é composta por 983 pontos e apresentada abaixo.

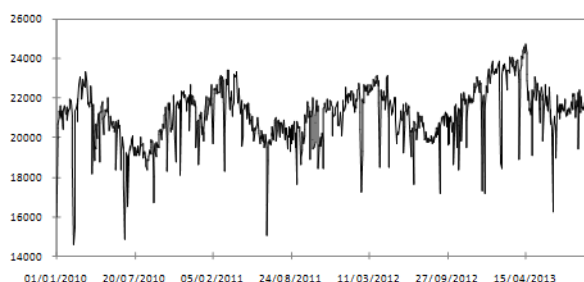


FIGURA 1: Série de energia distribuída da região Metropolitana sem finais de semana

Este trabalho tem por objetivo comparar duas modelagens para previsão de séries temporais. As modelagens comparadas são: uma linear SARIMAX e uma de Sistema de Inferência Fuzzy. Ambos os modelos são aplicados para a previsão da mesma série temporal de distribuição de energia elétrica da região Metropolitana do Estado da Bahia.

Para fins de comparação, se propõe assim realizar testes em ambos os modelos com o uso de diferentes combinações de três variáveis exógenas. As variáveis exógenas indicadas para ambos os modelos foram: a temperatura ambiente, a precipitação de chuvas e o número de clientes.

As combinações são feitas conforme a tabela abaixo:

Tabela1 – Combinações dos testes

Teste 12	Teste 13	Teste 23	Teste 123
Clientes e Temperatura	Clientes e Precipitação	Temperatura e Precipitação	Clientes, Temperatura e Precipitação

Com o uso das variáveis exógenas, os modelos buscam a previsão de um passo a frente, no dia anterior.

1.2. TRABALHOS RELACIONADOS

Diversos modelos são aplicados para a previsão de energia elétrica. Os modelos

matemáticos ARMA, ARIMA e SARIMAX e os sistemas de inferência Fuzzy, do inglês Fuzzy Inference System (FIS) foram encontrados nos trabalhos acadêmicos citados ao longo desta seção.

O modelo estatístico mais utilizado é o modelo autoregressivo integrado de média móvel (ARIMA). Na área de energia esse modelo está sendo utilizado desde 1978. O autor [32] o utilizou para prever a ocorrência mensal de picos de energia mediante o comportamento de alterações nas variáveis econômicas e climáticas.

Em 1999 [4] utilizaram o modelo de análise de séries temporais Box-Jenkins para previsão da produção e do consumo de energia elétrica. Os países utilizados foram: Áustria e Espanha.

Os autores [24] investigaram diferentes metodologias para se realizar a previsão de energia elétrica em Lebanon. Um de seus modelos foram o ARIMA. Na previsão do preço da geração e do consumo de energia elétrica na Espanha os autores [7] e [6] também utilizaram o modelo ARIMA.

Os estudos de [21] e [25] usaram técnicas de modelagem matemática para prever o comportamento de consumo de energia através de condições meteorológicas, jornadas de trabalho, dias da semana e a hora do dia.

O FIS vem sendo utilizada em modelos de previsão de séries temporais do setor elétrico no curto prazo, séries de meia hora, horárias e diárias. Os trabalhos [13] e [21] são alguns desses exemplos.

Pelas pesquisas de [21] e [25] comprovou-se que o comportamento de consumo de energia, em uma perspectiva de curto prazo, é sensível às condições meteorológicas, a parte da jornada de trabalho, os dias da semana e as horas do dia. Por este motivo, dispor de um bom sistema de previsão de carga, usando essas variáveis climáticas associadas a carga é de suma importância.

Os autores [26-28] propuseram o uso de sistemas Fuzzy em séries temporais nos trabalhos. E os autores [17] propuseram uma série temporal Fuzzy, onde a caracterizam como uma sequência linguística ordenada em termos de tempo.

O trabalho proposto pelos autores [1] utiliza um sistema Fuzzy, através do ANFIS, para a previsão de energia anual sobre países industrializados como Estados Unidos da América, Reino Unido, Japão, França e Itália. As variáveis de entrada do sistema são o produto interno bruto (PIB) e a população.

2. DESCRIÇÃO DAS TÉCNICAS DE MODELAGEM

Existem várias abordagens diferentes para modelagem de séries temporais. O modelo mais usual de séries temporais é o Box-Jenkins. O trabalho de Box-Jenkins denominado “Time Series Analysis: Forecasting and Control” foi criado em 1976 e tem sido utilizado em inúmeros trabalhos acadêmicos [2].

2.1. MODELO AUTOREGRESSIVO INTEGRADO DE MÉDIA MÓVEL (ARIMA)

A partir do operador diferença pode-se construir um modelo composto de uma parte autoregressiva, explicado pelo comportamento passado da própria variável, uma parte de médias móveis, explicado pelos distúrbios de períodos anteriores, e outra definida como integração, onde o número de defasagens (d) transforma a série em uma série estacionária.

O modelo ARIMA(p, d, q) pode ser escrito da seguinte forma [29]:

$$\nabla Z_t = (1 - B)Z_t = w_t \text{ e } \xi(B) = \phi(B)\nabla^d = \phi(B)(1 - B)^d \quad (1)$$

Onde: ∇Z_t é o operador diferença, Z_t é uma integral de w_t , ∇ é um polinômio do tipo AR, w_t é a diferença de Z_t , $\phi(B)$ é um operador autoregressivo não-estacionário de ordem $p + d$, B é o operador de atraso e μ é a expectativa de Y_t , geralmente assumida como 0.

2.2. IDENTIFICAÇÃO DOS MODELOS ARIMA (P, D E Q)

Diversos autores afirmam que para determinar a ordem dos termos p, d e q, as principais ferramentas a serem utilizadas são a função de autocorrelação (FAC), a função de autocorrelação parcial (FACP) e seus respectivos correlogramas [11], [18] e [20].

A FAC de ordem k é uma ferramenta amplamente utilizada nos modelos de Box-Jenkins [8]. Um coeficiente de autocorrelação p_1 mede a correlação entre a variável no período t em relação ao período defasado $t - 1$ e é chamado de coeficiente de autocorrelação de lag 1, assim como um coeficiente de autocorrelação p_k é dito de coeficiente de autocorrelação de lag k [22].

A FAC serve para medir a extensão para qual o valor tomado no tempo t, depende daquele tomado no tempo $t - k$ [29]. Assim o número ótimo de defasagens no modelo pode ser determinado com base nos valores da FAC [19].

A FAC também auxilia na análise de séries temporais, pois é uma importante ferramenta para a investigação das propriedades empíricas da série temporal, como: identificação da condição de estacionariedade e identificação de modelos potenciais a serem utilizados na modelagem e previsão da série temporal [11] e [20].

Assim como a FAC, a FACP também é utilizada para a identificação dos modelos a serem utilizados para a modelagem e previsão da série temporal. A FACP também pode auxiliar na identificação da condição de estacionariedade da série temporal [11], [19] e [20].

A correlação parcial entre duas variáveis y_t e y_{t-k} é a correlação simples existente entre elas menos aquela parte explicada linearmente pelas defasagens intermediárias [14].

2.3. MODELO MULTIPLICATIVO AUTOREGRESSIVO INTEGRADO DE MÉDIA MÓVEL (SARIMA)

A fórmula do modelo SARIMA contém uma parte sazonal na primeira parte da fórmula, e uma parte não sazonal dos parâmetros na segunda parte da fórmula [9].

Como extensão natural, pode-se definir o modelo SARIMA como [29]:

$$(B) \cdot \phi(B^s) \cdot \nabla_s^D \cdot \nabla^d Z_t = \phi(B) \cdot \theta B^s \cdot a_t \quad (2)$$

Onde: $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ é um operador não sazonal autoregressivo, $\nabla^d = (1 - B)^d$ é um operador diferença não sazonal de ordem d, $\phi(B^s) = (1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_p B^{ps})$ é um operador sazonal autoregressivo, $\nabla_s^D = (1 - B^s)^D$ é um operador sazonal de ordem D, $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$ é um operador não sazonal de médias móveis, $\theta(B^s) = (1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_q B^{qs})$ é um operador sazonal de médias móveis, $\nabla_s^D \cdot \nabla^d Z_t = w_t$ é um filtro não linear aplicado a série original Z_t que produz um processo estacionário w_t e a_t é um ruído branco.

2.4. MODELO MULTIPLICATIVO AUTOREGRESSIVO INTEGRADO DE MÉDIA MÓVEL COM VARIÁVEIS EXÓGENAS (SARIMAX)

A fim de melhorar o desempenho de previsão de valores futuros, outra série de tempo pode ser incorporada nos modelos de séries temporais [5].

A adição de uma entrada externa para um modelo é chamado usando uma variável exógena. Por exemplo, o seguinte é um modelo ARMA ordinária com uma única variável exógena, ARMAX.

$$Y_t = \sum_{i=1}^p (\phi_i Y_{t-i}) + \sum_{j=1}^q (\theta_j a_{t-j}) + \sum_{h=1}^b \beta_h X_{t-h} + \omega_t \quad (3)$$

Onde X_t são as variáveis exógenas.

3. SISTEMAS DE INFERÊNCIA FUZZY

Nesta seção está contida a lógica nebulosa, conhecida como Lógica Fuzzy, junto a conceitos de conjuntos Fuzzy, funções de pertinência, as variáveis linguísticas e as operações lógicas. A última subseção destina-se ao Sistema Neuro-Fuzzy proposto.

3.1. LÓGICA FUZZY

A lógica Fuzzy, conhecida como lógica nebulosa, ou difusa, é uma técnica de inteligência computacional que tem como objetivo modelar o modo aproximado de raciocínio do ser humano [3]. Assim é possível tomar decisões em um ambiente de incerteza e imprecisão [33], o que não ocorre com a lógica discreta.

Como a lógica Fuzzy é formada a partir da teoria de conjuntos Fuzzy. As subseções abaixo descrevem esse entendimento por completo.

Formalmente, um conjunto nebuloso do universo de discurso é definido por uma função de pertinência, $\mu_A: U \rightarrow [0,1]$.

Um conjunto Fuzzy é um conjunto sem uma fronteira nítida, ou claramente definida. Na forma clássica a associação de um elemento em relação a um conjunto é assumida com termos binários. Ou seja, o elemento pode somente pertencer ou não pertencer a um conjunto, 0 ou 1.

A função de pertinência é uma curva que define a forma como cada ponto no espaço de entrada é mapeado para um valor de pertinência. A função de pertinência deve estar entre 0 e 1.

É muito importante a escolha do tipo de função de pertinência, pois ela influencia diretamente nos resultados do sistema Fuzzy. Portanto a sua escolha depende do contexto do problema estudado e da experiência do usuário.

Uma variável linguística é aquela cujos valores são palavras em lugar de números [31]. Na lógica Fuzzy, uma variável linguística é uma variável cujos valores são nomes de conjuntos Fuzzy. A principal função das variáveis linguísticas é fornecer uma maneira sistemática para uma caracterização aproximada de fenômenos através da descrição linguística. Isto permite um melhor entendimento do problema que está sendo analisado, permitindo o tratamento de sistemas que são muito complexos para serem analisados através de mecanismos matemáticos convencionais [12].

3.2. O SISTEMA FUZZY

Os Sistemas Fuzzy representam uma das áreas de aplicação mais importantes da Teoria dos Conjuntos Fuzzy, sendo o modelo mais utilizado o baseado em regras. Eles são uma extensão dos sistemas clássicos baseados em regras, porque utilizam regras do tipo “se-então”, onde os antecedentes e consequentes são proposições Fuzzy [15] e [16].

Existem dois tipos de sistemas de inferência que variam conforme o tipo de saída determinado: Mamdani e Sugeno. O Mamdani é usado através da teoria dos conjuntos Fuzzy. O Sugeno é semelhante ao Mamdani em muitos aspectos. A principal diferença é que as funções de pertinência de saída Sugeno são ou linear ou constante [30].

As regras Fuzzy “se-então”, “ou”, “senão” e “e” são regras cuja saída pode ser número ou linguístico. Exemplificando, “se x é A então y é B” como um ponto Fuzzy A x B [33].

A estrutura mais usual de um processo controlado por um controlador Fuzzy é composta por uma interface de fuzzificação, uma unidade de inferência, uma interface de defuzzificação e por uma base de conhecimento, que por sua vez divide-se em Base de dados e a Base de Regras.

3.3. O SISTEMA NEURO-FUZZY

Sistemas Neuro-Fuzzy ganharam recentemente um grande interesse em pesquisa e aplicação. Modelos Neuro-Fuzzy são sistemas Fuzzy que usam estratégias de aprendizagem locais para aprender conjuntos Fuzzy e regras nebulosas. Diversas técnicas de Neuro-Fuzzy foram desenvolvidas para apoiar o desenvolvimento de, por exemplo, controladores Fuzzy e classificadores difusos. O algoritmo de aprendizagem é capaz de extrair regras Fuzzy a partir de um conjunto de dados de treinamento muito rapidamente, e ajusta-los, modificando os

parâmetros de funções de pertinência.

4. DESCRIÇÃO DOS DADOS

Apesar de não haver um consenso sobre as melhores variáveis que descrevem o funcionamento do serviço de distribuição de energia [23], optou-se pelo uso de três variáveis exógenas. A quantidade de clientes, a precipitação de chuvas e a temperatura ambiente foram utilizadas nos modelos.

Os dados disponíveis de normais climatológicas são do Instituto Nacional de Meteorologia (INMET). Os outros dados medidos foram fornecidos pela própria empresa em estudo.

4.1. NÚMERO DE CLIENTES

A variável “clientes” é a quantidade de unidades comerciais ou residenciais ligadas à rede de energia da distribuidora. Definiu-se como entrada do sistema a variação percentual da quantidade de consumidores dentre dois dias consecutivos. A fórmula abaixo destaca essa representação.

$$x_1 = \text{clientes}_t / \text{clientes}_{t-1} - 1 \quad (4)$$

Onde clientes_t é no instante em questão e clientes_{t-1} é no instante imediatamente anterior.

4.2. PRECIPITAÇÃO DE CHUVAS

Para a formação da variável, optou-se pela diferença do valor da precipitação diária em relação às normais climatológicas. Para uso nos modelos: fez-se o cálculo da diferença de chuva medida com a informada pela normal climatológica. A fórmula abaixo destaca essa representação.

$$x_2 = C_{medida} - C_{NC} \quad (5)$$

Onde C_{medida} a precipitação medida no dia em questão e C_{NC} é a normal climatológica de precipitação do mês em questão de Salvador.

4.3. TEMPERATURA AMBIENTE

Para a formação da variável, foi usada a diferença do valor da temperatura diária em relação às normais climatológicas. Para uso nos modelos: fez-se o cálculo de diferença entre a temperatura máxima diária medida e a normal climatológica máxima e a diferença entre a temperatura mínima diária medida e a normal climatológica mínima em cada dia. Pois assim, com a soma entra as diferenças das temperaturas máximas e das temperaturas mínimas obteve-se um único valor de sensibilização da temperatura. A fórmula abaixo destaca essa representação.

$$x_3 = (T_{máx} - NC_{máx}) + (T_{mín} - NC_{mín}) \quad (6)$$

Onde $T_{máx}$ e $T_{mín}$ são as temperaturas máxima e mínima medida no dia em questão. A $NC_{máx}$ e $NC_{mín}$ são as normais climatológicas de temperatura máxima e mínima do mês em questão de Salvador.

5. EXPERIMENTOS E MODELOS OBTIDOS

Antes de apresentar os modelos propostos é importante formalizar a avaliação dos modelos.

5.1. AVALIAÇÃO

Para avaliar os dois algoritmos foram utilizados o erro percentual médio (MAPE) e o desvio padrão (DP).

O MAPE mede, em termos absolutos, o desvio da série original com a série saída do modelo. A fórmula do MAPE é destacada abaixo:

$$M = 1/n \sum_{t=1}^n |(A_t - Y_t)/A_t| \quad (7)$$

Onde A_t é o valor original e Y_t é o valor previsto por cada modelo.

O DP utilizado é a diferença entre a série original e a série prevista pelo modelo. Ele pode ser escrito através da fórmula abaixo:

$$DP = \sqrt{\sum (A_t - \bar{Y}_t)^2 / (n - 1)} \quad (8)$$

A partir de ambos os avaliadores apresentados nesta seção é possível definir o modelo mais preciso para a previsão da série temporal de energia elétrica da região Metropolitana do Estado da Bahia.

5.2. MODELO DE PREVISÃO SARIMAX

As seções a seguir descrevem a formação do modelo SARIMAX a partir da escolha do método OLS. Para AST do modelo SARIMAX foram utilizados todos os 983 dados da série temporal. Esses mesmos 983 dados também foram utilizados para os testes.

5.2.1. Transformação da Série Temporal.

Como a série temporal não era estacionária foi necessário diferenciá-la para torná-la estacionária. Portanto foi necessário realizar a diferença da série proposta.

Como na imagem apresentada abaixo a FAC indica uma sazonalidade de 5 períodos, também foi feito a primeira diferença em 5 períodos.

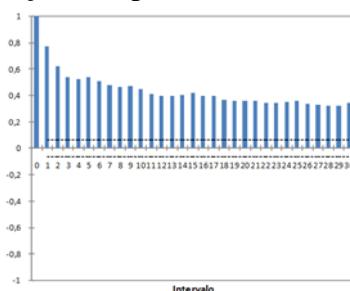


FIGURA 2: FAC

Neste momento é possível a análise da série para a busca dos parâmetros do modelo SARIMAX.

5.2.2. Análise Descritiva.

A partir da figura 3, foi assumida a hipótese de ruído branco para a série temporal de energia distribuída da região metropolitana da Coelba com o limite inferior e o limite superior de 95%, as duas linhas tracejadas apresentadas na imagem a seguir.

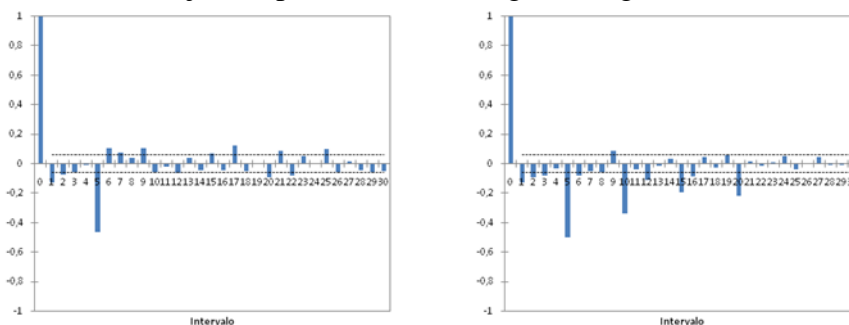


FIGURA 3: FAC e FACP

Seguindo a descrição da interpretabilidade da FAC e da FACP da seção 2. A FAC e a FACP indicam os parâmetros p , P , q e Q . Portanto, assumiu-se o modelo SARIMAX apresentado na tabela a seguir:

Tabela 2 – O modelo SARIMAX

p	d	q	P	D	Q	S
1	1	1	1	1	1	5

Os parâmetros e seus respectivos erros nos 4 testes aplicados são destacados na

tabela abaixo. Sendo os valores dos parâmetros o valor principal e seu erro entre parêntesis:

Tabela 3 – Parâmetros do modelo SARIMAX

Parâmetros	Teste 12	Teste 13	Teste 23	Teste 123
AR	0,054 (0,032)	0,561 (0,032)	0,491 (0,032)	0,491 (0,034)
SAR	0,022 (0,034)	0,082 (0,034)	0,081 (0,034)	0,081 (0,032)
MA	-0,480 (0,033)	-0,928 (0,033)	-0,937 (0,033)	-0,937 (0,014)
SMA	-0,999 (0,001)	-1,000 (0,001)	-1,000 (0,001)	-1,000 (0,000)
Constante	20.836,10 (-)	21.093,33 (-)	20.840,38 (-)	20.840,01 (-)
Clientes	5.598,59 (1,410)	3.278,408 (0,000)	-	5.697,09 (0,00)
Temperatura	282.055 (24,294)	-	287,472 (28,513)	287,533 (28,514)
Precipitação	-	28,057 (3,747)	-3,616 (5,822)	-3,631 (5,822)

5.3. MODELO DE PREVISÃO DO SISTEMA DE INFERÊNCIA FUZZY

Esta seção descreve o modelo de previsão do Sistema de Inferência Fuzzy. O modelo proposto é um Sistema de Inferência Neuro-Fuzzy Adaptativo (ANFIS). O Sistema proposto utiliza o método SUGENO com saída constante.

Utilizou-se 30 como a quantidade de épocas para o ajuste do modelo, que se ajusta através de um método híbrido. O método híbrido em questão é uma combinação do método dos mínimos quadrados e o método de gradiente descendente *backpropagation*.

Para o ajuste do ANFIS foram utilizados como input do modelo todos os 983 dados da série temporal. Assim como no modelo SARIMAX, esses mesmos 983 dados também foram utilizados para os testes. Pois no ANFIS é opcional utilizar a validação do modelo. Definiram-se quatro as funções de pertinência gaussianas para cada entrada do sistema.

Para fins de comparação, como foi informado no SARIMAX, foi definido como entrada do Sistema de Inferência Fuzzy a energia distribuída do dia anterior e a de cinco dias anteriores. Também foram utilizadas as mesmas combinações de variáveis exógenas para a modelagem dos testes do Sistema Fuzzy.

Tabela 4 – Parâmetros do Sistema Fuzzy

Funções de Pertinencia	Teste 12	Teste 13	Teste 23	Teste 123
Gbellmf	2,25%	2,40%	2,33%	2,07%
Gaussmf	2,18%	2,31%	2,23%	1,99%
Gauss2mf	2,79%	2,99%	2,86%	2,63%
Pimf	3,04%	3,32%	3,11%	2,94%
dsigmf	2,32%	2,71%	2,51%	2,32%

6. RESULTADOS E CONCLUSÃO

Cabe a esta seção apresentar os resultados e as diferenças entre os testes obtidos utilizando as diferentes ferramentas descritas na seção 5. Os resultados do modelo SARIMAX e do Sistema Fuzzy são apresentados na tabela abaixo com as diferentes combinações das variáveis exógenas.

Tabela 5 – Results

Testes	Clientes	Temperatura	Precipitação	SARIMAX MAPE	SARIMAX DP	ANFIS MAPE	ANFIS DP
Teste 12	X	X		3,20%	964	2,18%	705
Teste 13	X		X	2,69%	867	2,31%	762
Teste 23		X	X	3,07%	919	2,23%	727
Teste 123	X	X	X	3,08%	919	1,99%	662

O resultado do teste 13 indicado em negrito na tabela 5 destaca a melhor combinação de variáveis exógenas para o modelo SARIMAX. Assim como o resultado do teste 123 indicado em negrito destaca a melhor combinação de variáveis exógenas do Sistema de Inferência Fuzzy. Ou seja, as combinações que apresentaram o menor MAPE e o menor DP.

O teste 13 é feito a partir da combinação da variável exógena quantidade de clientes e de precipitação de chuvas. No modelo SARIMAX o MAPE do teste 13 é de 2,69% e o DP é de 867.

O teste 123 é feito a partir da combinação de todas as três variáveis exógenas: quantidade de clientes, temperatura ambiente e precipitação de chuvas. No Sistema Fuzzy o MAPE do teste 123 é de 1,99% e o DP é de 662.

A tabela acima destaca que a melhor combinação de variáveis para o modelo SARIMAX não foi a melhor combinação de variáveis para o Sistema Fuzzy. Ou seja, O Sistema de Inferência Fuzzy apresentou o menor erro no teste 123 e o SARIMAX no teste 13. Os resultados dos testes com o Sistema de Inferência Fuzzy apresentaram os melhores resultados em todas as combinações de variáveis exógenas.

Concluiu-se que, de uma forma geral, neste trabalho, com o uso da série temporal da região Metropolitana da Coelba nos dois modelos propostos, não há “melhores” variáveis exógenas para prever a quantidade diária de energia distribuída na região Metropolitana da Bahia. Pode-se afirmar que para o modelo SARIMAX aplicado a “melhor” combinação de variáveis exógenas foi a de quantidade de clientes e a precipitação de chuvas. E para o Sistema de Inferência Fuzzy aplicado a “melhor” combinação foi de todas as três variáveis exógenas.

É possível afirmar que no modelo SARIMAX, a inserção de mais “informação” ao sistema, ou seja, de mais uma variável de entrada no sistema não resultou em um menor erro na previsão. Diferentemente do Sistema de Inferência Fuzzy.

O Sistema de Inferência Fuzzy perdeu a sua interpretabilidade, pois gerou uma quantidade de regras muito grande. Sendo geradas 256 regras no teste 12, teste 13 e no teste 23 e no teste 123 gerado 1.024 regras. Para recuperar a interpretabilidade é necessário outro sistema para reduzir essa quantidade de regras.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AZADEH V. N. A., PAZHOUHESH FAR, P.; SABERIC, M. *An Adaptive-Network-Based Fuzzy Inference System for Long-Term Electric Consumption Forecasting (2008-2015): A Case Study of the Group of Seven (G7) Industrialized Nations: U.S.A, Canada, Germany, United Kingdom, Japan, France and Italy*. IEEE. 2010.
- [2] BOX, G. E. P.; JENKINS, G. M. *Time Series Analysis: Forecasting and Control*. San Francisco. Holden Day, 1970.
- [3] CAMPOS, R. J.; JESUS, T. A.; MENDES, E. M. A. M. *Uma abordagem fuzzy para a previsão de curto-prazo do consumo de energia elétrica*. In XXX Congresso Nacional de Matemática Aplicada - CNMAC, Brasil, Florianópolis, Setembro 2007.
- [4] CHAVEZ, G. S.; BERNAT, J. X.; COALLA, H. L. *Forecasting of energy production and consumption in Asturias (Northern Spain)*. Energy, 24, 183–98. 1999.
- [5] CHRISTO, E. S.; SOUZA, R. C. *Uma abordagem estatística para a previsão de potência reativa em sistemas elétricos*. Pesquisa Oper. Vol. 26. Número: 2. Rio de Janeiro. Maio-Agosto. 2006.
- [6] CONEJO, A. J.; CONTRERAS, J.; ESPÍNOLA, R.; PLAZAS, M. A. *Forecasting electricity prices for a day-ahead pool-based electric energy market*. International Journal of Forecasting, Vol. 21, pp. 435-65, 2005.
- [7] CUARESMA, J. C.; HLOUSKOVA, J.; KOSSMEIER, S.; OBERSTEINER, M. *Forecasting electricity spot prices using linear univariate time series models*. Applied Energy, Vol. 77, pp. 87-106, 2004.

- [8] ENDERS, W. *Applied Econometric Time Series*. 2 ed. United States of America: Ed. Wiley series in probability and statistics, 2004, 466 p. ISBN 0-471-23065-0.
- [9] ESPINOSA, M. M.; PRADO, S. M.; GHELLERE, M. *Uso do modelo SARIMA na previsão do número de focos de calor para os meses de junho a outubro no Estado de Mato Grosso*. Departamento de Estatística/ICET. Universidade Federal de Mato Grosso, Cuiaba, MT, 2010.
- [10] GUJARATI, D. N.; *Econometria básica*. 3 ed. São Paulo: Makron Books, 2000.
- [11] GOMIDE, F.; GUDWIN, R.; TANSCHKEIT, R. *Conceitos fundamentais da teoria de conjuntos fuzzy, lógica fuzzy e aplicações*. In: Proc. 6th IFSA Congress-Tutorials. 1995.
- [12] GOOIJER, de; JAN, G.; HYDMAN, R. J. *25 years of time series forecasting*. International Journal of Forecasting, 22(3):443-473. 2006.
- [13] GREENE, Willian H. *Econometric analysis*. 5. ed. New Jersey: Prentice Hall, 2003.
- [14] HERRERA, F. *Genetic fuzzy systems: taxonomy, current research trends and prospects*. Evolutionary Intelligence, v. 1, n. 1, p. 27-46, 1 mar. 2008.
- [15] JANG, J. S. R.; GULLEY, N. *The Fuzzy Logic Toolbox for use with MATLAB*. The MathWorks, Inc., Natick, Massachusetts, 1995.
- [16] LI, S.T.; CHENG, Y. *C.A Stochastic HMM-Based Forecasting Model for Fuzzy Time Series*. IEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetic, Vol. 40, No. 5. Outubro de 2010.
- [17] LIU, L; HUDAK, G. P. *Forecasting and Time Series Analysis Using the SCA Statistical System*. Scientific Computing Associated, 1994.
- [18] MATOS, O. C. *Econometria básica: teoria e aplicações*. São Paulo: Atlas, 2000.
- [19] MORETTIN, P. A. *Econometria Financeira: um curso de séries temporais financeiras*. ABE: São Paulo, 2008.
- [20] PANDIAN, S. C.; DURAISWAMY, K.; RAJAN, C. C. A.; KANAGARAJ N. *Fuzzy approach for short term load forecasting*. Electric Power Systems Research, 76: pp. 541-548, 2006.
- [21] PELLEGRINI, F.R. *Metodologia para implementação de sistemas de previsão de demanda*. Dissertação de Mestrado em Engenharia de Produção. Porto Alegre: UFRGS, 2000.
- [22] PINHEIRO, T. M. M. *Distribuidoras de Energia Elétrica no Brasil*. Dissertação de mestrado. Universidade de Brasília, 2012. Disponível em: http://www.aneel.gov.br/biblioteca/trabalhos/trabalhos/Dissertacao_Thelma_Pinheiro.pdf.
- [23] SAAB, S.; BADR, E.; NASR, G. *Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon*. Energy, 26, 1-14. 2001.
- [24] SERRÃO, F. C. C. *Modelo de previsão de carga de curto-prazo utilizando redes neurais e lógica fuzzy*. Dissertação de Mestrado. Pontifica Universidade Católica, Brasil, Rio de Janeiro, 2003.
- [25] SONG, Q.; CHISSOM, B. S. *Forecasting enrollments with fuzzy time series - Part I*, Fuzzy Sets Syst., vol. 54. pp. 1-9, 1993.
- [26] SONG, Q.; CHISSOM, B. S. *Fuzzy time series and its models*, Fuzzy Sets Syst., vol. 54. pp. 269-277, 1993.

- [27] SONG, Q.; CHISSOM, B. S. *Forecasting enrollments with fuzzy time series - Part II*, Fuzzy Sets Syst., vol. 62. pp. 1-8, 1994.
- [28] SOUZA, R. C.; CAMARGO, M. E. *Análise e previsão de séries temporais: os modelos ARIMA*. Sedigraf, 1996.
- [29] SUGENO, M. *Industrial applications of fuzzy control*. Elsevier Science Pub. Co., 1985.
- [30] TANSCHKEIT, R. *Sistemas fuzzy*, In: *Inteligência computacional aplicada a administração, economia e engenharia em Matlab*, pp. 229-264, São Paulo, Thomson Learning, 2007.
- [31] URI, N. D. *Forecasting peak system load using a combined time series and econometric model*. Applied Energy, Vol. 4, pp. 219-227, 1978.
- [32] ZADEH, L. A. *Fuzzy sets*. Fuzzy sets, Information and Control, 8:338-353, 1965.