

Previsão anual de demanda de energia elétrica

Cristiana da Silva Alexandre

Discente de graduação em Engenharia Elétrica – IFMA
cristiana.alexandre@acad.ifma.edu.br

Helton do Nascimento Alves

Professor Doutor em Engenharia Elétrica – IFMA
helton@ifma.edu.br

Resumo — Previsão de carga constitui uma atividade de grande importância, tendo em vista que o estudo do planejamento de sistemas de distribuição de energia elétrica somente poderá ser efetivado se houver a disponibilidade de uma boa estimativa de carga. Deste modo, visando contribuir para um seguro planejamento, foi desenvolvida uma ferramenta computacional utilizada na previsão de carga. Dados históricos da carga e temperatura média são requeridos da rede, assim como o índice médio pluviométrico e o produto interno bruto da região. O método foi desenvolvido no software MATLAB e baseia-se na regressão linear multivariável. O algoritmo foi testado e comparado computacionalmente com base nos dados reais, apresentando resultados significativos, confirmando a importância das variáveis no refinamento de métodos de previsão.

Palavras-chaves — planejamento de redes de distribuição, previsão de carga, regressão linear multivariável, variáveis exógenas

I. INTRODUÇÃO

O planejamento de sistemas de distribuição de energia elétrica consiste na obtenção de estudos para áreas com alternativas de expansão e suas obras programadas ao longo dos anos. Assim, a continuidade do processo de planejamento consiste em revisões periódicas e atualizações dos dados de carga. Pesquisar a evolução da carga consiste em obter a previsão de demanda da área para o período em estudo. A análise da evolução da carga e previsão do crescimento futuro geralmente aborda um período de até 10 anos, podendo ser dilatado em função de conveniências.

O estudo de previsão de carga proporciona um conjunto de informações vitais para que os sistemas de energia elétrica possam ser planejados de forma segura e confiável, fornecendo energia elétrica de qualidade e reduzindo possíveis investimentos desnecessários às concessionárias de energia.

A grande dificuldade de prever a distribuição da carga como ela efetivamente acontece está relacionada com a necessidade de dispor de dados em grande quantidade, elaboração de modelos complexos e de técnicas que possam processar tais dados e produzir resultados confiáveis e rápidos. São encontradas na literatura diferentes abordagens de previsão de carga a curto prazo, a médio prazo e a longo prazo, como: modelos de regressão [1], séries temporais [2] e inteligência artificial [3],[4].

O problema de previsão de carga para o planejamento da rede visa conhecer com antecedência padrões de carga, de modo a poder acompanhar o crescente consumo elétrico da forma mais eficiente possível. Prever com anos de antecedência o comportamento da carga pode levar a

diferentes decisões de traçados, necessidade de mudanças de equipamentos e prevenção de possíveis problemas na rede.

A carga de um sistema elétrico está sujeita a diversos fatores externos ao sistema e não lineares, de difícil modelamento, comumente chamadas de variáveis exógenas.

Neste trabalho é proposto um modelo de predição de cargas baseado na regressão linear múltipla que se ajusta após mudanças drásticas no comportamento dos dados. Tal proposta se faz levando em consideração que a carga é uma variável estatística de regressão, ou seja, uma combinação linear de variáveis com pesos empiricamente determinados, onde as variáveis escolhidas para o desenvolvimento do trabalho serão ano, temperatura, precipitação pluviométrica e PIB.

O modelo desenvolvido tem o objetivo de auxiliar no estudo do planejamento de redes em sistemas de distribuição de energia elétrica e minimizar as limitações do modelo de previsão de carga considerando a influência de variáveis exógenas ao problema sobre a predição da carga.

II. MATERIAL E MÉTODOS

A análise de regressão linear multivariável (RLM) é uma técnica da estatística multivariada utilizada para a predição de valores de uma ou mais variáveis dependentes (VDs) a partir de diversas variáveis independentes (VIs). A maximização da previsão é feita pela ponderação das VIs que indica a contribuição de cada uma delas para a previsão geral.

A regressão pode ser classificada de acordo com a quantidade de variáveis independentes. Quando existe apenas uma VI no modelo de regressão, a técnica estatística é denominada de regressão linear simples, mostrada em (1). Quando o modelo envolve mais de duas variáveis, a técnica é chamada de regressão linear multivariável, mostrada em (2).

$$Y_i = a_0 + a_1 x_i \quad (1)$$

Onde Y_i é o valor previsto da VD; a_0 é o intercepto; a_1 é o coeficiente de regressão; x_i representa o valor real de uma VI em relação ao i -ésimo valor real da VD.

$$Y_i = a_0 + a_1 x_{1i} + a_2 x_{2i} + \dots + a_n x_{ni} \quad (2)$$

Onde $a_1, a_2, \dots, a_n$ são os coeficientes de regressão; x_{ji} representa o valor real da j -ésima VI em relação ao i -ésimo valor real da VD; n é o número de VIs.

A análise regressiva consiste na determinação da curva que melhor se ajuste aos dados históricos de demanda, utilizando o método dos mínimos quadrados, segundo o qual uma curva

se ajusta a um conjunto de pontos dados quando a soma dos quadrados dos erros em cada ponto é mínima, isto é:

$$E^2 = \sum_{i=1}^m [Y_i - y_i]^2 \quad (3)$$

Onde y_i é o i -ésimo valor real da VD.

E^2 é um indicador de qualidade, quanto menor seu valor, mais fino é o ajuste, de modo que isso pode ser usado para se escolher entre duas curvas qual a que melhor se ajusta a uma sequência de dados.

Considerando que os dados históricos possuam m amostras, em (2) e (3) tornam-se:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_m \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{n1} \\ 1 & x_{12} & x_{22} & \cdots & x_{n2} \\ 1 & x_{13} & x_{23} & \cdots & x_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1m} & x_{2m} & \cdots & x_{nm} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \Rightarrow Y = XA \quad (4)$$

$$E^2 = (y - XA)^T (y - XA) \quad (5)$$

Onde y representa o vetor de tamanho m que contém os valores reais da VD.

Minimizando (5) obtém-se (6). Equação (6) é utilizada para calcular os coeficientes de regressão da RLM.

$$A = (X^T X)^{-1} X^T y \quad (6)$$

A seleção das variáveis é uma etapa de grande importância na construção de um modelo de RLM, já que todo modelo utilizado possui algum percentual de erro correlacionado, torna-se necessário a existência das melhores VIs que quando combinadas produzam a melhor previsão da VD.

A seleção das VIs possui uma problemática quanto ao erro de especificação, que diz respeito à inclusão de variáveis irrelevantes ou à omissão de variáveis relevantes do conjunto de VIs. A inclusão de variáveis irrelevantes pode causar dificuldades na interpretação dos resultados, podem diminuir os efeitos das variáveis relevantes e torna menos preciso o teste de significância estatística das VIs. Já a exclusão de variáveis relevantes pode afetar negativamente qualquer interpretação dos resultados e reduzir a capacidade preditiva geral do modelo.

As estratégias encontradas na literatura para a determinação de forma automatizada do melhor modelo para um determinado conjunto de dados são baseadas em mecanismos de busca exaustiva, algoritmos sequenciais ou metodologias heurísticas. Neste trabalho foi utilizado um método de busca sequencial.

Os métodos sequenciais ou Stepwise possuem a característica de avaliar os efeitos das variáveis através da

adição ou remoção de apenas uma variável em cada etapa, existindo três principais variações da metodologia.

A primeira variante é denominada de Seleção por Adição (Forward Selection) onde são adicionadas sequencialmente as variáveis candidatas ao modelo, a segunda é a Seleção por Eliminação (Backward Elimination), que por sua vez parte de um modelo com todas as variáveis e procede a eliminação sequencial das variáveis. A terceira modificação é uma mistura das duas técnicas anteriores. Sendo que a cada adição de variável no conjunto é realizado um procedimento para verificar se não há variáveis redundantes. Essa metodologia recebe a denominação de Stepwise Regression. Neste trabalho foi adotado o método de seleção por adição.

O procedimento de seleção por adição de variáveis parte de um modelo onde não há VIs. Criam-se então modelos com cada uma das VIs candidatas e a variável que apresentar melhor desempenho em relação ao um determinado índice de performance é adicionada ao conjunto de variáveis que irão compor o modelo.

Na segunda etapa de seleção são criados modelos contendo combinações da variável previamente incluída e cada uma das variáveis não selecionadas anteriormente. A variável que gerar o melhor modelo em combinação com aquela introduzida na etapa anterior é pré-selecionada para compor o modelo. A variável pré-selecionada é então submetida ao índice de performance para verificar se a sua adição acarreta melhora significativa ao modelo dentro de uma margem de confiança.

Em caso positivo, a variável é adicionada ao conjunto que irá fazer parte do modelo e repete-se o procedimento para as demais variáveis. Isso é repetido até que a adição de uma nova variável não traga melhoria significativa ao modelo ou se todas as variáveis tiverem sido incluídas.

Neste trabalho foi utilizado como índice de desempenho o coeficiente de correlação mostrado em (7). Considera-se que para $R \geq 0,95$ o ajuste da curva é considerado bom.

$$R^2 = 1 - \left(\frac{\sum_{i=1}^m [y_i - Y_i]^2}{\sum_{i=1}^m [y_i - \bar{y}_m]^2} \right) \quad (7)$$

Onde $\bar{y}_m$ é a média aritmética dos valores reais da VD.

Foi desenvolvido um modelo de predição de cargas elétricas baseado em regressão linear multivariável. Este estudo se faz necessário devido à necessidade de expansão da rede pelo aumento do consumo de carga. Além disso, uma análise através de métodos multivariados pode gerar previsões de cargas mais precisas e com maior impacto na análise da confiabilidade da rede, já que modelos exatos de previsão de carga são demasiadamente complexos.

Na definição da previsão de carga elétrica foram selecionadas quatro variáveis independentes: ano, temperatura média, precipitação pluviométrica e PIB.

Para avaliar a significância de cada uma dessas variáveis no modelo RLM foi utilizado o algoritmo apresentado na Fig. 1. O método é exemplificado na Tabela I. Verificou-se que as quatro VIs escolhidas possuem um poder preditivo significativo na obtenção do modelo de previsão. Depois do modelo concluído, das variáveis de regressão especificadas e da confirmação dos resultados obtidos torna-se necessário avaliar a equação preditiva (8) obtida.

$$Y_i = a_0 + a_1 T_1 + a_2 T_2 + a_3 T_3 + a_4 T_4 \quad (8)$$

Onde T_1 é o ano; T_2 é temperatura média no ano; T_3 é a precipitação pluviométrica máxima registrada no ano; T_4 é o produto interno bruto médio no ano.

Além de prever cargas futuras de uma determinada região, os coeficientes de regressão também apresentam um meio de avaliar a importância de cada variável independente na previsão geral da carga e desta forma tomar medidas de gerenciamento de manutenção adequada para minimizar a contribuição de cada variável na interrupção de fornecimento de energia elétrica.

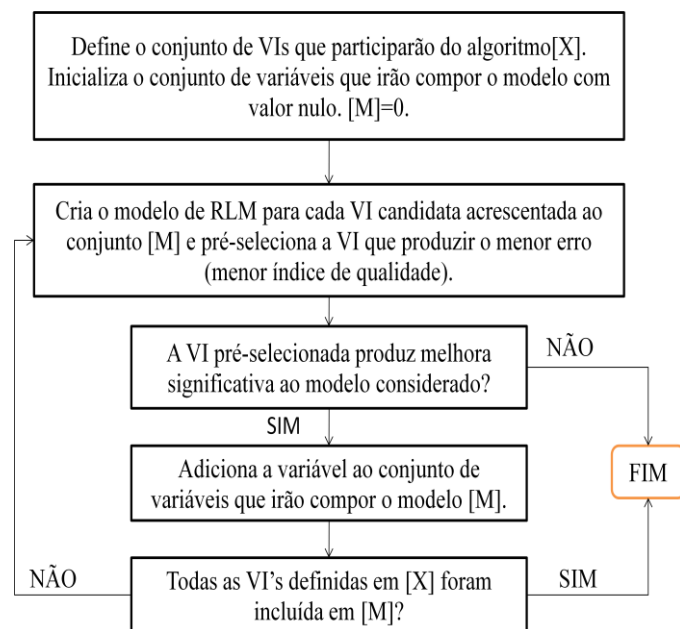


Fig. 1. Fluxograma da seleção por adição

TABELA I. MÉTODO DE ANÁLISE DA SIGNIFICÂNCIA DAS VARIÁVEIS

Número de VIs		T_1 Ano	T_2 Temperatura máxima	T_3 Precipitação pluviométrica	T_4 PIB
1	E^2	0.0003	3.5747	3.5749	3.5743
	R^2	0.9760			
2	E^2		2.5770	2.6649	2.663
	R^2		0.9760		
3	E^2			0.6405	2.5770
	R^2			0.9941	
4	E^2				3.2235
	R^2				0.9993

III. RESULTADOS E DISCUSSÕES

O modelo de previsão de carga baseado em regressão multivariável foi aplicado aos dados de carga encontrados no banco de dados da Empresa de Pesquisa Energética (EPE). A análise através de métodos multivariados pode gerar previsões de cargas mais precisas e com maior impacto na análise da confiabilidade da rede, já que modelos exatos de previsão de carga são demasiadamente complexos. A Tabela II apresenta o consumo de energia das regiões brasileiras. Os dados originais são relativos a 3 anos (2012 a 2014). As colunas 5 e 6 da Tabela II mostram os valores previstos para a carga obtida pela RLM para o ano de 2015 e 2016. A partir

da análise desse conjunto de dados foram determinados os parâmetros dos modelos de regressão multivariável para a equação preditiva de carga. Verificou-se então, a previsão para os 2 anos posteriores (2015 e 2016). No entanto, foi necessária a predição de valores futuros para as variáveis independentes. Dessa forma, utilizou-se uma regressão linear simples. Depois de previstos os valores da variável dependente do modelo, foram calculados o coeficiente de correlação por região. O coeficiente de correlação obtido por região foi: 0.9993 (Norte); 0.9996 (Nordeste); 0.9989 (Sudeste); 0.9991 (Sul); 0.9987 (Centro-Oeste). Para validação do modelo foi realizada uma comparação entre os valores de carga previstos e os valores reais por região. Os índices avaliados foram o erro percentual e coeficiente de correlação. Por região, o erro percentual foi aproximadamente abaixo de 5% e o coeficiente de correlação, superior a 0.99.

TABELA II. CONSUMO DE ENERGIA DAS REGIÕES BRASILEIRAS NO PERÍODO DE 5 ANOS

	2012	2013	2014	2015	2016
Norte	29.048.752	30.195.682	32.363.651	33.575.317	34.731.466
Nordeste	75.610.402	79.693.889	80.746.115	83.219.381	87.465.127
Sudeste	235.237.240	240.104.803	243.025.384	223.295.389	231.931.425
Sul	77.490.628	80.392.671	84.818.504	84.971.725	90.101.333
Centro-Oeste	30.717.728	32.755.449	34.380.929	35.423.159	38.958.689

TABELA III. VALORES REAIS E VALORES PREVISTOS

	Valor Real em 2015	Valor Previsto em 2015	Valor Real em 2016	Valor Previsto em 2016
Norte	33.420.014	33.575.317	34.256.889	34.731.466
Nordeste	80.040.979	83.219.381	82.986.134	87.465.127
Sudeste	234.607.807	223.295.389	242.131.907	231.931.425
Sul	82.004.951	84.971.725	87.120.196	90.101.333
Centro-Oeste	34.648.191	35.423.159	36.169.567	38.958.689

IV. CONCLUSÃO

Em termos de previsão de carga o modelo proposto atingiu um nível elevado de precisão e os resultados obtidos o validaram. Todas as variáveis escolhidas para o desenvolvimento do modelo demonstraram-se relevantes na previsão da variável carga. Outras variáveis climáticas, como incidência solar e velocidade do vento, outras variáveis sazonais, como meses e estações do ano e outros efeitos aleatórios, como aumento ou redução da demanda em função de aquecimento no mercado, eventos esportivos, podem ser utilizadas com variáveis independentes no modelo apresentado. O grande problema nesse tipo de estudo é a escassez de informações que limita o desenvolvimento de um modelo mais adequado. Como trabalho futuro, pretende-se utilizar esse modelo de previsão de carga para a análise de confiabilidade em um planejamento de redes de distribuição de energia elétrica.

V. REFERÊNCIAS

- [1] J. R. Cancelo, “Modelling and forecasting daily series of electricity demand”, *Invest. Econ.*, vol. 20, pp. 359 – 376, 1996.
- [2] R. J. Campos, “Previsão de séries temporais com aplicações a série de consume de energia elétrica. 110 f. Mestrado em Engenharia Elétrica. UFMG, Minas Gerais, 2008.
- [3] J. F. Nunes, “Previsão de carga multimodal usando a rede neural ART-ARTMAP Fuzzy. 110 f. Doutorado em Engenharia Elétrica. Universidade Estadual Paulista, 2013
- [4] C. R. Guirelli, “Previsão da carga de curto prazo de áreas elétricas através de técnicas de inteligências artificial”. 127 f. Doutorado em Engenharia Elétrica, USP, São Paulo, 2006.