

Previsão de Carga no estudo de Planejamento de Redes de Distribuição de Energia Elétrica

Helton do Nascimento Alves^[1]

Cristiana da Silva Alexandre^[2]

Kelly dos Santos Silva^[3]

^[1] Departamento de Eletroeletrônica (DEE), ^[2] ^[3] Engenharia Elétrica Industrial
Grupo de Pesquisa em Sistemas Elétricos (GPSE)

São Luís, Brasil

helton@ifma.edu.br; cristianaalexandre@hotmail.com; santos.kelly@outlook.com

Resumo— A 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 que o planejamento dos sistemas de distribuição de energia elétrica ocorra de forma segura, foi desenvolvida uma ferramenta computacional utilizada na previsão. Essa ferramenta utiliza a regressão linear multivariável. Dados históricos da carga e informação climáticas como a temperatura máxima e mínima são requeridos da rede, assim como o índice médio pluviométrico.

Palavras-chave— Sistema de Distribuição, Previsão de Carga, Regressão Linear Multivariável.

1. 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 relacionado 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 [1], a médio prazo [2] e a longo prazo [3], como: modelos de regressão [4]; séries temporais [5] e inteligência artificial [6].

Neste trabalho é proposto um modelo de predição de cargas baseado na regressão linear múltipla que se ajusta após mudanças drásticas no comportamento de alguns dados, como temperatura e precipitação pluviométrica. O modelo desenvolvido tem o objetivo de auxiliar no estudo do planejamento de redes em sistemas de distribuição de energia elétrica.

2. FORMULAÇÃO DO PROBLEMA

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 tais como precipitação e temperatura, também chamadas de variáveis exógenas.

A temperatura é a variável climática mais importante que afeta o consumo de energia. Uma mudança de temperatura influencia as condições de conforto térmico de uma determinada região, produzindo mudanças de comportamento no consumidor final e, consequentemente, mudanças no consumo de energia [7]. Neste trabalho é proposta uma abordagem da previsão de carga como uma variável estatística de regressão, ou seja, uma combinação linear de variáveis com pesos empiricamente determinados, onde as variáveis serão ano, temperatura máxima, temperatura mínima e precipitação pluviométrica. Esta abordagem tem como objetivo minimizar as limitações do modelo de

previsão de carga considerando a influência de variáveis exógenas ao problema sobre a predição da carga.

2.1. Análises de Regressão Linear Multivariável

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 na equação (1). Quando o modelo envolve mais de duas variáveis, a técnica é chamada de regressão linear multivariável, mostrada na equação (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.

A regressão múltipla:

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

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 processo 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, a equação (2) e (3) tornam-se:

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

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

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

Minimizando a equação (5) obtém-se que:

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

A equação (6) é utilizada para calcular os coeficientes de regressão da RLM.

2.2. Seleção de variáveis

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.

2.2.1. 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. Cria-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 melhora significativa ao modelo ou se todas as variáveis tiverem sido incluídas.

Neste trabalho foi utilizado como índice de desempenho o coeficiente de correlação [8] mostrado na equação (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.

O algoritmo de seleção por adição é mostrado na Figura 1.

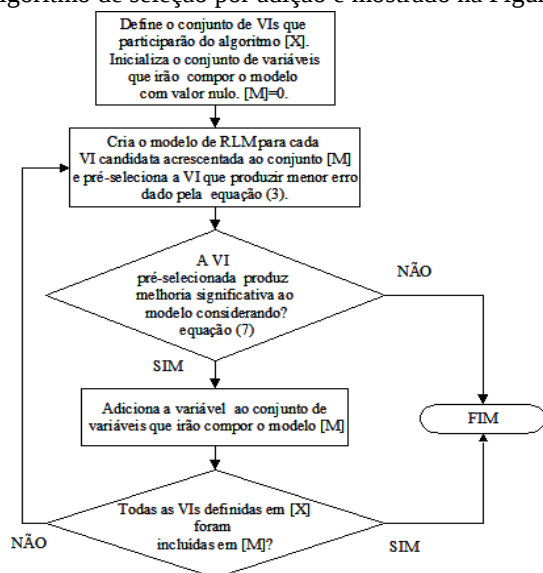


Figura 1. Fluxograma da seleção por adição.

3. ALGORITMO DESENVOLVIDO

Neste trabalho é desenvolvido um modelo de predição de cargas elétricas baseado em regressão linear multivariável. Este estudo se faz necessário devido a 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.

Tal modelo foi aplicado aos dados de carga encontrados no banco de dados da Empresa de Pesquisa Energética (EPE) e os dados de temperatura máxima, temperatura mínima e precipitação encontrados no Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP).

3.1. Análise das variáveis

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

Para avaliar a significância de cada uma dessas variáveis no modelo RLM foi utilizado o algoritmo apresentado na Figura 1. O método é exemplificado na tabela 1. Verificou-se que as quatro VIs escolhidas possuem um poder preditivo significativo na obtenção do modelo de previsão.

Tabela 1. Método de análise da significância das variáveis referente à região Norte

Número de VIs		T_1	T_2	T_3	T_4
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				0.9973

3.2. Interpretação da Variável Estatística de Regressão

Depois do modelo concluído, das variáveis de regressão especificadas e da confirmação dos resultados obtidos torna-se necessário avaliar a equação preditiva obtida.

$$Y_i = a_0 + a_1T_1 + a_2T_2 + a_3T_3 + a_4T_4 \quad (10)$$

Onde:

T_1 é o ano;

T_2 é a máxima temperatura registrada no ano;

T_3 é a mínima temperatura registrada no ano;

T_4 é a precipitação pluviométrica máxima registrada 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.

4. RESULTADOS OBTIDOS

A tabela 2 apresenta o consumo de energia das regiões brasileiras. Os dados originais são relativos a 4 anos (2011 a 2014). As colunas 6 e 7 da tabela 2 mostram os valores previstos para a carga obtida pela RLM para o ano de 2015 e 2016. A tabela 3, 4 e 5 referem-se respectivamente a temperatura máxima, temperatura mínima e a precipitação máxima anual por região de 4 anos. 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. No entanto, foi necessária a predição de valores futuros para as variáveis independentes. Dessa forma, utilizou-se uma regressão linear simples e os resultados encontram-se nas linhas 6 e 7 das tabelas 3, 4 e 5. Depois de previstos os valores da variável dependente do modelo, foram calculados o coeficiente de correlação por região. O indicador de qualidade assim como o coeficiente de correlação foram mostrados por regiões nas tabelas 1, 6, 7, 8 e 9.

5. 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, podem ser utilizadas com variáveis independentes no modelo apresentado.

Esse trabalho focou-se na previsão a longo prazo para os anos de 2015 e 2016, no entanto, pode ser aplicado a uma análise a curto e a médio prazo, desde que seja levado em consideração os dados históricos de temperatura (máxima e mínima) e precipitação daquele período analisado, o que

acarretaria uma mudança no erro correlacionado do modelo em questão.

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.

AGRADECIMENTO

Agradecimento ao Instituto Federal do Maranhão (IFMA), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo a Pesquisa do estado do Maranhão (FAPEMA) pelo apoio ao desenvolvimento do projeto.

REFERÊNCIAS

- [1] J. F. Antunes, "Previsão de carga multinodal a curto prazo usando uma rede neural ART-ARTMAP-FUZZY", *Brazilian Conference on Intelligent Systems BRACIS 2012*, v. 1., pp. 1-12.
- [2] A. F. Guimarães, "Previsão da evolução da carga no médio prazo em redes de distribuição", Tese de mestrado, Dept. de Engenharia Eletrotécnica e de Computadores, Universidade Técnica de Lisboa, 2008.
- [3] R. S. Dias, "Metodologia de previsão de carga de longo prazo de energia elétrica", Monografia, Dept. de Engenharia Elétrica, Universidade Federal do Ceará, 2010.
- [4] S. Caldas e H. Henri, "Previsão do consumo de energia elétrica na cidade de Salvador – BA utilizando regressão linear", *XVI Congresso Brasileiro de Energia*, 2015.
- [5] R. J. Campos, "Previsão de séries temporais com aplicações a séries de consumo de energia elétrica", Tese de Mestrado, Dept. de Engenharia Elétrica, Universidade Federal de Minas Gerais, 2008.
- [6] C. R. Guirelli "Previsão da carga de curto prazo de áreas elétricas através de técnicas de inteligência artificial", Tese de Doutorado, Dept. de Engenharia de Energia e Automação Elétricas, Escola Politécnica, Universidade de São Paulo, 2006.
- [7] K.Y. Lee e S. D. Du, "Short-term load forecasting using semigroup based system type Neural Network. *IEEE on Intelligent System Applications to Power Systems*, n.978-1- 4244-5098-5, p.1 – 6.
- [8] D. C. Montgomery and E. A. Peck. *Introduction to linear regression analysis*. John Wiley and Sons, New York, USA, 1982.

Tabela 2. Consumo de energia das regiões brasileiras no período de 4 anos.

	2011	2012	2013	2014	2015	2016
Norte	27.777.104	29.048.752	30.195.682	32.363.651	33.670.467	35.200.134
Nordeste	71.913.805	75.610.402	79.693.889	80.746.115	84.607.830	87.654.541
Sudeste	230.668.383	235.237.240	240.104.803	243.025.384	247.529.183	251.637.276
Sul	74.469.787	77.490.628	80.392.671	84.818.504	87.739.727	91.118.459
Centro-Oeste	28.204.522	30.717.728	32.755.449	34.380.929	36.867.514	39.008.657

Tabela 3. Temperatura máxima por região (°C).

	Norte	Nordeste	Sudeste	Sul	Centro-Oeste
2011	33.593548	31.453571	31.171429	27.774194	30.01
2012	34.406667	31.741935	30.806452	28.717241	30.23871
2013	34.123333	32.780645	29.989286	26.906452	29.193333
2014	34.283333	30.009677	32.880645	29.200000	30.190000
2015	34.548225	30.673214	32.289573	28.766129	29.781666
2016	34.726827	30.343916	32.720621	29.012791	29.737128

Tabela 4. Temperatura mínima por região (°C)

	Norte	Nordeste	Sudeste	Sul	Centro-Oeste
2011	18.396774	19.090323	11.403333	7.9266670	13.751613
2012	17.680645	19.722581	12.941935	9.1419350	14.141935
2013	17.506452	19.954839	12.006452	8.5129030	14.176667
2014	18.438710	20.32500	13.200000	10.274194	13.996774
2015	17.993548	20.757258	13.501559	10.567311	14.209300
2016	17.988710	21.150886	13.947011	11.208666	14.286322

Tabela 5. Precipitação (mm)

	Norte	Nordeste	Sudeste	Sul	Centro-Oeste
2011	289.70	331.0	285.40	338.80	333.80
2012	452.40	411.50	241.20	254.60	444.60
2013	411.20	308.40	258.10	308.90	491.80
2014	512.20	247.20	195.20	242.50	353.60
2015	572.95	235.90	181.55	227.55	432.60
2016	635.58	200.45	156.18	204.09	443.26

Tabela 6. Método de análise da significância das variáveis referente à região Nordeste

Número de VIs		T_1	T_2	T_3	T_4
1	E ²	0.0002	2.3760	2.3760	2.3760
	R ²	0.9501			
2	E ²		0.3306	2.3324	2.3324
	R ²		0.9920		
3	E ²			0.3435	3.3062
	R ²			0.9993	
4	E ²				3.1983
	R ²				0.9993

Tabela 7. Método de análise da significância das variáveis referente à região Sudeste

Número de VIs		T_1	T_2	T_3	T_4
1	E^2	0.0004	2.2526	2.2526	2.2526
	R^2	0.9894			
2	E^2		1.9521	9.3136	9.3136
	R^2		0.9976		
3	E^2			1.5442	1.9521
	R^2			0.9980	
4	E^2				1.4945
	R^2				0.9982

Tabela 8. Método de análise da significância das variáveis referente à região Sul

Número de VIs		T_1	T_2	T_3	T_4
1	E^2	0.0001	2.5208	2.5201	2.5205
	R^2	0.9891			
2	E^2		2.8243	6.2841	6.2845
	R^2		0.9954		
3	E^2			2.5838	2.8243
	R^2			0.9956	
4	E^2				2.3931
	R^2				0.9958

Tabela 9. Método de análise da significância das variáveis referente à região Centro-Oeste

Número de VIs		T_1	T_2	T_3	T_4
1	E^2	0.0004	3.9940	3.9941	3.9943
	R^2	0.9907			
2	E^2		1.6118	1.9733	1.9721
	R^2		0.9925		
3	E^2			0.6602	1.6118
	R^2			0.9971	
4	E^2				6.5396
	R^2				0.9972