

## **PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA EM PETROLINA - PE**

Alexandre Alex Silva e Caldas<sup>1</sup>

Eucymara França Nunes Santos<sup>2</sup>

### **RESUMO**

Este artigo científico apresenta equações matemáticas como forma de previsão do consumo de energia elétrica utilizando o método estatístico de regressão linear múltipla. As equações que são apresentadas utilizam o modelo do método dos mínimos quadrados para determinação dos coeficientes de regressão, e respeitando a divisão de classes consumidoras industrial e residencial na cidade de Petrolina (PE). Este artigo possui o objetivo de prover as empresas e as pessoas que atuam nas áreas de geração, transmissão e distribuição de energia elétrica na região com ferramenta de previsão que garanta um patamar confiável de projeções do consumo de energia elétrica, garantindo o fornecimento necessário de energia. Visando que situações como as ocorridas no Brasil em meados de 2000 e 2001 de racionamento de energia elétrica não se repitam. Então o artigo apresenta uma relevância significativa, visto a ausência de trabalhos nessa área voltados aos interesses da região. Ao final consegue-se alcançar métodos confiáveis, a partir das equações de regressão linear múltipla, que garantem uma previsão do consumo de energia elétrica correta.

Palavras-chave: Consumo, previsão, energia elétrica, regressão linear múltipla.

---

1 Graduando em Engenharia Elétrica, Universidade Federal do Vale do São Francisco. E-mail: eng.alexandrecaldas@gmail.com Avenida Antônio Carlos Magalhães, 510 – Santo Antônio. CEP: 48.902-300 Juazeiro, BA.

2 Professora do Colegiado de Engenharia Elétrica, Universidade Federal do Vale do São Francisco. E-mail: eucymara.nunes@univasf.edu.br Avenida Antônio Carlos Magalhães, 510 – Santo Antônio. CEP: 48.902-300 Juazeiro, BA



## ABSTRACT

This article presents mathematical equations as a way to forecast electricity consumption using the statistical method of multiple linear regression. The equations are presented using the model of least squares method for determining the regression coefficients, and respecting the division of industrial and residential customer classes in the city of Petrolina (PE). This article has the objective of providing companies and individuals who work in the areas of generation, transmission and distribution of electricity in the region a forecasting tool that ensures a reliable level projections of energy consumption, ensuring the necessary supply of energy, aiming to avoid situations such as occurred in Brazil in mid 2000 and 2001 from energy rationing. Then the paper has a significant relevance because there are not many works in this area of knowledge that address the interests of the region. At the end we present reliable methods derived from multiple linear regression equations, which guarantee a correct forecast of electricity consumption.

Keywords: Consumption, forecast, electricity and multiple linear regression

## 1. INTRODUÇÃO

A eletricidade assume condição *sine qua non* e ubíquo a toda e qualquer atividade cotidiana moderna, seja nas mais complexas de risco de vida em hospitais, seja em grandes atividades industriais, comerciais ou mesmo em usos mais supérfluos. Passando desde a fase de irrigação para a plantação do alimento, quanto ao seu transporte, industrialização, venda e consumo final pela população (FARIAS & SOUZA 2009).

Cid Tomanik Pompeu (1997) conceitua que:

“A Energia do sistema é seu potencial de trabalho. A energia elétrica é uma das formas de energia dentro de um sistema; ela corresponde ao produto de uma diferença de potencial (volt's) por uma corrente elétrica (Âmpere) pelo tempo (segundo) em que é fornecida.”

O Vale do São Francisco, mais especificamente a cidade de Petrolina (PE), constitui um pólo em grande expansão tendo como principal foco da economia, a agricultura de exportação, mas que desde o início do século



XXI volta sua economia também para a industrialização desses produtos agrícolas e atração de indústrias diversas nessa nova visão de economia cosmopolita. Possuindo ainda grandes taxas de aumento populacional devido a bons atrativos como educação, saúde e emprego. Para a garantia desse crescimento exacerbado da cidade é necessário corroborar uma garantia futura de oferta energética que supra essa necessidade emergente. Então esse estudo visa à previsão da demanda de consumo de energia elétrica na cidade colaborando com órgãos de geração e distribuição visando à garantia de oferta de energia elétrica num futuro próximo.

Esse tipo de previsão é muito importante, principalmente para concessionárias de distribuição de energia, empresas que trabalham com a geração e os órgãos estatais que trabalham com o gerenciamento da energia elétrica no país, pois prevê o consumo em função da quantidade de consumidores, temperatura e tempo, Fogliatto (2005). Assim essas empresas poderão fazer análises para alocar projetos de expansão ou melhoramento em áreas com previsão de aumento acelerado. Para que acontecimentos como o racionamento de energia, ocorridos entre 2000 e 2001 e a interrupção abrupta do fornecimento, os apagões, não voltem a acontecer, caso semelhante que a Venezuela viveu em 2009, também pela ausência de projetos de aumento da demanda por energia.

Com isso, o presente artigo possui o objetivo de prever o consumo de energia elétrica na região para um futuro próximo, a fim de que concessionárias de Energia Elétrica possam se preparar para os desafios de crescimento dessa região, colaborando assim com um crescimento planejado da região do Vale do São Francisco, mais especificamente a cidade de Petrolina (PE).

## **2. MATERIAIS E MÉTODOS**

O objetivo do modelo de regressão é estimar parâmetros desconhecidos a esse modelo de previsão utilizando regressão estatística. O modelo utilizado neste trabalho foi o de regressão linear múltipla, onde se estimam os coeficientes das variáveis independentes pelo método dos mínimos quadrados, (COLEHO-BARROS 2008 e BARBETTA 2004).

Foram utilizados neste trabalho, dados mensais de consumo de energia elétrica e número de consumidores, fornecidos pela Companhia Energética de Pernambuco (CELPE). Dados de temperatura média, forneci-



dos pela Embrapa Semi-Árido, Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), que possui uma estação agro-meteorológica no distrito de Bebedouro, em Petrolina.

Com base na série histórica, totalizando 144 dados de cada item, compreendidos entre os meses de janeiro de 1998 à dezembro de 2009. Foram utilizadas as variáveis que mais atenderam as expectativas quanto à regressão em termos de “Correlação”, o valor de “R<sup>2</sup>” e o teste “F” da regressão múltipla.

Utilizando como meio de previsão a regressão linear múltipla, quando utilizado mais de 30 observações já se admite que a amostra possui uma distribuição normal, Devore (2006). Grandes valores de dados ou experimentos geram gráficos com grande dispersão que implica numa boa possibilidade de se observar os valores da variável dependente, mas como meio de comprovação foi feito a plotagem dos dados previstos e dados reais e ambos ficaram bem dispersos o que indica uma normalidade na amostra dos dados.

## 2.1. Regressão Linear Múltipla

O modelo estatístico utilizado consiste em um modelo com mais de um regressor. De modo que se representa da seguinte forma, (HOLLANDER & WOLFE 1999):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \varepsilon_i \quad (1)$$

É chamado de modelo de regressão linear múltipla onde:

$\beta_j$  ( $j= 1,2,\dots,k$ ) são parâmetros dos coeficientes de regressão;

$X_1, X_2, \dots, X_k$  são os regressores;

“ $\varepsilon$ ” um erro aleatório provindo das medições e análise dos dados e;

$Y_i$  o  $i$ -ésimo valor observado da variável dependente.

O modelo de regressão linear múltipla descreve um hiper plano de  $k$ -dimensões, com  $k$  sendo a quantidade de variáveis dependentes. O parâmetro  $\beta_j$  descreve a variação esperada no valor de “ $Y_i$ ” por variação de  $X_j$ , quando os outros regressores são mantidos constantes,  $X_i$  com  $i \neq j$ , (MONTGOMERY et al. 2003).



## 2.2. Estimação pelo Método dos Mínimos Quadrados

É um método utilizado para estimar os coeficientes da regressão múltipla. Com essas observações satisfazendo o modelo, (MINGOTI 2005 e MEYER 1983):

$$Y - \hat{y} = \varepsilon$$

$$Y - (a + bX) = \varepsilon$$

$$\varepsilon = Y - (a + bX)$$

$$\varepsilon^2 = [Y - (a + bX)]^2$$

$$\sum \varepsilon^2 = \sum [Y - (a + bX)]^2$$

$$Z = \sum [Y - (a + bX)]^2$$

$$\frac{\partial Z}{\partial a} = 2 \sum [Y - (a + bX)]$$

$$\frac{\partial Z}{\partial b} = 2 \sum [Y - (a + bX)]X$$

$$\sum [Y - (a + bX)] = 0$$

$$\sum [Y - (a + bX)]X = 0$$

$$\sum Y = \sum a + b \sum X$$

$$\sum XY = \sum ab = b \sum X^2$$

$$\sum Y = na + bX$$

$$\sum XY = a \sum X + b \sum X^2$$

Formando as equações normais do modelo linear múltiplo com duas variáveis independentes; Temos:

$$\sum Y = na + b_1 X_1 + b_2 X_2$$

$$\sum X_1 Y = a \sum X_1 + b_2 \sum X_1 X_2 + b_1 \sum X_1^2$$

$$\sum X_2 Y = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2$$

Atualmente existem softwares que resolvem essas contas com o auxílio do computador. No caso dos nossos dados utilizaremos o Excel.



## 2.3. Testes de Hipóteses para a Regressão Linear Múltipla

### 2.3.1. Teste para a Significância da Regressão

Determina a existência de uma relação linear entre a variável resposta “Y” e o conjunto de regressores  $X_1, X_2, \dots, X_k$ , as hipóteses são:

$$H_0 : \beta_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0 \quad (2)$$

$$H_a : \beta_j \neq 0 \text{ para no mínimo um regressor } j.$$

A rejeição de  $H_0$  implica que pelo menos um regressor contribui significativamente para o modelo de regressão.

Tabela 1 - Análise da variância de testes da regressão linear múltipla

Fonte de Variação	Graus de Liberdade	Soma Quadrática	Média Quadrática	F0
Regressão	k	SQr	MQr	$\frac{MQr}{MQe}$
Erro ou Resíduo	n - p	SQe	MQe	
Total	n - 1	SQt		

Rejeita-se  $H_0$  se o valor calculado  $F_0$  for maior que  $F_{\alpha, p-1, n-p}$ , valor tabelado com “p-1” e “n-p” graus de liberdade, embora isso não implica necessariamente que a relação encontrada seja um modelo apropriado para previsão. Apenas informa que algum dos regressores se relaciona linearmente com a variável resposta e assim contribui significativamente com a regressão.

### 2.3.2. Teste F

Compara-se populações com distribuição normal, variáveis dependentes com as independentes, havendo variáveis que possuem variâncias parecidas ou iguais, o que pode empobrecer a regressão, verifica-se a possibilidade de sua exclusão, (MAGALHÃES & LIMA 2004).

As hipóteses são:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2 \quad (3)$$

## Intervalos de Confiança para os Coeficientes Individuais da Regressão

O procedimento para obtenção dos intervalos de confiança requer que os erros sejam normais e independentemente distribuídos com média zero e variância  $\sigma^2$ , com média  $\beta_0 + \sum_{j=1}^k \beta_j X_j$ , MARTINS (2005). Um intervalo de confiança de  $100(1 - \alpha)\%$  para coeficiente e regressão  $\beta_j$ , com  $j = 0, 1, 2, \dots, k$  no modelo de regressão múltipla é dado por:

$$\beta_j - t_{\alpha/2, n-p} \sqrt{\sigma^2 C_{jj}} \leq \beta_j \leq \beta_j + t_{\alpha/2, n-p} \sqrt{\sigma^2 C_{jj}} \quad (4)$$

### 2.4. Análise Residual

Útil no julgamento da adequação do modelo de previsão utilizando regressão linear múltipla. Utiliza-se do gráfico dos resíduos dos  $X_k$  contra  $Y$  para verificação da aleatoriedade desses resíduos, que demonstra o quão melhor será a previsão na utilização desse método (FARIAS et al. 2003).

## 3. RESULTADOS E DISCUSSÕES

### 3.1. Gráficos com os Dados Fornecidos

Dados do consumo industrial em Petrolina, como mostra a figura 1, que oscila muito por causa dos meses de safra e entressafra da agroindústria.

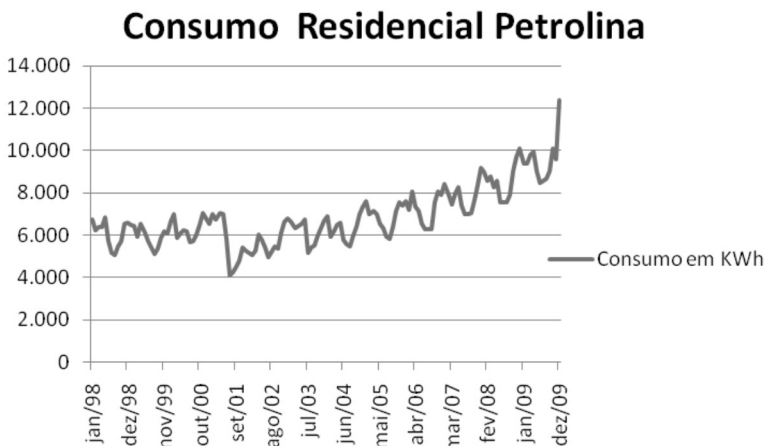


Figura 1: Consumo industrial em Petrolina, dados em KWh.



Dados do Consumo residencial em Petrolina, na figura 2, apresenta decréscimo entre os anos de 2000 e 2001, causado pelo racionamento de energia elétrica. Apresenta certa sazonalidade devido às variações de temperaturas ocorridas na região.

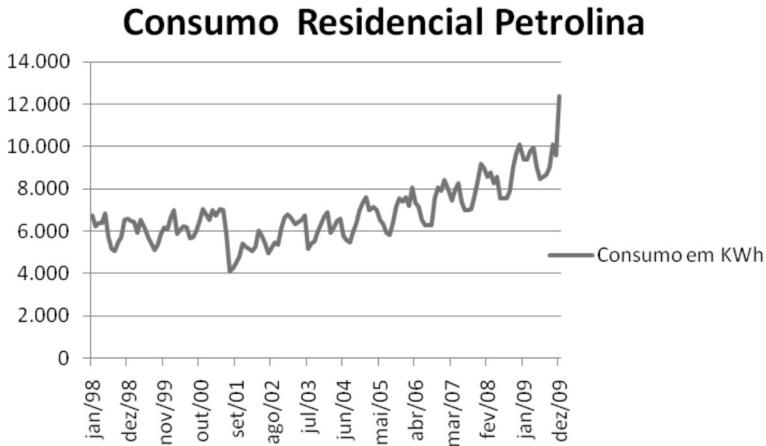


Figura 2: Consumo Residencial em Petrolina, dados em KWh

### 3.2. Regressões e Análises

Os coeficientes das equações de previsão do consumo de energia elétrica foram encontrados respeitando a metodologia clássica do método dos mínimos quadrados, sendo feito, ao final, uma análise de plotagem de dados previstos e observados mostrando um bom ajuste do modelo aos dados mostrando-se dispersos dos dados reais previstos, Hogg & Craig (2004). Sendo utilizadas variáveis de quantidade de consumidores, temperatura e tempo, que foram escolhidas e arranjadas dessa forma por apresentarem bons valores de correlação, como pode ser visto na tabela 2 e na tabela 5, BUNCHAFT & KELLNER (1998).

#### 3.2.1. Previsão para Consumo Industrial de Petrolina

A correlação é o primeiro teste que fazemos para análise de boa previsão da regressão, como mostrado na tabela 2, a variável dependente apresenta baixa correlação com as variáveis independentes.



Tabela 2 - Correlação entre as variáveis da Previsão de consumo industrial

Correlação	Consumo	Consumidores	Tempo
Consumo	1		
Consumidores	0,508177997	1	
Tempo	0,702207212	0,794660842	1

O teste F de significação testa na equação a possibilidade de que um ou mais variáveis ou estimadores tenham significado, para isso o F de significação calculada precisa ser maior que o F de significação tabelado, o que nas regressões apresentadas constituem bons valores validando as regressões, como mostra a tabela 3.

Tabela 3 - MANOVA de previsão de consumo industrial

	gl	SQ	MQ	F	F de significação
Regressão	2	845471397,3	422735698,7	3283,758276	2,3423E-118
Resíduo	141	18151681,25	128735,328		
Total	143	863623078,6			

Os coeficientes da equação são apresentados na tabela 4, bem como o valor do  $R^2$ .

Tabela 4 - Coeficiente das variáveis da equação de previsão do consumo industrial

<b>Variável Y= Consumo Industrial em Petrolina</b>	
Nº de Consumidores ( $x_1$ )	7,158
Tempo ( $x_2$ )	4,431
$R^2$	0,979

Os intervalos de confiança apresentam níveis máximos e mínimos dos coeficientes das equações de previsão, no caso com variação de até 5%, Holanda (2000) e Dias et. al. (2009), que configuram bons valores estatísticos de previsão nessa análise.

Intervalos de Confiança para 95%:

Número de Consumidores Industriais em Petrolina:  $6,686 \leq X_1 \leq 7,630$

Tempo:  $2,786 \leq X_2 \leq 6,075$



### 3.2.2. Previsão de Consumo Residencial em Petrolina

A tabela 5 apresenta a baixa correlação da variável dependente com as variáveis independentes, comprovando uma boa análise da confiabilidade da regressão.

Tabela 5 - Correlação entre as Variáveis da Previsão de Consumo Residencial

Correlação	Consumo	Consumidores	Temp. Média	Tempo
Consumo	1			
Consumidores	0,791078329	1		
Temp. Média	0,319717223	-0,00943348	1	
Tempo	0,735500512	0,983900397	-0,051878657	1

A tabela 6 apresenta a MANOVA da equação de previsão do consumo residencial, onde utilizamos o valor do teste "F" para testar a significação dos regressores. Estando de acordo com o valor tabelado, considera-se uma boa regressão com valores confiáveis.

Tabela 6 - MANOVA de previsão do consumo residencial

	gl	SQ	MQ	F	F de significação
Regressão	3	6854854934	2284951645	3711,551382	1,9686E-132
Resíduo	140	86188549,56	615632,4969		
Total	143	6941043484			

Os coeficientes da equação e o valor de "R<sup>2</sup>" são apresentados na tabela 7.

Tabela 7 - Coeficientes das variáveis da equação de previsão do consumo residencial

Variável Y= Consumo Residencial em Petrolina	
Nº de Consumidores (X <sub>1</sub> )	0,072
Temperatura média (X <sub>2</sub> )	96,178
Tempo (X <sub>3</sub> )	5,036
R <sup>2</sup>	0,988



Intervalos de Confiança para 95%:

Número de Consumidores Residenciais em Petrolina:  $0,021 \leq X_1 \leq 0,124$

Temperatura Média em Petrolina:  $24,530 \leq X_2 \leq 167,825$

Tempo:  $-8,970 \leq X_3 \leq 19,042$

Para obtenção desses dados foi utilizado o *software* Excel®, versão 2007. Nas duas equações de previsão foram feitos testes de validação da regressão, sendo eles: teste de normalidade dos resíduos, em que se observou uma normalidade na curva de plotagem dos resíduos e teste de homocedasticidade, onde foi observada uma variância nos erros praticamente constante.

Essas duas equações mostram matematicamente uma previsão confiável do consumo de energia elétrica na cidade de Petrolina (PE). Como as empresas de distribuição possuem mecanismos de previsão para a quantidade de consumidores, através de metas ou estimativas, sendo a temperatura de fácil previsão, pois existem órgãos que fazem essa previsão, citando alguns desses órgãos na região EMBRAPA, com suas duas subestações agrometeorológicas em Petrolina e Juazeiro, INMET, Instituto Nacional de Meteorologia, com estações urbanas de meteorologia e a Estação Meteorológica da UNIVASF, Universidade Federal do Vale do São Francisco. Considerando essas variáveis independentes de fácil previsão, esse modelo de regressão linear múltipla se torna uma importante ferramenta para previsão do consumo de energia elétrica na região do Vale do São Francisco.

#### 4. CONCLUSÃO

Considerando a análise estatística feita podemos prever, mensalmente, a demanda de energia elétrica na região, partindo de regressões feitas com alta confiabilidade, temos então a garantia de uma previsão segura dos índices de consumo de energia elétrica na cidade de Petrolina (PE).

O emprego da regressão linear múltipla possibilitou a construção de um modelo de previsão confiável, cruzando dados de consumidores, temperatura e tempo, possibilitando a verificação de características relevantes dos dados apresentados, como visualização do período de racionamento, assim contribuindo para a adequação do melhor modelo de previsão utilizando regressão.



Comparando com, SANTOS & CALDAS (2010), temos uma previsão de energia elétrica bastante válida. E que juntos constituem trabalhos de demanda do maior pólo de consumo na região, colaborado assim para o crescimento planejado da região do Vale do São Francisco.

Assim o presente trabalho contribui significativamente com o seu propósito de oferecer métodos matemáticos de previsão utilizando pressupostos estatísticos. Assim as empresas da área e a própria população pode utilizar essas equações matemáticas para previsão do consumo de energia elétrica na região em um futuro próximo e assim prever melhorias nas redes de geração, transmissão e distribuição para melhor atender a população consumidora. Sendo muito importante considerando a falta de trabalhos com esse proposto na região do Vale do São Francisco.

## REFERÊNCIAS BIBLIOGRÁFICAS

Farias, Douglas Alexander, et Reinaldo Souza. Previsão da Carga de Energia em Patamares Combinando Técnicas Estatísticas e de Inteligência Computacional. Rio de Janeiro, 2009.

Pompeu, Cid Tomanick Enciclopédia Saraiva do Direito. Volume 32, Editora Saraiva, São Paulo, 1997.

Fogliatto, Flávio Sanson. Previsão de demanda por Energia Elétrica – Método e Aplicação. XXV Congresso Nacional de Eng. de Produção. Porto Alegre, RS, 2005.

Coelho-Barros, Emílio Augusto, et al. Métodos de estimação em regressão linear múltipla: aplicação a dados clínicos. Revista Colombiana de Estatística, junho de 2008. V. 31, n °1, p. 111 a 129.

Barbetta, Pedro Alberto, et al. Estatística para Cursos de Engenharia e Informática. Editora ATLAS. São Paulo, 2004.

Devore, Jay L., Probabilidade e Estatística: Para Engenharia e Ciências. 6ª Edição. Editora Thomson. São Paulo, 2006.

Hollander, Myles et Douglas Wolfe. Nonparametric Statistical Methods. 2ª Edição. Editora Wiley-Interscience. Toronto, 1999.

Montgomery, Douglas et al. Estatística Aplicada e Probabilidade para Enge-



nheiros. 2ª Edição. Editora LTC. Rio de Janeiro, 2003.

Mingoti, Sueli Aparecida. Análise de dados através de Estatística Multivariada: uma abordagem aplicada. Editora UFMG. Belo Horizonte, 2005.

Meyer, Paul. PROBABILIDADE: Aplicações à Estatística. 2ª Edição. Editora LTC. Rio de Janeiro, 1983.

Magalhães, Marcos Nascimento. et Antonio Carlos Pedroso de Lima, Noções de Probabilidade e Estatística. 6ª Edição. Editora Edusp. São Paulo, 2004

Martins, Gilberto de Andrade. Estatística Geral e Aplicada. 3ª Edição. Editora ATLAS. São Paulo, 2005.

Farias, Alfredo Alves et al. Introdução à Estatística. 2ª Edição. Editora LTC. Rio de Janeiro, 2003.

Bunchaft, Guenia et Sheila Rubino Kellner de Oliveira. Estatística sem Mistérios. 2ª Edição Revista, Volume II. Editora Vozes. São Paulo, 1998.

Hogg, Robert et Alen Craig. Introduction to Mathematical Statistics. 5ª Edição. Editora Pearson Educations. Dehli, 2004.

Holanda, Mônica Calixto Ribeiro de. et. al. Natimortalidade e Mortalidade até 21 dias de Idade em Leitões da Raça Large White. Revista Brasileira de Zootecnia. Recife. V.1. s.2, 2000.

Dias, J. L. G. et al. Metodologia para determinação de equações de regressão para avaliação das torres de transmissão. XIII ERIAC, 28 de maio de 2009, Porto Iguazú, Argentina.

Santos, Eucymara França Nunes et Alexandre Alex Silva e Caldas. "Equações de Regressão para Estimar o Consumo de Energia Elétrica na Cidade de Juazeiro-BA." III SEMAEST, UFS. Aracajú, 2010.