



**XXII SNTPEE
SEMINÁRIO NACIONAL
DE PRODUÇÃO E
TRANSMISSÃO DE
ENERGIA ELÉTRICA**

BR/GPL/03
13 a 16 de Outubro de 2013
Brasília - DF

GRUPO - VII

GRUPO DE ESTUDO DE PLANEJAMENTO DE SISTEMAS ELÉTRICOS – GPL

APLICAÇÃO DE TÉCNICAS DE MINERAÇÃO DE DADOS PARA PREVISÃO DE MERCADO DE ENERGIA ELÉTRICA – APRESENTAÇÃO DE METODOLOGIA E ESTUDO DE CASO DE UMA CONCESSIONÁRIA DE ENERGIA

**Raimundo C. GhizoniTeive(*)Fabiano F. AndradeEdison A. C. Aranha Neto
SEENERGIASEENERGIASEENERGIA**

**C. Celso de BrasilCamargoJorge Coelho Lucas M. Rosário João Airto de Bettio
SEENERGIA UFSCSEENERGIACelesc Distribuição**

RESUMO

Conforme o PRODIST, as distribuidoras devem realizar campanhas de medição para caracterização das cargas dos seus consumidores. Diversos trabalhos abordam esse tema para consumidores residências, porém para consumidores industriais, a literatura e as metodologias existentes deixam a desejar. Neste sentido, este artigo apresenta uma metodologia, baseada em mineração de dados, para previsão do mercado de energia elétrica de uma distribuidora, envolvendo tanto o levantamento das curvas de cargas típicas, quanto a previsão do crescimento de carga para os consumidores industriais, considerando as atividades industriais de forma segregada. O artigo apresenta ainda todos os procedimentos adotados para definição do universo amostral capaz de representar a classe industrial de toda área de concessão da distribuidora. Resultados preliminares com dados reais da concessionária apontam a viabilidade da metodologia.

PALAVRAS-CHAVE

Previsão de carga, Curva de carga, Mineração de Dados, Descoberta de Conhecimento.

1.0 - INTRODUÇÃO

O crescimento da demanda na área geográfica servida pela distribuidora é o fator mais importante e que mais influencia a necessidade de expansão do sistema. Portanto, o estudo da reação do sistema a esses incrementos de carga é essencial ao processo de planejamento. Entretanto, o comportamento do crescimento da demanda é influenciado por diversos fatores, especialmente o desenvolvimento social e econômico da área atendida pela concessionária.

Conforme previsto no Módulo de Planejamento da Expansão dos Sistemas de Distribuição do PRODIST(1), as distribuidoras devem caracterizar a carga de suas unidades consumidoras e o carregamento de suas redes e transformadores, por meio de informações oriundas de campanhas de medição. Pesquisas de posse de equipamentos e hábitos de consumo das classes consumidoras devem ser realizadas para caracterização das curvas de carga. As cargas e as redes devem ser caracterizadas por curvas de demanda típicas para dia útil, sábado e domingo.

A partir de dados de medição e resultados de questionários, a distribuidora de energia deve dispor de metodologias e modelos computacionais que possam processar estes dados e informações, gerando conhecimento novo e útil para auxiliar o seu processo de planejamento e projeção de mercado consumidor. Entretanto, esta tarefa não é trivial, envolvendo normalmente a utilização de métodos estatísticos e técnicas da Inteligência Artificial (IA). Uma técnica da IA que tem sido empregada com sucesso neste problema é a Mineração de Dados (MD). A MD é a principal etapa do processo de Descoberta de Conhecimento em Base de Dados (KDD - *Knowledge Discovery in Databases*), sendo responsável pela busca efetiva por conhecimento a partir dos dados transformados. Nessa fase,

são definidas as técnicas e os algoritmos a serem utilizados na tarefa em questão, sendo a parte mais complexa do processo de KDD.

Na literatura podem ser encontrados diversos trabalhos relacionados ao levantamento dos perfis típicos para diversas classes de consumidores de energia elétrica, e previsão de carga de forma geral (2),(3),(4). Curvas de carga típicas de consumidores residenciais já são bem conhecidas e não sofrem grandes alterações de padrão ao longo dos anos, bem como o problema de previsão de carga de consumidores residenciais, que é bem explorado na literatura. Por outro lado, devido à dinâmica de suas atividades, as classes comercial e industrial apresentam uma diversidade maior de perfis de carga diários. Além disso, a quantidade e profundidade dos trabalhos para essas classes consumidoras, considerando os problemas de levantamento do perfil de curva de carga e previsão de carga ainda não atingiram o mesmo patamar da classe residencial.

Neste artigo, é proposta uma metodologia e um modelo computacional, baseado em MD, para previsão do mercado de energia elétrica de uma distribuidora, envolvendo tanto o levantamento dos padrões do consumo de energia (curvas de cargas típicas), quanto a previsão do crescimento de carga (variação do montante de energia) para os consumidores industriais, considerando todos os segmentos CNAE (Classificação Nacional de Atividades Econômicas) relevantes do Estado. Neste caso, foram implementadas regras de associação, a partir das respostas de questionários aplicados a uma amostra selecionada de consumidores industriais, bem como o algoritmo de clusterização *k-means* foi aplicado sobre os dados de medição destes consumidores, gerando oito tipologias de agrupamentos dos perfis de consumo. Além disso, para cada divisão da indústria foi realizada a previsão da variação do montante de energia utilizando Redes Neurais Artificiais (RNAs), ajustadas para extrair informações relevantes da economia que afetam distintamente cada tipo de indústria.

O artigo apresenta ainda todos os procedimentos adotados para definição do universo amostral capaz de representar a classe industrial de toda a área de concessão da distribuidora. Todas as análises foram realizadas a partir de uma campanha de medições e levantamento de séries históricas de clientes horossazonais do banco de dados da concessionária. O nível de disponibilidade dessas informações foi determinante para o sucesso do levantamento estatístico.

A aplicação desta metodologia está prevista no escopo de um projeto de P&D ANEEL e resultará numa nova metodologia para dimensionamento dos transformadores de distribuição da empresa. Além de reduzir os custos com medições futuras de curva de carga de clientes individuais, a aplicação desta metodologia aproxima as necessidades dos consumidores das disponibilidades da distribuidora, melhorando a eficiência dos investimentos na rede de distribuição. Resultados preliminares com dados reais da concessionária apontam a viabilidade desta metodologia.

2.0 - AQUISIÇÃO DE DADOS E INFORMAÇÕES DOS CONSUMIDORES

Para que as técnicas de MD (Associação, Clusterização e Previsão) pudessem ser aplicadas neste trabalho, foi necessário inicialmente a aquisição de dados de medição dos consumidores e a obtenção de informações complementares sobre o comportamento destes consumidores industriais. No caso de dados de medição, a concessionária disponibilizou apenas dados de consumidores horo-sazonais (HS), sendo necessária a medição de consumo dos demais consumidores (grupo A e B), optando-se pela medição do consumo semanal. A estratégia utilizada para medição do consumo dos consumidores industriais e a respectiva aquisição da informações complementares é descrita a seguir.

2.1 Definição das Atividades CNAE

Para alcançar os objetivos previstos neste trabalho, foi necessário definir uma amostra com elevada significância estatística. Em nível de informação, considerando-se dados de 2011, o número de consumidores industriais no estado de Santa Catarina era de 26.488, abrangendo 29 divisões CNAE (CNAE 10 a 37 e 40). Neste caso, considera-se a versão 1.0 do CNAE – Classificação Nacional de Atividades Econômicas.

O número de regionais da concessionária também foi um fator que contribuiu para a definição das atividades CNAE a serem consideradas, buscando-se manter uma boa representatividade em termos de consumo de energia e de número de consumidores. Das 16 regionais existentes, buscou-se selecionar oito regionais que pudessem representar mais da metade do consumo e do número de consumidores das classes industriais do estado. Além disso, a pesquisa deveria se concentrar nas divisões do CNAE mais representativas, ou seja, as divisões selecionadas deveriam contemplar pelo menos 75% do consumo dentro de cada uma das duas classes de consumidores. Após muitas análises foi possível encontrar oito regionais que tivessem a representatividade desejada, considerando todas as divisões CNAE. Estas regionais são: Grande Florianópolis, Blumenau, Concórdia, Chapecó, Joaçaba, Joinville, Lages e Videira.

A partir da definição das regionais a serem consideradas no processo de medição foi possível definir as atividades CNAE da indústria a serem consideradas no cálculo da amostra para medição. Neste caso, manteve-se como parâmetro de referência a representatividade de consumo e de número de consumidores. Assim, considerando as oito regionais escolhidas, conseguiu-se representar, com apenas 8 divisões CNAE, mais de 76%

do consumo de energia e cerca de 65% dos consumidores industriais de Santa Catarina. As atividades CNAE definidas para a indústria estão apresentadas na Tabela 1.

Tabela 1 - Atividades Consideradas - CNAE Indústria

Divisão CNAE	Descrição CNAE	Consumo (%)	Num. Cons. (%)	Num. Amostras
15	Fabricação de produtos alimentícios e bebidas	19,28	12,71	65
17	Fabricação de produtos têxteis	14,09	8,14	64
18	Confeção de artigos de vestuário e acessórios	2,16	10,92	65
20	Fabricação de produtos de madeira	4,92	14,63	66
25	Fabricação de artigos de borracha e plástico	7,30	3,17	59
27	Metalúrgica básica	12,62	1,83	54
28	Fabricação de produtos de metal	7,28	10,39	65
29	Fabricação de máquinas e equipamentos	8,49	3,13	59
Total		76,14	64,93	497

2.2 Cálculo da Amostra para Medição

Ainda que a concessionária armazene e disponibilize dados de medição da parcela de consumidores com tarifas horossazonais (HS), a quantidade restante de consumidores a serem medidos nas oito regionais selecionadas ainda seria muito grande. Assim, houve a necessidade da definição de uma amostra adequada para medição em campo.

Para o cálculo das amostras de cada divisão, consideraram-se os seguintes objetivos:

- preservar as características das 8 regionais da concessionária no estado;
- manter todas as divisões CNAE (Tabela 1) selecionadas com os mesmos erros amostrais;
- erro amostral fixo de 10%;
- nível de confiança de 90%.

Ao ser estabelecido o universo amostral da pesquisa com as oito regionais selecionadas e oito divisões CNAE industriais, foram estabelecidas as condições adequadas para a construção de um plano de medições viável técnica e economicamente para o projeto. Os resultados da pesquisa terão validade e significância estatística, ao passo que o número de medições de clientes não extrapola o orçamento previsto para esta atividade. Cabe ressaltar que alguns clientes apresentam medição horossazonal, dispensando assim eventuais medições em campo. Na Tabela 1 tem-se o número de amostras necessárias para representar estatisticamente cada divisão CNAE selecionada para a classe industrial.

A partir da Tabela 1 foi possível construir uma proposta de plano de medições por regionais, definindo-se finalmente o número de medições necessárias para cada regional e cada subdivisão CNAE, em nível de consumidor industrial. A Tabela 2 apresenta um exemplo de proposta de plano de medição obtido. Neste caso, considerou-se o CNAE 15 da indústria, "Fabricação de produtos alimentícios e bebidas", o qual contém 33 subdivisões.

Tabela 2 - Plano de Medição CNAE 15 e Subdivisões

CNAE	Descrição	Florianópolis	Blumenau	Chapécó	Joaçaba	Joinville	Lages	Videira	Concórdia
1500	FABRICACAO DE PRODUTOS ALIMENTICIOS E B		1						
1543	FABRICACAO DE SORVETES	1	1	1	1	1			
1551	BENEFICIAMENTO ARROZ E FABR PROD ARRO		1	1					
1581	FABR PRODUTOS DE PADARIA, CONFEI PASTEL	9	6	3	1	9	2	1	1
1589	FABRICACAO DE OUTROS PRODUTOS ALIMENTI	1	1	1		1			

Como pode ser observado na Tabela 2, cada regional terá alguma subdivisão amostrada, sendo que o número de amostras resulta dos diferentes níveis da atividade de fabricação de produtos alimentícios em cada uma delas. O total de amostra de medições para este CNAE é 65, sendo que na Tabela 3 são apresentadas somente algumas subdivisões de um total de 33 subdivisões existentes nesta divisão CNAE.

2.3 Aplicação de Questionários

Além dos dados de medição dos consumidores industriais e comerciais do estado, foi necessário obter algumas informações dos consumidores medidos, para identificação do comportamento destes consumidores e da atividade econômica correspondente. Isto é importante para que os algoritmos relacionados às tarefas de mineração de dados, utilizados neste trabalho, possam ser executados adequadamente, gerando conhecimento novo e útil para a concessionária. Assim, uma equipe foi contratada para aplicar os questionários nos consumidores selecionados para medição. De forma geral, o questionário foi dividido em cinco partes, envolvendo identificação da unidade consumidora, informações sobre a instalação, informações sobre a operação da empresa em 2011, influência da economia e do clima no negócio. Maiores detalhes sobre o questionário podem ser encontrados em (www.seenergia.com.br/questionario).

3.0 - MINERAÇÃO DE DADOS

3.1 Clusterização dos Perfis de Carga Diários

Uma das abordagens mais fundamentais para o entendimento e aprendizagem é a organização de dados em grupos ou agrupamentos. A análise de clusters ou clusterização é um estudo formal de métodos e algoritmos para um agrupamento natural ou clusterização de objetos de acordo com suas características mensuradas ou intrínsecas de similaridade. As amostras para agrupamento são representadas como um vetor de medições, ou mais formalmente, como um ponto em um espaço multidimensional (5). Amostras dentro de um cluster válido são mais similares entre si do que são para amostras pertencentes a diferentes clusters. A clusterização *K-means*, utilizada nesse trabalho para agrupar os perfis das curvas de carga, agrupa os dados de perfil de carga pela determinação de certo número de clusters e do ponto central de cada cluster. Após a determinação do ponto central, cada conjunto de dados deve ser associado ao ponto central mais próximo, recalculando então o novo ponto central. Isto será feito iterativamente até a estabilidade do ponto central.

3.2 Classificação Funcional das Tipologias

Uma das técnicas utilizadas para encontrar relacionamentos em bases de dados é a análise de regras de associação. Esta técnica tem sido intensamente pesquisada e consiste em encontrar correlações entre os dados minerados. O algoritmo APRIORI analisa as regras de associação em uma base de dados, descobrindo e elegendo padrões verdadeiros e eliminando padrões que ocorrem simplesmente ao acaso (6). Utiliza regras de associação do tipo: $X \rightarrow Y$ (X leva a Y) onde X e Y são conjuntos mutuamente exclusivos de itens. O cálculo do suporte e confiança são dados pela equação (1), onde o suporte representa a frequência em que as associações surgem, enquanto que a confiança mede a confiabilidade da inferência feita por uma regra. Para uma determinada regra $X \rightarrow Y$, quanto maior a confiança, maior a probabilidade de que Y esteja presente em transações que contenham X.

$$\text{Suporte} = \frac{\text{Número de registros com } X \text{ e } Y}{\text{Número total de registros}} \quad \text{Confiança} = \frac{\text{Número de registros com } X \text{ e } Y}{\text{Número de registros com } X} \quad (1)$$

No caso particular deste trabalho, os antecedentes das regras de associação são os padrões observados nos resultados obtidos dos questionários aplicados nos consumidores de interesse; enquanto que os consequentes das regras de associação são os *cluster* resultantes obtidos da etapa de clusterização, envolvendo os perfis de cargas dos consumidores industriais. Maiores detalhes sobre o levantamento de tipologias de curvas de carga através de algoritmo k-means e de regras de associação podem ser encontrados em (7).

Utilizou-se neste trabalho o software livre para exploração de dados, aprendizado de máquina e mineração de dados, chamado Tanagra (8). Para trabalhar com as regras de associação, a função escolhida foi a SPV (*Supervised association generator*), que é uma variante do algoritmo APRIORI, o método de regra por indução supervisionada.

4.0 - RESULTADOS

4.1 Comportamento do Consumo de Energia Elétrica da Indústria

Para compreender melhor a dinâmica da variação do consumo de energia elétrica da indústria no Estado de Santa Catarina foi necessário construir uma série histórica do consumo das principais divisões (CNAE) atendidas pela concessionária. Foram utilizados dados de 1997 a 2011, separados conforme o CNAE. A Figura 1a apresenta a evolução do comportamento do consumo médio mensal (CMM) das 6 principais divisões consumidoras. A Figura 1b apresenta os mesmos consumos médios mensais, porém com uma média móvel de 12 meses. Pode-se verificar quão heterogêneo é o comportamento da demanda entre as divisões ao longo dos anos.

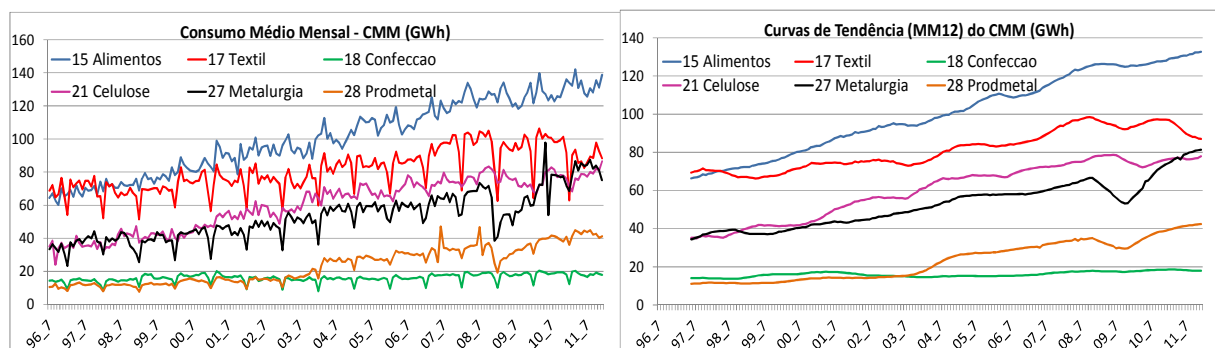


FIGURA 1 a) – Consumo Médio Mensal (CMM). b) – Média Móvel de 12 meses (MM12) do CMM.

4.2 Estudo sobre Comportamento do Consumo de Energia Elétrica da Indústria de Alimentos (Divisão 15)

A Figura 2a apresenta o comportamento do consumo médio mensal da divisão 15 desde 1998 até 2011. Pode-se notar que não existe um padrão de variação do consumo de um mês para outro e, pelo contrário, as variações são rápidas e difíceis de ser previstas ou correlacionadas com outras variáveis explicativas. Portanto, foram realizadas simulações considerando não o CMM de cada mês individualmente (linha preta), mas sim as médias móveis de 6 (MM6) ou de 12 meses (MM12). A série construída MM6, por não ser uma janela tão longa consegue reagir relativamente bem às variações da economia ou do clima ao mesmo tempo em que apresenta boas correlações parciais com as suas variações explicadas. Médias móveis mais curtas são difíceis de correlacionar e as mais longas demoram muito para carregar os efeitos de variações da economia, mesmo as mais intensas.

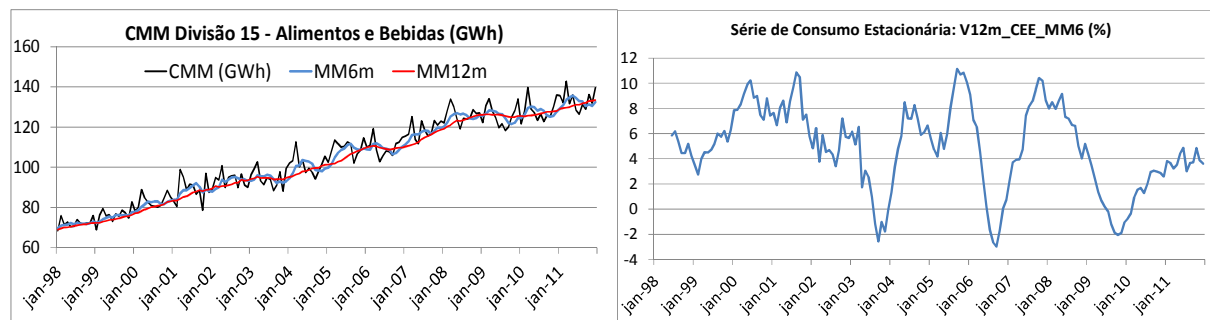


FIGURA 2 – a) Médias Móveis de 6 (MM6) e 12 meses (MM12). b) Variação da MM6 do CEE em Relação ao Mesmo Período do Ano Anterior (V12m_CEE_MM6).

A MM6 não é estacionária e, portanto não pode ser diretamente inserida como variável dependente do modelo econométrico. Para torná-la estacionária, optou-se pelo cálculo da variação percentual em relação ao mesmo período do ano anterior. Denominada aqui de variação 12 meses ou V12m (Figura 2b), esta operação consegue extrair as mudanças no ritmo de crescimento ou decrescimento do consumo de um ano para outro.

4.3 Formação do Conjunto de Variáveis Explicativas

A escolha das variáveis explicativas do comportamento da demanda é uma das etapas cruciais da elaboração de um modelo econométrico. Focado na divisão CNAE 15 (alimentos e bebidas), optou-se pela seleção de variáveis predominantemente macroeconômicas. Pôde-se utilizar as respostas obtidas junto aos consumidores industriais da divisão 15 quando da aplicação do Questionário de Caracterização Funcional da etapa anterior. Por meio daquele questionário, os empresários e gestores já puderam indicar quais as variáveis de maior influência para seus negócios.

Após consultar diversas bases de dados, a saber: Banco Central do Brasil, IBGE, Fundação Getúlio Vargas e Confederação Nacional da Indústria, foi possível obter 18 variáveis econômicas: PIB (cinco variantes), valor agregado, imposto, consumo das famílias, consumo da administração pública, valor da formação bruta do capital fixo, exportação e importação, TCER, IGPM, TJLP, Selic, TR e IPCA. Foi possível então construir 24 várias séries históricas estacionárias para o exemplo de aplicação.

As técnicas utilizadas para tornar as séries estacionárias foram:

- V12m – variação percentual em relação ao mesmo mês do ano anterior ou
- a12m – valor acumulado nos últimos 12 meses, utilizado para taxas de inflação.

A Figura 3a, 3b, 4a e 4b apresentam o comportamento de algumas das variáveis explicativas em suas respectivas formas estacionárias. Todos os valores medidos são percentuais (%).

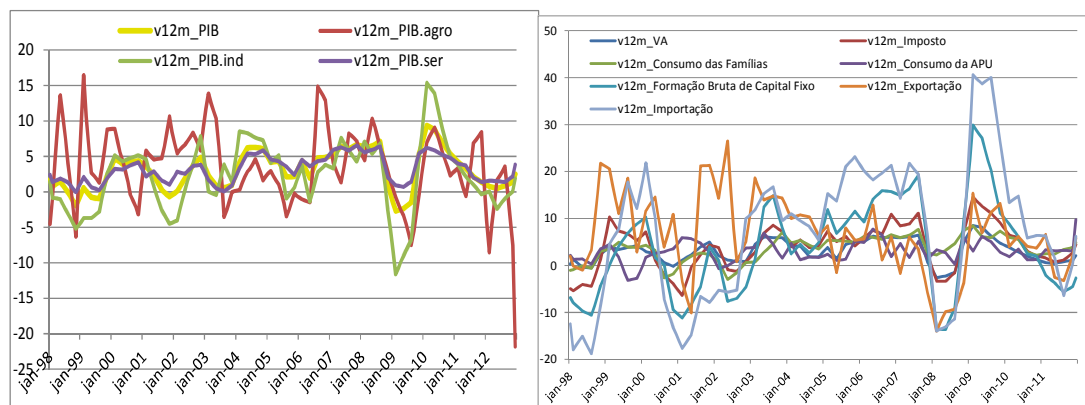


FIG. 3 – a) PIB por setores.

b) PIB por componentes

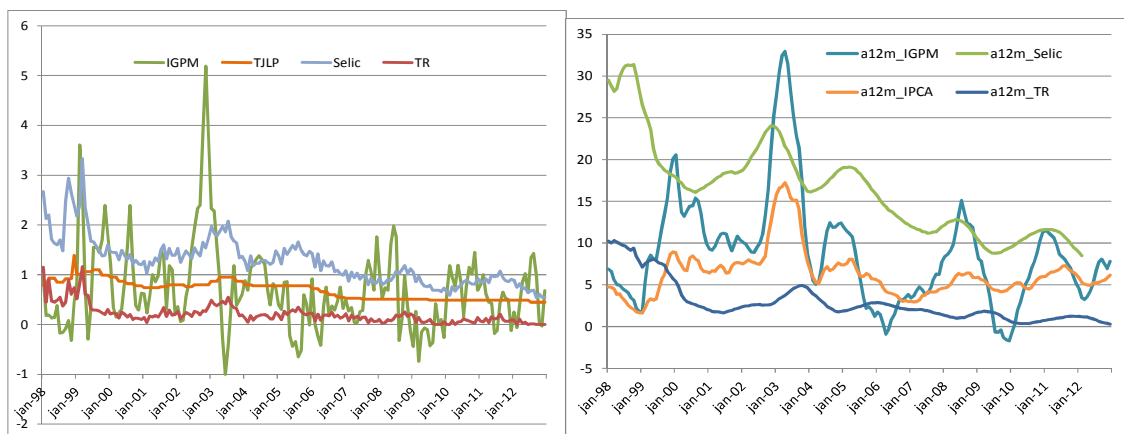


FIGURA 4- a) Juros e Inflação. b) Var. e Acum. de Juros e Inflação.

4.4 Resultados do Modelo Econométrico.

Como não se tem conhecimento prévio sobre os intervalos de tempo em que ocorrem os impactos das variáveis explicativas (econômicas) sobre a variável dependente (consumo de energia elétrica), todas as 24 variáveis explicativas foram correlacionadas com a série histórica de V12m_CEE_MM6 para defasagens de 0,1, 2, 3... até 12 meses. Esse procedimento permite identificar os efeitos de políticas econômicas cuja constante de tempo encontra-se no médio prazo. A redução da taxa de juros ou o aumento da renda, por exemplo, poderiam impactar positivamente na cadeia produtiva da divisão 15, porém o impacto no consumo de energia elétrica pode não ser imediato.

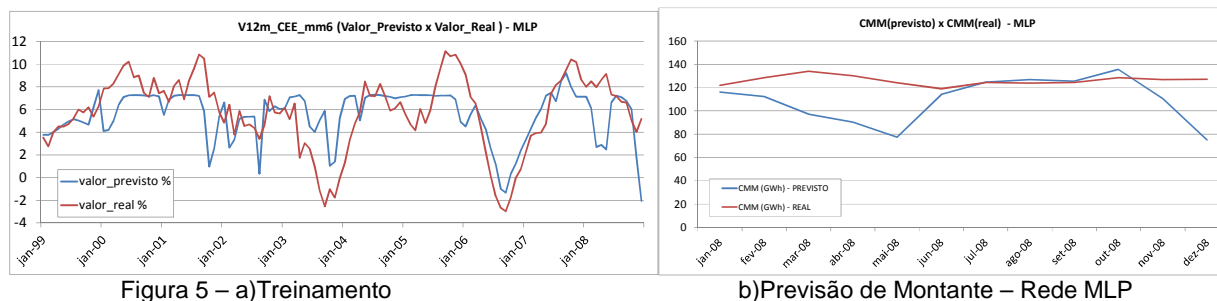
Os resultados obtidos mostram-se coerentes com as expectativas de um modelo econômico, onde o crescimento do PIB, renda e crescimento das exportações induzam ao aumento da produção industrial e por consequência do consumo de energia elétrica. Os atrasos mensais na ordem 3 a 6 meses para variáveis que estimulam a produção industrial também estão dentro do esperado. Entre todas as variáveis testadas, destacam-se a influência do PIB da indústria de transformação e as exportações, com correlações parciais moderada, de 0,425 e 0,405, respectivamente. Ressalta-se que, individualmente, nenhuma variável econômica ou climática conseguiria obter uma grande correlação com uma série de consumo específica da indústria. Como o modelo econométrico é multivariável, é possível que com algumas correlações individuais mais fracas ou moderadas se obtenha uma boa explicação do comportamento do consumo como um todo.

Finalmente, conclui-se esta etapa com a indicação das melhores variáveis econômicas capazes de descrever o comportamento do consumo de energia elétrica da indústria de alimentos e bebidas de Santa Catarina:

- V12m_PIB Indústria de Transformação (0,425 com atraso de 6 meses)
- V12m_Exportações (0,405 com atraso de 3 meses)
- V12m_TJLP (-0,258 com atraso de 4 meses)

4.5 Previsão de montante

Considerando duas variáveis econômicas (V12m_PIB Indústria e V12m_exportações) e o consumo médio mensal dos seis meses anteriores (média móvel MM6) como entradas, foram implementadas duas RNAs: MLP backpropagation e RBF. Na Figura 5 é apresentado os resultados com a RNA MLP.



Na Figura 5 os resultados foram obtidos com os dados de treinamento de 1999 a 2007 e testes feitos com os dados de 2008. A rede MLP três camadas (5 neurônios na camada intermediária) teve os seguintes parâmetros de treinamento: taxa de aprendizagem 0,07; momentum 0,7 e erro de 0,03.

Os resultados obtidos com a rede RBF três camadas (8 neurônios na camada intermediária) são apresentados na Figura 6. Neste caso, a rede RBF foi modelada com centro inicial 0,5; peso inicial 0,5 e *spread* entre 0,4 e 0,5. Os dados de treinamento foram os mesmos da rede MLP, assim como os dados usados para teste.

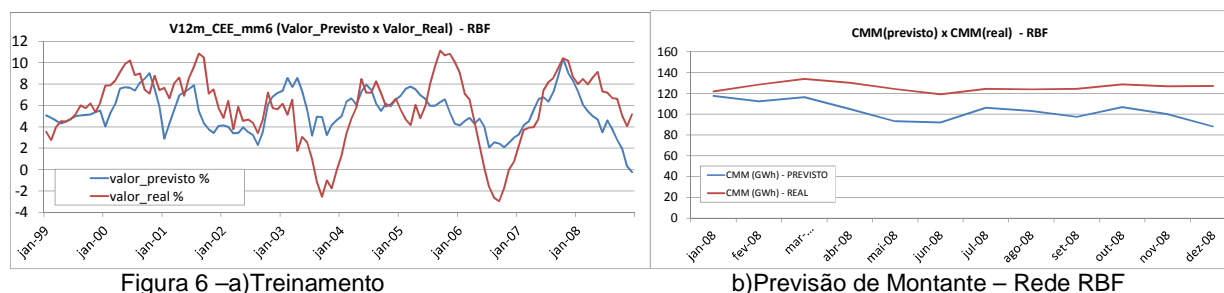


Figura 6 –a)Treinamento

b)Previsão de Montante – Rede RBF

A parte de previsão de montante ainda está em desenvolvimento e os resultados atuais ainda são incipientes. Porém as variáveis de entrada da rede já foram definidas a partir da análise econométrica. O foco do trabalho neste momento é a melhoria do modelo neural, em termos dos parâmetros de treinamento e de topologia.

5.0 - CONCLUSÃO

A caracterização do comportamento da demanda, inserida no processo de planejamento da distribuição, é um problema complexo e relevante para uma concessionária de energia. É complexo, pois o comportamento da demanda é influenciado por diversos fatores, entre eles destacam-se as características funcionais das empresas consumidoras e o nível de desenvolvimento social e econômico da área atendida pela concessionária. Além disto, este problema é relevante, pois a estimativa da demanda é fundamental para que a concessionária mantenha seus padrões de qualidade de serviço e não seja penalizada pelo agente regulador.

Neste trabalho foi apresentada uma metodologia com todos os procedimentos adotados para a definição do universo amostral representativo da classe industrial em Santa Catarina, para criação de regras de associação que conseguem estimar com boa precisão, os perfis de carga diárias, tomando como base as características funcionais de cada cliente, e finalmente para previsão do montante de energia baseada em modelos neurais, cujas variáveis de entrada são de ordem econômica e customizadas para cada divisão da indústria.

De um modo geral, o processo de mineração aqui apresentado conseguiu agregar conhecimento novo e útil sobre o comportamento da demanda de consumidores industriais, permitindo aos gestores do sistema de distribuição antecipar sua tomada de decisão às variações do consumo de seus clientes, tanto ao longo do dia quanto ao longo dos meses. Esse novo patamar de conhecimento traz consequências positivas diretas para a otimização de dimensionamento dos equipamentos das redes de distribuição.

6.0 - REFERÊNCIAS BIBLIOGRÁFICAS

- (1) ANEEL (2010). Prodist - Procedimentos de distribuição de energia elétrica no Sistema Elétrico Nacional. Disponível em: <http://www.aneel.gov.br>.
- (2) N. Anuar, Z. Zakaria. "Cluster Validity Analysis for Electricity Load Profiling". *Electrical Engineering*, p.35-38. 2010.
- (3) J. F. M. Pessanha, L. C. Laurencel, R. C. Souza. "Mapa de Kohonen na Construção de Tipologias de Curvas de Carga," in XXXVI SBPO - *Simpósio Brasileiro de Pesquisa Operacional*, São João del Rei, 2004.
- (4) J. F. M. Pessanha, V. L. O. Castellani, A. L. A Araújo. "Uma Nova Ferramenta Computacional para Construção de Tipologias de Curvas de Carga," in SEPOPE2006 - *X Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica*, Florianópolis, 2006.
- (5) M. Kantardzic. *Data Mining: Concepts, Models, Methods, and Algorithms*. 2nd ed. Wiley-IEEE Press. 2011.
- (6) R. Agrawal, Srikant, Ramakrishnan. "Fast Algorithms for Mining Association Rules", In: *20th VLDB Conference*, 487-498, Santiago, Chile, 1994.
- (7) F. F. Andrade, R. C. G. Teive, E. A. C. Aranha Neto, L. Oliveira, C. C. de B. Camargo, L. M. Rosário, J. Coelho, J. A. de Bettio Metodologia para Descoberta de Clusters Mínimos e Significativos dos Perfis de Cargas Típicos de Consumidores Industriais e Comerciais de uma Distribuidora de Energia Elétrica. Artigo a ser apresentado no XV ERIAC. Foz do Iguaçu. Pr. Maio. 2013.
- (8) R. Rakotomalala. "Tanagra". Disponível em: < <http://chirouble.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html> >. Acesso em 10 de janeiro de 2013.

7.0 - DADOS BIOGRÁFICOS

Raimundo C. GhizoniTeive - engenheiro eletricitista formado pela UFSC, com mestrado e doutorado pela própria UFSC, em 2001 e 2007, respectivamente. É professor titular na Universidade do Vale do Itajaí e pesquisador participante do programa de pós-graduação em engenharia elétrica da UFSC, onde atua principalmente nas áreas de planejamento da transmissão e distribuição, manutenção de sistemas elétricos e comercialização de energia.

Fabiano F. Andrade possui graduação em Engenharia Elétrica pela UFU (2000) e mestrado e doutorado pela UFSC (2003 e 2009). Atualmente é professor adjunto da Universidade do Estado de Santa Catarina. Tem experiência na área de Planejamento de Sistemas de Energia Elétrica, com ênfase em sistemas de distribuição e comercialização.

Edison A. C. Aranha Neto possui graduação em Engenharia Elétrica pela UFU (2000) e mestrado e doutorado pela UFSC (2006 e 2012). Atualmente é pesquisador em diversas empresas, atuando principalmente com P&Ds. Tem experiência em sistemas de distribuição, metodologia multicritério, perdas e qualidade da energia elétrica.

C. Celso B. Camargo - engenheiro eletricitista pela UFJF e mestrado em engenharia elétrica pela COPPE em 1977. Doutor em Engenharia de Produção pela UFSC em 1996. Coordenador e pesquisador de P&D com várias empresas e professor aposentado da UFSC. Tem experiência na área de Confiabilidade, GLD, aplicação de Métodos Probabilísticos em Sistemas de Energia Elétrica. Possui 05 livros publicados.

Jorge Coelho - graduação e mestrado pela UFSC em 1977 e 1980, respectivamente, doutorado pela PUC_RJ em 1990. Coordenador e pesquisador de P&D com várias empresas e professor titular da UFSC. Tem experiência na área de Confiabilidade, Qualidade de Energia, Proteção, Modelagem de Incertezas, Sistemas de Grande Porte.

Lucas M. Rosário – Engenheiro de Computação formado pela Universidade do Vale do Itajaí em 2005. Atualmente está realizando mestrado pela mesma universidade, tendo como foco o desenvolvimento e aplicação de modelos de previsão de carga para descrição de consumo de consumidores industriais.

João Airto de Bettio – Engenheiro eletricitista, lotado na Divisão de Engenharia de Medição – DVMD, Celesc Distribuição.