



XVIII Seminário Nacional de Distribuição de Energia Elétrica

SENDI 2008 - 06 a 10 de outubro

Olinda - Pernambuco - Brasil

Gestão da Previsão de Consumo e Energia Não Faturada

Carlos Alberto Fróes Lima	Marley Apolinário Saraiva	Juliano Andrade
KNBS	KNBS	CPFL
froes@knbs.com.br	saraiva@knbs.com.br	julianoandrade@dominio.com.br
Glauber Renato Colnago	Rodolfo Miyasaki	Edson Amaral
KNBS	KNBS	CPFL
colnago@knbs.com.br	miyasaki@knbs.com.br	edsonamaral@dominio.com.br

Palavras-chave

Energia Não Faturada

Perdas Comerciais

Previsão do Consumo de Energia Elétrica

Rede Neural Artificial

SARIMA

Resumo

Este projeto inclui métodos e algoritmos para estimar a quantidade de energia não faturada, permitir previsões de consumo e receita e desagregar as perdas comerciais do valor da energia não faturada. Foi realizado com base na pesquisa do conhecimento existente na concessionária, propondo-se novos padrões de análise, implementado em um protótipo computacional para realizar a gestão da previsão de consumo de energia e receita não faturadas.

A ferramenta permite a geração de um critério único de avaliação do consumo para uso na concessionária, aplicável em diversas áreas: financeira, mercado, comercial e perdas. Além disto, apresenta facilidades para avaliar o consumo ainda não faturado (considerando-se a projeção do consumo dos clientes) e os dados de demanda atendida no mesmo período, ou seja, em função do volume de energia e do total do consumo projetado para o período, gerando indicadores mais precisos de perdas. Esta avaliação baseia-se em valores históricos de consumo relativos a um período de até cinco anos, definindo o consumo faturado do mês e o consumo do mês civil após a data de leitura, adotando de técnicas matemáticas e estatísticas de modelagem, como modelos de séries temporais (SARIMA), redes neurais artificiais (RNA) e interpolação cúbica monótona por partes.

1. INTRODUÇÃO

Este trabalho foi desenvolvido através do Programa de Pesquisa e Desenvolvimento da ANEEL (Agência Nacional de Energia Elétrica), pela KNBS, empresa de tecnologia e gestão do conhecimento, em parceria da CPFL (Companhia Paulista de Força e Luz).

Devido a operação diferenciada, inclusive de suas fontes de dados, a pesquisa desenvolveu-se de forma independente sobre as duas empresas envolvidas do grupo, a CPFL Paulista e a CPFL Piratininga. Para caracterizar a quantidade de energia não faturada, permitir previsões de consumo e receita e desagregar as perdas comerciais do valor da energia não faturada, inclusive como respaldo às análises futuras das influências específicas de cada grupo nas perdas comerciais, foram considerados os grupos A e B destas empresas em separado. Foi assertiva esta separação, constatando-se que apresentavam comportamentos de consumo diferentes. Desta forma, cada um destes quatro grupos foi tratado separadamente, obtendo-se assim, resultados mais precisos com relação à energia não faturada e à previsão de consumo de energia elétrica.

O cálculo da energia fornecida tem forte influência no balanço energético. O método atual da companhia para este cálculo considera como energia fornecida a diferença entre as quantidades obtidas em duas medições de cada consumidor, resultando na soma da energia faturada no mês atual com a energia não faturada do mês anterior, geralmente não coincidindo com início e fim do mês civil. Este método gera um valor chamado de *float*, que é devido à diferença entre a energia não faturada de meses consecutivos.

Como alternativa a este método, foi proposto estimar a energia fornecida no mês civil através de interpolação cúbica monótona por partes, proposta em 0, eliminando o *float*. Esta estimativa é realizada até o mês anterior ao mês de referência do balanço energético. Em seguida, a estimativa da energia fornecida no mês civil de referência do balanço energético foi feita com métodos preditivos, tendo como dados a série de consumos estimada pelos métodos interpolativos. Os métodos utilizados para previsão foram: redes neurais artificiais (RNA) **Erro! Fonte de referência não encontrada.** e modelos SARIMA (*Seasonal Auto-Regressive Integrated Moving-Average*)¹ 0 para previsão de séries temporais. A utilização e comparação de dois métodos distintos, um analítico e outro estocástico, foi proposital e atende aos anseios apresentados de se estabelecer critérios de análise imediatos e apoiar a previsibilidade de evolução do sistema.

Uma vez que se tenha previsto o consumo do mês civil referente ao balanço energético, é possível o cálculo da energia não faturada do próprio mês e assim calcula-se a energia fornecida do mês atual. Com isto, o balanço energético se torna mais preciso, uma vez que todos os valores considerados no cálculo são referentes ao mês civil e a variação do erro de estimação é menor que a variação do *float*.

Esta análise foi realizada em cada grupo de consumidores, e quando possível ou necessário, a consumidores específicos, permitindo verificar sua influência nas perdas comerciais.

2. DESENVOLVIMENTO E RESULTADOS

2.1. Metodologia de trabalho

¹ Sazonal Auto-Regressivo Integrado de Média-Móvel

Na Tabela 1 estão os dados com percentuais de consumo e consumidores dos grupos A e B da CPFL Paulista e Piratininga. Nota-se que o grupo A, em ambas, apresenta maior consumo com uma quantidade de consumidores muito menor.

Tabela 1 – Porcentagem de Consumo e Consumidores Por Grupo

CPFL	GRUPO	% CONSUMO (Médio)	% CONSUMIDORES (Jul/2007)
PAULISTA	Grupo A	56,06613%	0,36031%
	Grupo B	43,93387%	99,63969%
PIRATININGA	Grupo A	70,624%	0,3034%
	Grupo B	29,376%	99,6966%

Para obter o valor da energia fornecida do mês atual, é necessário o cálculo da energia não faturada. Para obter este valor, é necessário calcular o consumo de energia no mês civil e também prever o próximo valor deste consumo. O método atual da companhia para este cálculo considera a soma da energia faturada no mês atual com a energia não faturada do mês anterior, geralmente não coincidindo com início e fim do mês civil, principalmente devido ao processo de leituras realizados. Este método gera um valor chamado de *float*, que é devido à diferença entre a energia não faturada de meses consecutivos.

Como alternativa a este método, foi proposto estimar a energia fornecida no mês civil através de interpolação cúbica monótona por partes, eliminando o *float*, apresentado na Figura 1. Esta estimativa é realizada até o mês anterior ao mês de referência do balanço energético. Os procedimentos utilizados para estes cálculos serão descritos a seguir.

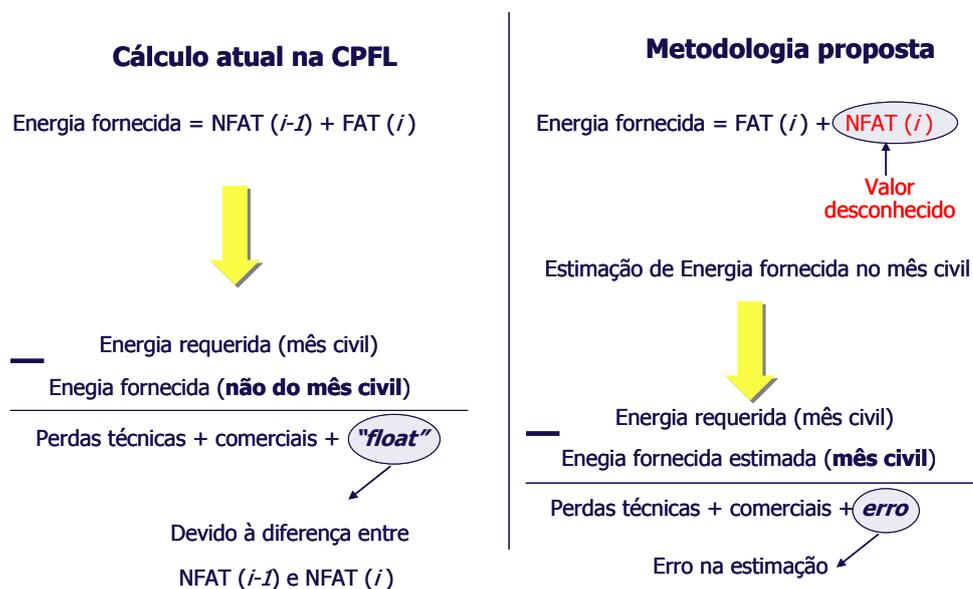


Figura 1 – Balanço Energético

2.2. Cálculo da Energia Consumida no Mês Civil

Obtidos os dados de consumo por período de faturamento e as respectivas datas de leitura do consumo, é possível estimar o consumo por mês civil através de interpolação cúbica monótona por partes 0.

As equações dos polinômios cúbicos são da seguinte forma:

$$c_i(x) = \frac{\Delta_i}{h_i^2} \left[(\alpha + \beta - 2) \cdot (x - x_i)^3 - (2\alpha + \beta - 3) \cdot h_i \cdot (x - x_i)^2 + \alpha \cdot h_i^2 \cdot (x - x_i) \right] + y_i$$

e suas derivadas de 1º e 2º graus:

$$c'_i(x) = \frac{\Delta_i}{h_i^2} \left[3 \cdot (\alpha + \beta - 2) \cdot (x - x_i)^2 - 2 \cdot (2\alpha + \beta - 3) \cdot h_i \cdot (x - x_i) + \alpha \cdot h_i^2 \right]$$

$$c''_i(x) = \frac{2\Delta_i}{h_i^2} \left[3 \cdot (\alpha + \beta - 2) \cdot (x - x_i) - (2\alpha + \beta - 3) \cdot h_i \right]$$

onde

$$\Delta_i = \frac{(y_{i+1} - y_i)}{h_i} \text{ e } h_i = x_{i+1} - x_i.$$

Estas equações garantem que ocorra a interpolação dos dados e que a derivada de primeiro grau seja contínua. A derivada de segundo grau não é necessariamente contínua.

Uma condição necessária para que $c_i(x)$ seja monótona em $[x_i, x_{i+1}]$ é que $\alpha \geq 0$ e $\beta \geq 0$. Além disso, a monotonicidade de $c_i(x)$ está diretamente relacionada ao ponto ótimo da função quadrática $c'_i(x)$:

$$x^* = x_i + \frac{h_i}{3} \left[\frac{2\alpha_i + \beta_i - 3}{\alpha_i + \beta_i - 2} \right]$$

Assim, uma condição necessária e suficiente para que $c_i(x)$ seja monótona em $[x_i, x_{i+1}]$ é $x^* \notin [x_i, x_{i+1}]$ ou $x^* \in [x_i, x_{i+1}]$ com os sinais de $c'_i(x^*)$ e Δ_i iguais.

Uma vez obtidas as curvas de consumo para cada consumidor, é possível estimar o consumo por mês civil da forma como ilustrada na Figura 2. Nesta ilustração, considera-se o valor do consumo no último dia do mês menos o valor do consumo no dia da data de leitura. Esta diferença resulta na energia não faturada do mês. A diferença entre o consumo na data de leitura e o consumo no primeiro dia do mês resulta na energia faturada do mês. Somados os valores de energia faturada e não faturada do mês tem-se a energia consumida no mês civil. Após esta estimativa para cada consumidor, somam-se os valores de cada mês, obtendo-se os valores de consumo totais da concessionária.

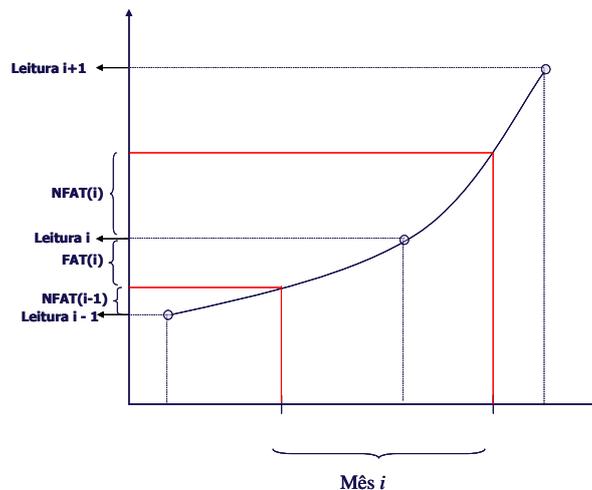


Figura 2 – Consumo por mês civil

Com este método, é necessário prever o valor do consumo no mês $i + I$ para que seja possível estimar o valor da energia não faturada no mês i . Esta previsão é baseada nos valores de consumo por mês civil. As técnicas utilizadas para previsão são descritas a seguir.

2.3. Método SARIMA

Os modelos SARIMA são apropriados para descrever séries temporais que apresentam sazonalidade e tendência. Em geral, os modelos SARIMA(p,d,q)S(P,D,Q) são dados pela expressão 0:

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D Y_t = \theta(B)\Theta(B^S)a_t$$

onde

$$(1-B)Y_t = Y_t - Y_{t-1}, \quad \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

$$\Phi(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_p B^{pS}, \quad (1-B)^d = 1 - B - B^2 - \dots - B^d,$$

$$(1-B^S)^D = 1 - B^S - B^{2S} - \dots - B^{DS}, \quad \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q,$$

$$\Theta(B^S) = 1 + \Theta_1 B^S + \Theta_2 B^{2S} + \dots + \Theta_Q B^{QS},$$

a_t é ruído branco e Y_t é a observação no tempo t . Os valores de p e P determinam o grau dos polinômios auto-regressivos da parte regular $\phi(B)$ e da parte sazonal $\Phi(B^S)$, respectivamente. De forma análoga, os valores de q e Q determinam o grau dos polinômios de médias-móveis da parte regular $\theta(B)$ e da parte sazonal $\Theta(B^S)$, respectivamente. Semelhantemente, d e D determinam quantas vezes a série temporal deve ser diferenciada na parte regular e na parte sazonal, respectivamente, para que se torne estacionária. Os parâmetros do modelo que devem ser estimados são ϕ , θ , Φ e Θ . Os valores de p , d , q , S , P , D e Q devem ser determinados na modelagem e não estimados **Erro! Fonte de referência não encontrada.**

Neste tipo de modelagem é possível fazer previsões para vários meses com erros muito pequenos. Foi feita previsão para os próximos doze meses a partir do período considerado. Foram reservados seis meses dos últimos valores dos dados para avaliar previsão. Esta avaliação foi feita de duas formas. Na primeira foi feita previsão para seis meses (intervalo) e na segunda foi feita previsão para os mesmos seis meses, porém sempre atualizando os dados para prever o mês seguinte (pontual).

Para modelagem com este método foi utilizado o software livre R **Erro! Fonte de referência não encontrada.** Neste software há pacotes específicos para este tipo de modelagem. Para seleção e validação dos modelos escolhidos para cada grupo, foram utilizadas técnicas descritas em **Erro! Fonte de referência não encontrada.** e 0.

No ajuste do grupo A foi considerado o período de julho de 2003 até agosto de 2007.

Para o grupo A Paulista foi proposto o modelo SARIMA(0,1,1)12(1,1,0), dado pela equação

$$Y_t = Y_{t-1} + 0,5843504 \cdot Y_{t-12} - 0,5843504 \cdot Y_{t-13} + 0,5843504 \cdot Y_{t-24} - 0,4156496 \cdot Y_{t-25} \\ + a_t - 0,3221046 \cdot a_{t-1}.$$

A Figura 3 mostra o ajuste dos dados do grupo A da CPFL Paulista.

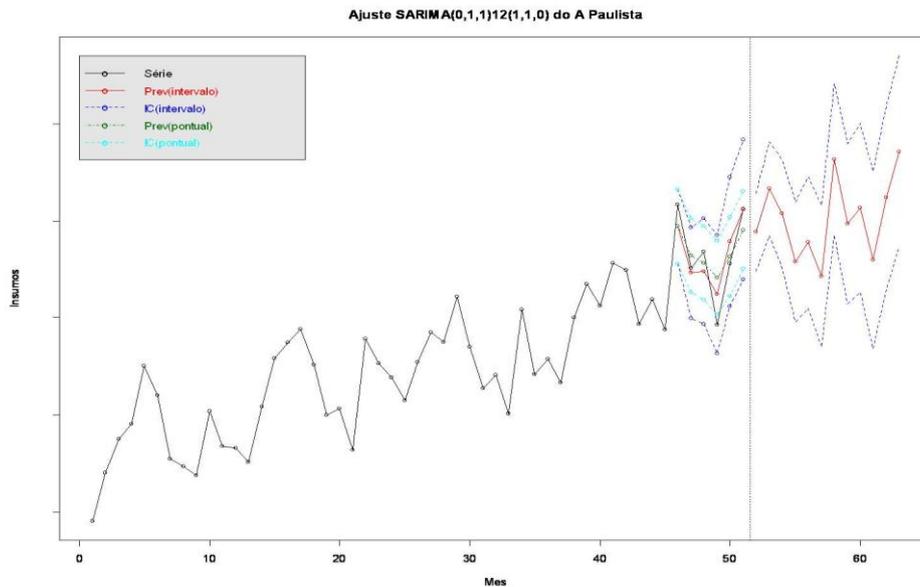


Figura 3 – Grupo A CPFL Paulista (SARIMA)

Para o grupo A Piratininga foi proposto o modelo SARIMA(2,1,0)12(0,1,0) dado pela equação

$$Y_t = 0,6431763 \cdot Y_{t-1} - 0,2575739 \cdot Y_{t-2} + 0,6143976 \cdot Y_{t-3} + Y_{t-12} - 0,6431763 \cdot Y_{t-13} - 0,2575739 \cdot Y_{t-14} - 0,6143976 \cdot Y_{t-15} + a_t.$$

A Figura 4 mostra o ajuste dos dados do grupo A da CPFL Piratininga.

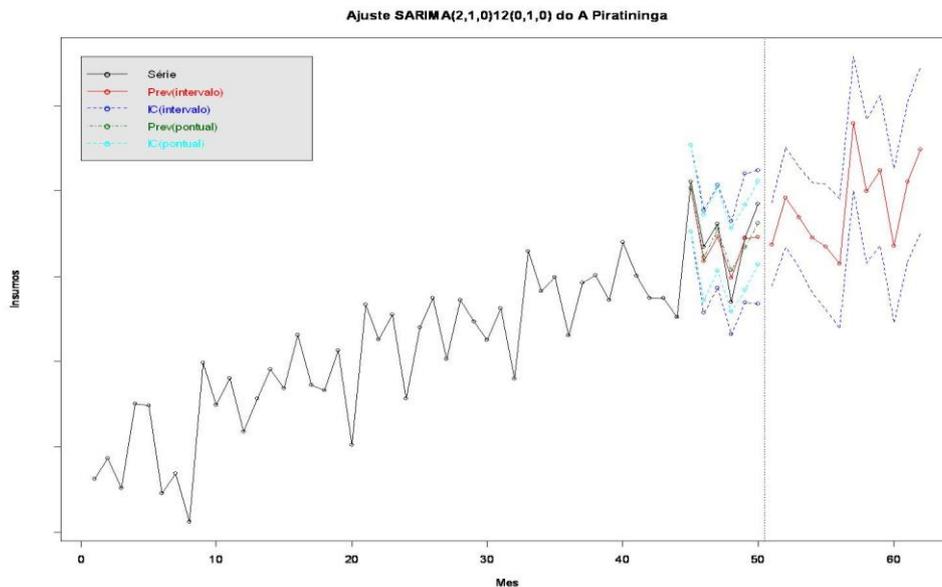


Figura 4 – Grupo A – CPFL Piratininga (SARIMA)

2.3. Método RNA

As Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no cérebro humano, tentando reconhecer regularidades e padrões nos dados. São capazes de aprender com a experiência e fazer generalizações baseadas no seu conhecimento previamente acumulado **Erro! Fonte de referência não encontrada.** As RNA são utilizadas em diversas áreas de aplicação e neste trabalho é utilizada para a previsão de consumo de energia elétrica. As entradas da Rede Neural para a previsão foram baseadas em **Erro! Fonte de referência não encontrada.** Dada uma série a ser modelada, devem ser realizados alguns procedimentos e análises. Segue a descrição da metodologia adotada.

O primeiro procedimento é a detecção e retirada de tendência da série, caso a mesma não seja estacionária. Na Figura 5 pode-se ver uma série temporal, referente ao consumo do Subgrupo A4 da CPFL Paulista. Percebe-se que há tendência de crescimento ao longo do tempo. Esta tendência aparentemente é linear, ou seja, pode ser descrita por uma reta. Para verificar se esta tendência é mesmo linear, calcula-se a diferença entre a série de consumo e a reta. A nova série, calculada esta diferença, é apresentada na Figura 6. Verifica-se, também, que a nova série tem comportamento igual à série original (com exceção da tendência), preservando a sazonalidade.

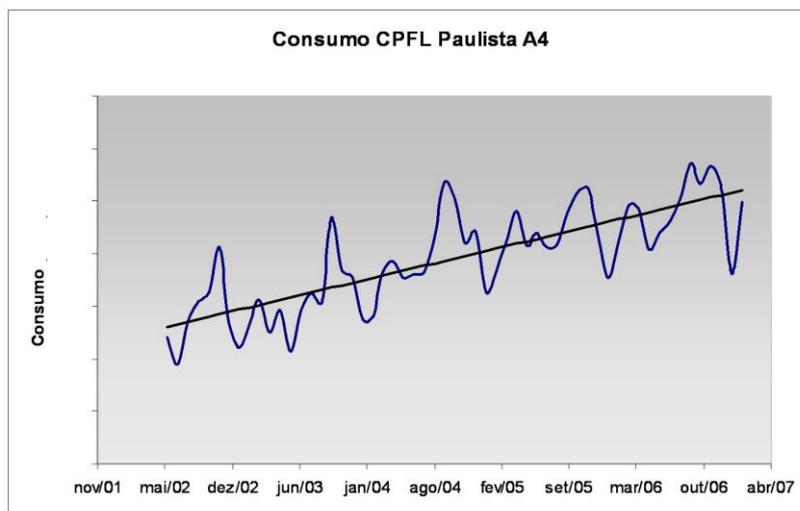


Figura 5 - Série temporal com tendência

Após a retirada da tendência da série, o próximo passo é a detecção da sazonalidade da mesma. Pelas Figuras 5 e 6, percebe-se que há sazonalidade nas séries e que aparentemente tem período de 12 meses (anual). Em algumas séries é mais difícil esta detecção e, então, a função de auto-correlação 0 é de grande ajuda. A Figura 7 apresenta a função de auto-correlação da série sem tendência. O padrão de seqüência de valores negativos e positivos indica que a série é sazonal. O período é indicado pelas observações que são significativas, ou seja, que estão fora do intervalo de aceitação. Na figura, verifica-se que, começado do primeiro valor significativo positivo, a cada 12 valores de correlação, o valor é positivo e significativo. O mesmo ocorre para os valores negativos. Isto é um indício de que a sazonalidade da série é anual.

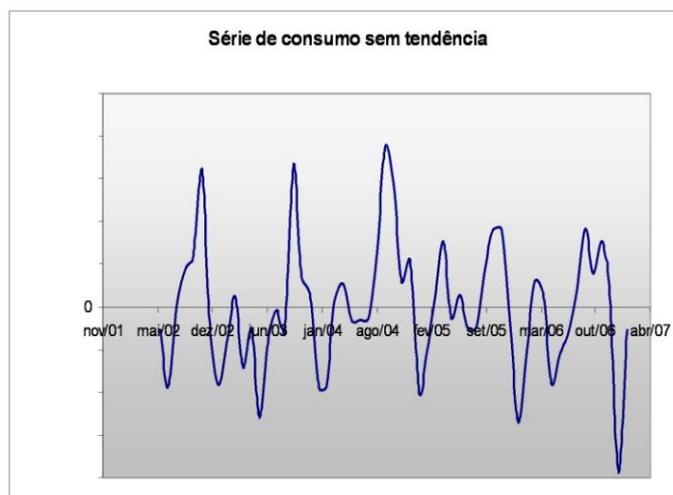


Figura 6 - Série temporal estacionária

É mais simples a modelagem de dados estacionários do que com tendência, por isto são retiradas as tendências.

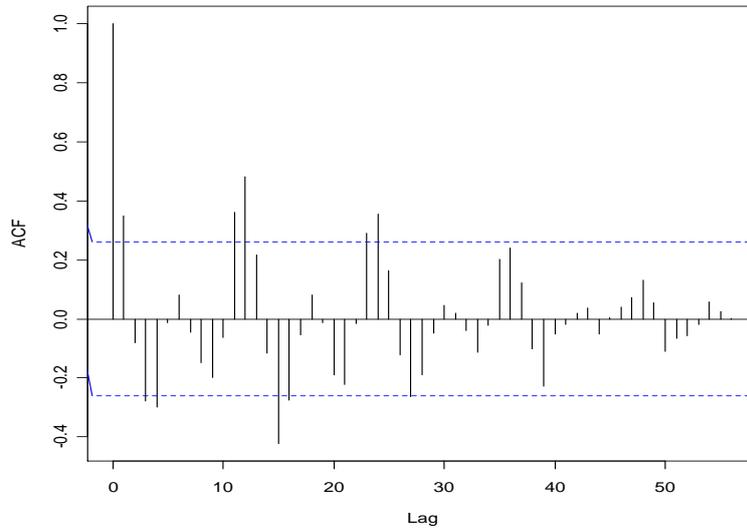


Figura 7 - Função de auto-correlação da série sem tendência

Com relação a sazonalidade, a inserção de dados que indiquem a sazonalidade de uma série facilita o aprendizado da Rede Neural. Desta forma, adotou-se utilizar como entrada da Rede Neural, um vetor com variáveis binárias. Suponha, por exemplo, que uma série tenha sazonalidade semestral. O vetor que indica a sazonalidade tem 6 entradas, da seguinte forma:

[jan/jul fev/ago mar/set abr/out mai/nov jun/dez].

Quando a variável de saída da rede for março ou setembro, mar/set será igual a 1 e todas as outras entradas do vetor, iguais a zero. Além da sazonalidade, outro fator importante da função de auto-correlação são as indicações das próprias observações que possuem correlação. Na Figura 9, indica-se que uma observação no tempo t , possui correlação com as observações nos tempos $t - 1$, $t - 3$, $t - 4$, $t - 11$, $t - 12$, e assim por diante.

Foram adotadas como entradas da Rede Neural, combinações destas observações defasadas. A Figura 8 exemplifica a estrutura de uma Rede Neural multicamadas com três camadas (uma de entrada, uma intermediária e uma de saída). Na figura estão indicadas as entradas e a saída da Rede Neural. Nas entradas temos variáveis de sazonalidade, como no exemplo anterior, e observações defasadas consideradas relevantes analisando o gráfico da função de autocorrelação.

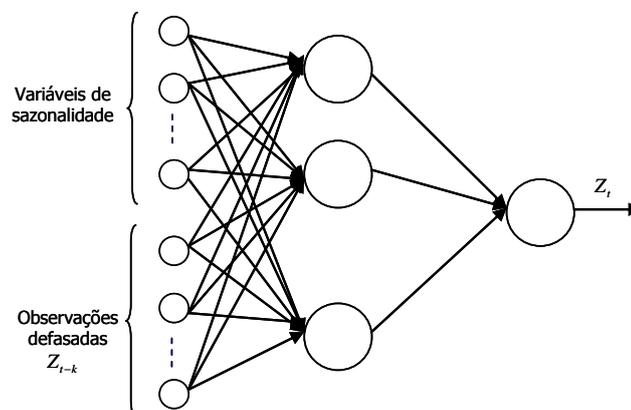


Figura 8 - Entradas e saída da Rede Neural de previsão

Um conjunto de dados é utilizado para o treino da Rede Neural e, então, o restante dos dados é utilizado para a validação. As entradas da rede, como já mencionado, são as variáveis que indicam sazonalidade e observações defasadas que possuem correlação com a observação que se deseja que a rede indique como saída.

Foi feita a previsão do consumo de energia elétrica estimado por mês civil, do grupo A CPFL Paulista e CPFL Piratininga. Foi utilizado o software Clementine **Erro! Fonte de referência não encontrada.** para as previsões com Redes Neurais. O período considerado foi de janeiro de 2004 a agosto de 2007, desenvolvendo dois modelos de RNA.

O primeiro modelo faz a previsão do próximo mês a partir do período considerado, ou seja, considerando o consumo entre janeiro de 2004 a fevereiro de 2007 como conjunto de treinamento, fez-se a previsão do consumo da energia elétrica para o mês de março de 2007. Posteriormente, utiliza-se o período de janeiro de 2004 a março de 2007 como conjunto de treinamento e estima-se o consumo de abril de 2007, assim sucessivamente até a previsão de agosto de 2007.

O segundo modelo faz a projeção do consumo de energia elétrica para seis meses. Para este modelo, consideram-se os consumos de janeiro de 2004 a fevereiro de 2007 como dados de treinamento para a RNA e estima-se o consumo de março de 2007 a agosto de 2007.

A Figura 9 mostra o ajuste da série de consumo da energia elétrica do grupo A CPFL Paulista no período de janeiro de 2004 a fevereiro de 2007, prevendo os meses de março a agosto de 2007.

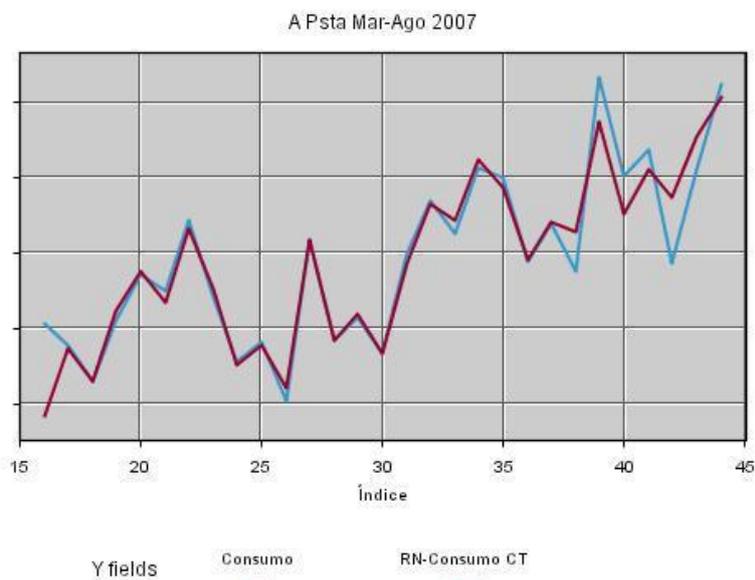


Figura 9 - Previsão dos meses de março a agosto de 2007 do Grupo A CPFL Paulista

No gráfico da Figura 10 tem-se o ajuste com a previsão para o mês de agosto de 2007, considerando a série de consumo no período de janeiro de 2004 a julho de 2007.

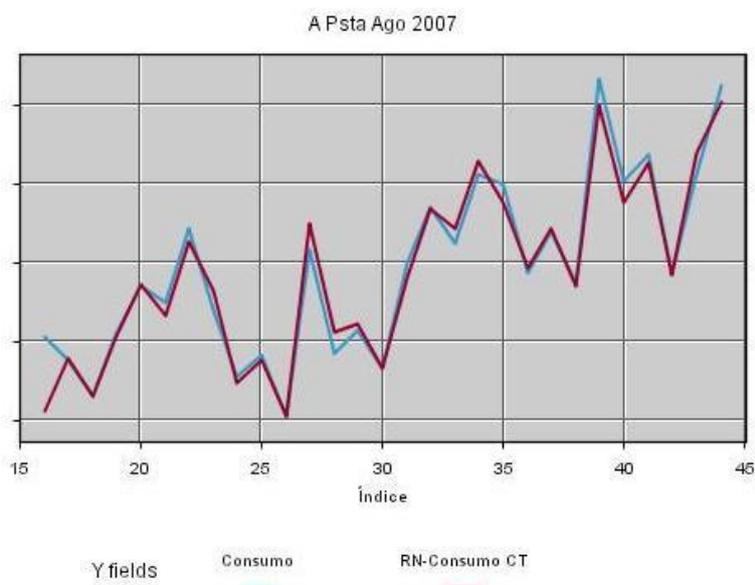


Figura 10 - Previsão para o mês de agosto de 2007 do grupo A Paulista.

3. CONCLUSÕES

A aplicação de dois métodos distintos, preditivo e analítico (Sarima e Redes Neurais artificiais) ao invés da escolha de uma única direção de desenvolvimento foi também um grande mérito do desenvolvimento realizado. Os métodos utilizados para previsão, assim como estas técnicas, possuem um erro associado que, segundo resultados obtidos no primeiro ano do projeto, têm uma variação menor que a do “float”, método atual da concessionária. Este considera, como Energia Fornecida a diferença das quantidades lidas entre duas medições de cada consumidor, geralmente não coincidindo com início e fim do mês civil.

Foram realizadas análises sobre toda a base (5 milhões de consumidores) das concessionárias CPFL Paulista e Piratininga, demonstrando que os métodos utilizados apresentaram resultados satisfatórios, com erros percentuais em geral, em torno de 1 a 2 por cento.

Para avaliação e comparação dos resultados dos dois métodos de previsão de consumo de energia, SARIMA e RNA, foram calculados os erros percentuais absolutos (EP), dados pela equação

$$|EP| = \left| \frac{(R - P)}{R} \cdot 100 \right|.$$

Na Tabela 2 estão os valores calculados para os grupos A CPFL Paulista e CPFL Piratininga.

Tabela 2 – Erros percentuais absolutos do grupo A

Meses		Ago	Jul	Jun	Mai	Abr	Mar	Mar-Ago
Grupo A Paulista	RNA	0,981	2,003	4,302	1,251	3,867	2,722	2,260
	SARIMA	1,946	0,697	4,879	1,065	1,234	2,034	1,627
Grupo A Piratininga	RNA	0,081	1,208	1,431	1,064	1,444	1,008	1,103
	SARIMA	0,956	0,843	1,650	3,669	2,201	3,963	1,617

Na Tabela 3 estão os valores calculados para os grupos B Paulista e Piratininga. Em geral estes valores estão, em média, entre um e dois por cento, indicando que estes métodos de previsão são bastante adequados para previsão do consumo de energia elétrica. Também se pode notar que para CPFL Paulista, o melhor método é o SARIMA e para a CPFL Piratininga, o melhor é o RNA.

Utilizando os valores previstos para o cálculo da energia não faturada, obtêm-se as curvas dos valores mensais do *float* (diferença entre a energia não faturada do mês atual com a do mês anterior) e dos erros percentuais absolutos dos métodos SARIMA e RNA. Estes valores são obtidos considerando seis meses de dados reais que foram reservados para avaliar as previsões. O que se pode notar é que a variação dos métodos de previsão é menor que a variação do *float*. No gráfico da Figura 11 estão as curvas do *float* e dos erros percentuais absolutos para cada grupo.

Tabela 3 – Erros percentuais absolutos do grupo B

Tabela de erros percentuais absolutos de previsão EP								
Meses		Jul	Jun	Mai	Abr	Mar	Fev	Fev-Jul
Grupo B Paulista	RNA	1,258	0,487	2,703	5,485	4,863	0,810	2,988
	SARIMA	0,263	1,780	2,504	1,542	2,270	1,867	1,475
Grupo B Piratininga	RNA	0,126	1,636	0,174	3,642	1,904	2,525	1,680
	SARIMA	3,047	1,525	3,696	0,330	3,888	3,970	2,742

Assim, pode-se concluir que o método de cálculo do balanço energético (Figura 1) será mais preciso que o método atual. Esta melhoria na precisão se dá pelo fato de substituir um valor desconhecido, *float*, por um valor previsto, sendo que esta previsão tem uma eficiência relativamente boa. Desta forma, a ferramenta computacional a ser construída, fará de forma automatizada o cálculo do balanço energético utilizando os valores de consumo de energia elétrica por mês civil e o valor previsto de consumo para o mês atual.

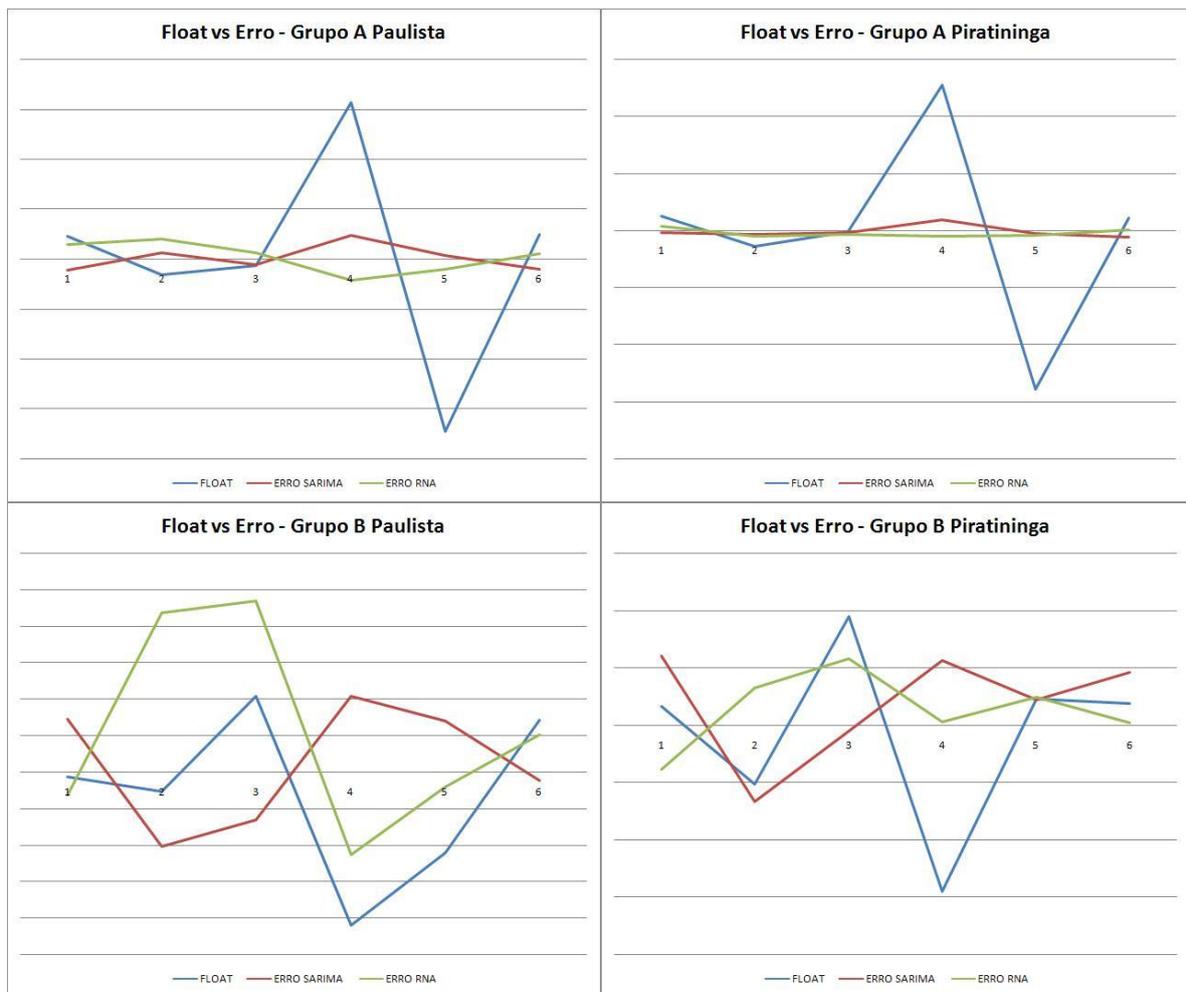


Figura 11 – Variação do float e dos erros percentuais absolutos

4. Referências bibliográficas

- [1] F. N. Fritsch; R. E. Carlson. Monotone Piecewise Cubic Interpolation, SIAM Journal on Numerical Analysis, 17 (1980), pp. 238-246.
- [2] Fernandes, L. G. L.; M. S. Portugal; P. O. Navaux. A. Previsão de séries de tempo: Redes Neurais Artificiais e modelos estruturais. 1995. Disponível em: http://www.ufrgs.br/ppge/pcientifica/1995_09.pdf. Acesso em abril de 2007.
- [3] Brockwell, Peter J.; Davis, Richard A. Introduction to time series and forecasting, 2nd ed. p. cm- (Springer texts in statistics), USA, 2002.
- [4] Morettin, Pedro A.; Toloi, Clélia M. Análise de séries temporais – 2ª ed. Edgard Blucher, São Paulo, 2006.
- [5] R 2.6.0 - R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [6] Clementine 8.5 – Copyright © Integral Solutions Ltd., 1994 – 2004. All rights reserved. www.spss.com/clementine
- [7] Calôba, G. M.; L. P. Calôba. E. Saliby. Cooperação entre Redes Neurais Artificiais e técnicas ‘clássicas’ para previsão de demanda de uma série de vendas de cerveja na Austrália. Em Pesquisa Operacional, v.22, n.3, 2002.