

# MODELAGEM DA DEMANDA RESIDENCIAL DE ENERGIA ELÉTRICA POR MEIO DE REDES NEURAIS E ALGORITMOS GENÉTICOS

SAMUEL B. DEFILLIPO

*Programa de Mestrado em Modelagem Computacional, UFJF*  
*Campus Universitário - CEP 36036 330 - Juiz de Fora, MG*  
*E-mail: samueldefilippo@gmail.com*

HENRIQUE S. HIPPERT

*Depto. de Estatística / ICE, UFJF*  
*Campus Universitário - CEP 36036 330 - Juiz de Fora, MG*  
*E-mail: henrique.hippert@ufjf.edu.br*

**Abstract.** This paper describes the application of both statistical techniques and computer intensive methods to the modelling of monthly residential electrical load of a household. The input data are the household's characteristics (area, number of rooms, etc.), its location (de-fined by metric co-ordinates), and the socio-economic profile of its residents. We experimented with linear regression models, with neural networks (feed-forward multilayer perceptrons) defined by means of genetic algorithms, and with a model with two components: one that modelled the characteristics of the household and the profile of its residents by linear regression, and another that modelled the location effect by means of a small neural network. Such a model may prove useful as a tool for designing a typology that associates the consumer's profile to their most probable level of electricity demand, and serves as input for a load forecasting model, particularly for relatively small areas where the socio-economic characteristics of the population are known.

**Keywords**— neural networks, genetic algorithms, load forecasting, spatial data

**Resumo**— Este artigo descreve a aplicação de técnicas estatísticas e de computação intensiva à modelagem do consumo mensal de energia elétrica de uma residência, a partir de dados sobre as características físicas desta residência (área construída, número de cômodos, etc.), de sua localização (definida por meio de coordenadas), e das características sócio-econômicas de seus moradores. Foram experimentados modelos de regressão linear, modelos de redes neurais (perceptrons feedforward multicamadas) cuja arquitetura foi definida por meio de simulações usando algoritmos genéticos, e um modelo com dois componentes: uma regressão linear que modelava as características da residência e de seus moradores, e uma rede neural que modelava o efeito de localização. Um modelo assim desenvolvido irá servir de base para a criação de uma tipologia que associe o perfil de cada consumidor ao seu nível de consumo mais provável, e forneça subsídios para a previsão de carga, especialmente em áreas relativamente restritas onde as características sócio-econômicas da população sejam conhecidas.

**Palavras-chave**— redes neurais, algoritmos genéticos, previsão de carga, dados geo-referenciados

## 1 Introdução

O setor elétrico brasileiro tem passado nos últimos anos por transformações profundas causadas pela substituição de um modelo inteiramente centralizado e monopolista por um modelo onde a geração e a distribuição foram privatizadas. Esta substituição tem se processado gradualmente, e uma das últimas mudanças foi a centralização da aquisição da energia, realizada em 2004. Neste ano, foi criada a *Câmara de Comercialização de Energia Elétrica*, responsável por todas as compras da energia gerada; as distribuidoras, desde então, se limitam a informar a esta câmara as cargas que prevêem necessárias nas suas áreas de concessão, para os próximos cinco anos.

Esta modificação aumentou a importância da previsão de demanda para as distribuidoras, porque os erros passaram a ser penalizados por multas. Métodos para a previsão de demanda de áreas relativamente pequenas passaram a ser necessários, principalmente aqueles que levam em conta as característi-

cas sócio-econômicas da população abrangida pela área de concessão da distribuidora.

O trabalho descrito neste artigo faz parte de um projeto de estudo de tais métodos: o objetivo é criar modelos que associem informações geo-referenciadas sobre o perfil sócio-econômico dos consumidores e as características das residências a seu consumo provável de energia. A partir destes modelos, será possível criar uma tipologia de demanda que deverá ser utilizada na previsão da carga elétrica, especialmente para a previsão a médio ou longo prazo de demanda em áreas restritas, cujas características populacionais sejam conhecidas.

Os métodos estudados serão técnicas de computação intensiva, como as redes neurais artificiais e os algoritmos genéticos, e a combinação destas técnicas com métodos tradicionais como a regressão linear. Em particular, procuramos tirar proveito do aspecto geo-referenciado dos dados. Como no banco de dados disponível os valores de demanda estão associados às coordenadas que localizam cada domicílio avaliado, foi possível verificar se a localização geográfica de um domicílio era uma das características que influenciam no seu consumo de energia. Técni-

cas de análise de dados geográficos ou georeferenciados têm geralmente sido pouco exploradas em pesquisas sobre consumo de energia (e.g., Chow & Tram, 1997; Willis *et al.*, 1995), devido ao alto custo da obtenção destes dados, em comparação ao da obtenção de dados agregados. Uma versão preliminar deste trabalho foi apresentada em Hippert (2006).

## 2 Dados usados para análise

O banco de dados utilizado resultou de um levantamento feito pelo Departamento de Estatística da UFJF em 2004-2005. Um total de 557 residências foram escolhidas por meio de um processo de amostragem em várias etapas, e os chefes das famílias responderam questões sobre as características de cada residência (área construída, número de cômodos, etc.), as características sócio-econômicas de seus moradores (escolaridade, número de carros, número de empregadas, etc.), aos hábitos de consumo dos moradores (frequência e horário em que eram utilizados os diversos equipamentos elétricos existentes) e

às atitudes frente a políticas de conservação de energia e de racionalização do consumo. Além disso, foram registradas as coordenadas em metros (norte e leste) da residência, por meio de um aparelho de Global Positioning System (GPS).

Para este estudo, não utilizamos as variáveis relativas aos hábitos de consumo de energia e à conservação e racionalização do consumo; mantivemos apenas as variáveis que descreviam as características físicas da residência e o perfil sócio-econômico dos moradores. A Tabela 1 mostra as variáveis usadas, indicando também a amplitude dos valores assumidos pelos dados. Como havia muitos valores faltantes entre as variáveis que mediam o consumo mensal, procuramos identificar os três meses consecutivos em que havia maior quantidade de dados disponíveis. Adotamos então como variável indicadora do consumo a média entre estes três meses (março a maio de 2004). A seguir, eliminamos todos os casos com valores faltantes, com erros óbvios ou com registros duvidosos em qualquer variável; restaram 444 casos para análise.

**Tabela 1 – variáveis disponíveis para estudo**

variáveis	valores assumidos	descrição
1grau, 2grau, 3grau	0 : 1	grau de instrução do chefe da família:
monofas, bifasic, trifasic	0 : 1	padrão de abastecimento:
empregd	0 : 1	tem empregada doméstica?
carros	0 : 5	número de carros no domicílio
tipo	1 : 3	tipo da residência
area	1 : 6	área construída
comodos	2 : 20	número de cômodos na residência
banheiro	1 : 7	número de banheiros na residência
resident	1 : 17	número de residentes no domicílio
consumo	28 : 666	consumo, em kWh
coord_N	7588406 : 7601199	coordenada Norte, em metros
coord_E	661042 : 673269	coordenada Este, em metros

## 3. Modelos e Resultados

### 3.1. Modelo de regressão linear

O objetivo dos experimentos com modelos de regressão linear relacionando as variáveis com o consumo foi, primordialmente, o de fornecer um *benchmark* ao qual o desempenho das redes neurais deveria ser comparado. A primeira tarefa foi a de identificar as variáveis que poderiam estar associadas ao consumo de energia elétrica.

O consumo médio foi significativamente diferente entre os grupos definidos pelo padrão de abastecimento (**trifasic** ou **monofas**) e pela existência ou não de empregada doméstica (**empregd**); em ambas comparações,  $p = 0.000$  no teste-t. A ANOVA foi usada para verificar a associação entre o consumo e as variáveis nominais ou ordinais com mais de duas categorias. O consumo foi significativamente diferen-

te ( $p=0,000$ ) entre os grupos definidos pela variável **area** (área construída, classificada em seis níveis, de “0-50 m<sup>2</sup>” até “mais de 200 m<sup>2</sup>”) e pelo grau de instrução do chefe da família (representados pelos *dummies* **1grau**, **2grau** e **3grau**). Por outro lado, não houve diferença significativa ( $p = 0,255$ ) entre o consumo dos dois tipos de residência, casas ou apartamentos (variável **tipo**).

Algumas das variáveis numéricas existentes no banco de dados mostraram uma correlação pequena, mas significativa ( $p = 0,000$ ), com o consumo: **carros** ( $r = 0,362$ ), **comodos** ( $r = 0,375$ ), **banheiro** ( $r = 0,458$ ) e **resident** ( $r = 0,394$ ). As coordenadas geográficas, no entanto, praticamente não mostraram correlação: **coord\_N** ( $r = -0,156$ ), **coord\_E** ( $r = 0,045$ ); nos modelos de regressão linear, elas foram excluídas pelo procedimento *stepwise*. Esta falta de correlação era esperada, uma vez que a relação probabilística entre as coordenadas e o consumo, se exis-

tir, deverá ser altamente não-linear, o que justifica a tentativa de modelagem por meio de redes neurais artificiais.

Em primeiro lugar, experimentamos modelos de regressão linear que incluíam todas as variáveis cuja associação com a variável dependente, **consumo**, foi verificada acima. As variáveis cujos coeficientes resultaram significativos foram **empregd**, **3grau**, **carros**, **monofas**, **trifasic** e **resident**. As variáveis restantes, **1grau**, **area**, **comodos** e **banheiro**, tiveram coeficientes não-significativos. Em segundo, lugar experimentamos aplicar um procedimento *stepwise* de inclusão e eliminação de variáveis, o que resultou num modelo contendo as mesmas seis variáveis cujos coeficientes tinham sido considerados significativos anteriormente. Os valores do  $R^2$  ajustado para estes dois modelos são bastantes semelhantes; para o modelo que incluía todas as 10 variáveis,  $R^2 = 0,4965$ ; para o modelo com seis variáveis escolhidas por procedimento *stepwise*,  $R^2 = 0,4891$  (em ambos,  $p = 0,000$ .)

### 3.2. Modelos de redes neurais (RNs)

Os modelos de rede neural empregados a seguir são baseados em percéptrons com uma camada oculta. O treinamento foi interrompido por um procedimento de parada por validação cruzada (Haykin, 1999, p. 213). Para a definição do número de neurônios ocultos, da taxa de aprendizado e das variáveis de entrada, recorremos a simulações por meio de algoritmos genéticos (Dréo et al, 2005).

Os algoritmos genéticos (AG) usados foram de forma *steady-state*, com a seleção ocorrendo através de torneios. Escolhíamos aleatoriamente dois indiví-

duos da população e tomávamos o melhor; depois, escolhíamos aleatoriamente outros dois indivíduos e também tomávamos o melhor. Verificávamos se seria feito *cross-over*; em seguida, se seria aplicado o operador de mutação, de acordo com probabilidades definidas. Por fim, calculávamos o valor da aptidão do indivíduo final. Se ele fosse melhor que o pior dos quatro indivíduos que participaram do torneio, ele era incluído na população, substituindo o pior indivíduo dos quatro.

Para o cálculo da aptidão, utilizamos como critério o erro médio percentual absoluto (*mean absolute percent error* - MAPE), definido pela Eq. (1), onde  $\hat{y}_i$  é o valor previsto,  $y_i$  é o valor observado. O MAPE tem-se tornado atualmente um padrão de medida de erro, nas pesquisas relacionadas ao consumo de energia, na engenharia elétrica.

$$MAPE = 100 * \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (1)$$

Para a modelagem, os 444 casos disponíveis foram divididos aleatoriamente em três amostras: 364 casos para treinamento da rede (isto é, para a estimação de seus parâmetros), 30 para a validação cruzada, e 50 para o teste *out-of-sample* da rede ajustada.

Foram definidos 11 modelos de RNs (por meio de quatro configurações diferentes do AG), que foram então treinados 30 vezes cada. As características destas RNs, e os MAPEs produzidos por elas nos 30 treinamentos, são exibidos na Tabela 2.

**Tabela 2 - Resultados de 30 rodadas dos modelos propostos**

modelo	população	gerações	probabilidades		neurônios	MAPE		
			mutação	<i>crossover</i>		médio	máximo	d-padrão
AG1/1	30	10.000	0,01	0,80	3	32,8	49	4,6
/2					12	28,4	48	5,4
/3					58	35,0	181	29,9
/4					52	43,9	157	30,7
AG 2/1	80	20.000	0,01	0,80	12	27,3	43	4,1
/2					6	33,0	49	4,7
AG 3/1	30	10.000	0,05	0,80	61	24,9	49	5,0
/2					17	27,8	75	9,8
AG 4/1	80	20.000	0,05	0,80	25	34,2	151	24,2
/2					19	26,2	38	4,5
/3					45	38,5	130	26,3
modelo combinado RN + regressão linear						27,6	29,6	1,0
regressão linear						30,6	-	-

### 3.3. Modelo combinado RN e regressão linear

O terceiro modelo experimentado foi uma combinação de um modelo de regressão linear com uma RN. O objetivo foi verificar se seria possível usar as variáveis de geo-referenciamento (coord\_N e coord\_E) para explicar os desvios do consumo medido em cada residência em relação a um padrão (no caso, definido por um modelo de regressão linear). A carga L será prevista como a soma de uma previsão linear,

fornecida pelo modelo de regressão, com uma parcela fornecida por uma RN. O modelo de regressão fornece uma previsão básica; o segundo componente, fornecido pela RN, explica os desvios observados em relação ao padrão linear como função das coordenadas. A RN adotada seguiu o mesmo padrão dos modelos anteriores, mas as entradas foram apenas as duas coordenadas. A Figura 1 mostra um esquema deste modelo combinado.

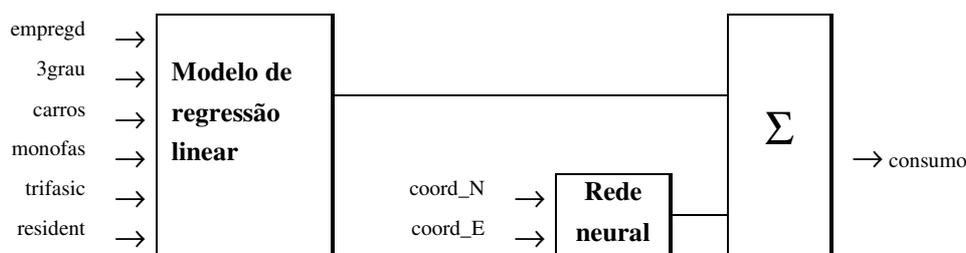


Figura 1. Modelo combinando regressão linear e rede neural

## 4. Discussão e conclusão

Os resultados exibidos na Tabela 2 mostram que houve uma variação muito grande entre as RNs especificadas por cada rodada do AG, tanto no número de neurônios, quanto no resultado médio alcançado (MAPE médio). Das 11 RNs testadas, a maioria foi pior do que o modelo combinado RN+regressão, e mesmo do que um modelo simples de regressão linear (apenas três RNs obtiveram MAPEs médios melhores do que o do modelo combinado, e apenas cinco melhores a regressão linear). A Figura 2, contudo, nos permite uma visão um pouco diferente destes resultados. O gráfico mostra que a maioria das RNs (sete em 11) obtiveram resultados *medianos* melhores do que

o do modelo RN+regressão, e quase todas foram melhores, em mediana, do que a regressão linear (nove em 11).

(Para fazer este gráfico, comparando por diagramas de Tukey (*boxplots*) os MAPEs de cada modelo, foi preciso eliminar os valores discrepantes encontrados, a fim de que a escala vertical fosse adequada. Como pode ser visto na última coluna da Tabela 2, quase todas as RNs tiveram MAPEs máximos muito grandes, chegando até a 181%. Excluímos do gráfico todos os pontos discrepantes com MAPEs superiores a 80%; a existência destes pontos é denotada no gráfico por uma seta vertical colocada acima do diagrama correspondente.)

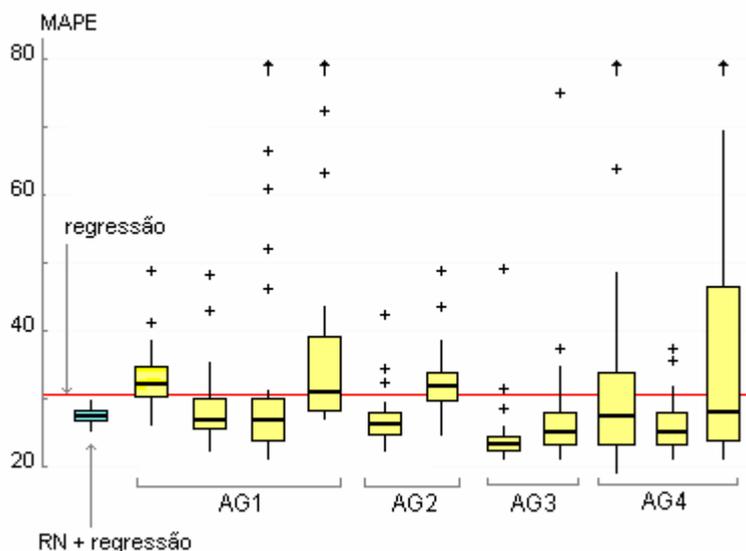


Figura 2. MAPEs obtidos por cada modelo, em 30 rodadas (as setas denotam a existência de um ou mais pontos discrepantes, acima de 80 %)

Os diagramas mostram a razão da diferença de interpretação entre valores médios e medianos: todas as RNs obtiveram às vezes erros muito altos, que se destacam como pontos discrepantes no gráfico. Estes pontos discrepantes tiveram o efeito de inflacionar os valores das médias e dos desvios-padrões das RNs, tornando-os em geral mais altos do que os do modelo RN+regressão, que não teve pontos discrepantes.

É importante lembrar que cada diagrama representa os MAPEs obtidos por uma mesma RN, sobre uma mesma amostra de teste, em simulações feitas a partir de valores iniciais diferentes. Embora algumas RNs tenham conseguido resultados médios ou medianos muito bons em relação à regressão linear, estas mesmas redes mostraram uma dispersão muito grande nos resultados, e obtiveram algumas vezes erros excessivamente altos. Para os objetivos deste trabalho, que é o de definir uma tipologia, que sirva de *input* para a previsão de demanda a médio e longo prazo (ao lado de outras variáveis como as séries de cargas passadas e indicadores da atividade econômica), estes erros são inaceitáveis, pois poderiam tornar instável o modelo de previsão usado. A combinação RN+regressão, por outro lado, embora fosse um pouco inferior em mediana a várias das RNs, ainda foi sempre melhor do que um modelo de regressão linear, mostrou menor dispersão que todos as RNs, e não apresentou valores discrepantes. Estas características fazem com que os resultados deste modelo talvez possam ser mais adequados do que os das RNs estudadas, para uma utilização como entrada de um modelo de previsão.

Contudo, o fato de algumas redes terem conseguido resultados pontuais muito bons (e.g. AG 4/1, cujo melhor resultado foi de 19 %), nos leva a crer que existem boas possibilidades de aplicação de redes neurais neste tipo de problema; o que é preciso é encontrar arquiteturas que mostrem menos sensibilidade aos valores iniciais. Em prosseguimento a este trabalho, pretendemos estudar outras metodologias para a definição das características das RNs, tanto em relação à suas arquiteturas quanto em relação à seleção de variáveis de entrada, quanto modificações no treinamento, de forma evitar a parada das buscas em mínimos locais.

### Agradecimentos

Parte deste trabalho foi apoiada pela FAPEMIG – Fundação de Amparo a Pesquisa de Minas Gerais, através do Edital 005 / 2004 (Projeto no. EDT 164 / 05).

### Referências Bibliográficas

- Bishop, C. M. (1998), *Neural Networks for Pattern Recognition*, Oxford University Press.
- Chow, M.-Y. & Tram, H. (1997), Application of fuzzy logic technology for spatial load forecasting, *IEEE Trans Power Systems*, v .12, n. 3, pp. 1360-1366.
- Dréo, J.; Pétrowski, A.; Siarry, P.; Taillard, E. (2005), *Metaheuristics for Hard Optimization*. Springer.
- Haykin, S. (1999), *Neural Networks – a Comprehensive Foundation*. 2<sup>nd</sup> ed., Prentice Hall, Upper Saddle River, NJ.
- Hippert, H. S. (2006), Modelagem e previsão de demanda de energia elétrica com base em dados geo-referenciados. *Anais do IX Encontro de Modelagem Computacional*, Belo Horizonte.
- Willis, H. L., Engel, M. V. & Buri, M. J. (1995), Spatial load forecasting, *IEEE Computer Applications in Power*, April, pp. 40-43