

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AMBIENTAL**  
**MESTRADO EM ENGENHARIA AMBIENTAL**  
**MODALIDADE PROFISSIONAL**

**PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA PARA  
COMUNIDADES ISOLADAS BASEADA EM REDES NEURAIIS**

**ROBERTA FERNANDES MENDIONDO NUNES**

**MACAÉ/RJ**

**2012**

**ROBERTA FERNANDES MENDIONDO NUNES**

**PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA PARA  
COMUNIDADES ISOLADAS BASEADA EM REDES NEURAIAS**

**Dissertação de Mestrado em Engenharia Ambiental**

**Mestrado em Engenharia Ambiental**

**Área de Concentração Desenvolvimento Regional**

**Orientador D.Sc Jader Lugon Junior,**

**MACAÉ, RJ**

**2012**

## FICHA CATALOGRÁFICA

MENDIONDO, Roberta F. N.  
Previsão de Consumo de Energia Elétrica para  
Comunidades Isoladas Baseada em Redes Neurais. 96 pg.

Dissertação (Mestrado) – Programa de Pós-Graduação em  
Engenharia Ambiental. Instituto Federal Fluminense.

1. Previsão de Consumo de energia elétrica
2. Sistemas Individuais de Geração de Energia
4. Desenvolvimento socioeconômico
5. Redes Neurais

Dissertação intitulada PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA PARA COMUNIDADES ISOLADAS BASEADA EM REDES NEURAIS elaborada por ROBERTA FERNANDES MENDIONDO NUNES e apresentada publicamente perante a Banca Examinadora, como requisito para obtenção parcial para obtenção do título de Mestre em Engenharia Ambiental pelo Programa de Pós-graduação em Engenharia Ambiental, na área de concentração Sustentabilidade Regional, linha de pesquisa Desenvolvimento Regional, do Instituto Federal de Educação, Ciência e Tecnologia Fluminense.

Aprovada em .....

Banca Examinadora:

.....

Jader Lugon Junior, D.Sc., Universidade do Estado do Rio de Janeiro UERJ, Instituto Federal Fluminense IFF, Orientador

.....

Luis Biondi Neto, D.Sc. Universidade Federal do Rio de Janeiro, UFRJ. Universidade do Estado do Rio de Janeiro, UERJ.

.....

Marcos Antonio Cruz Moreira, D.Sc. Universidade Federal do Rio de Janeiro, UFRJ. Instituto Federal Fluminense, IFF.

.....

Maria Inês Paes Ferreira, D.Sc. Universidade Federal do Rio de Janeiro, UFRJ. Instituto Federal Fluminense, IFF.

## **Dedicatória**

Dedico essa dissertação ao meu esposo que tanto me apoiou nessa caminhada, a qual é de extrema importância para nossa família, nesse momento tão especial à espera de nossos filhos, gêmeos, Davi e Pedro, frutos do nosso amor, que é lindo E a meu pai, que na data da defesa está completando 64 anos de idade, de dedicação e amor à família.

## **Agradecimentos**

Agradeço a Deus por ter me sustentado e abençoado em tudo até aqui, especialmente pela saúde e família que me concedeu, a qual me apóia sempre, em especial meu esposo por estar junto nessa caminhada, suportando ausências com amor.

Agradeço ao amor, incentivo e educação recebidos da minha tão especial família, meus pais e irmãos, que fizeram parte da mais importante fase da minha formação, na vivência familiar tão intensa, da infância a fase adulta e da qual ainda desfruto plenamente hoje, acrescida das vidas de meus cunhados e sobrinhos muito amados.

Agradeço a essa Instituição de tão alto grau de excelência, pública e que proporciona condições para que seus alunos possam desenvolver trabalhos extremamente relevantes para o desenvolvimento socioambiental do nosso país.

Agradeço aos professores do PPEA pelos ensinamentos que vão muito além do estudo da ecologia e da engenharia, e dão o tom a cada trabalho produzido nesse programa, em especial, ao meu orientador, professor Jader Lugon Junior pela paciência e ajuda imprescindíveis no caminho desse estudo, bem como aos professores que ministraram disciplinas que revelam o viés multidisciplinar do Programa e tanto contribuíram para a formação do meu problema de estudo e essencialmente da própria preocupação em representar esse viés no desenvolvimento do trabalho, como a professora Maria Inês Paes Ferreira, Rodrigo Serra, Hélio Gomes Filho, Luiz Pinedo e José Augusto Ferreira Silva.

De forma especial agradeço a cada família que participou da pesquisa de campo, sem as quais este trabalho não seria possível.

Agradeço ao corpo docente e administrativo do IFF por lutar a cada dia pela manutenção de uma instituição pública e de excelência.

**Epígrafe**

Se você eliminar do seu meio o jugo opressor, o dedo acusador e a falsidade do falar; se com renúncia própria você beneficiar os famintos e satisfizer o anseio dos aflitos, então a sua luz despontará nas trevas, e a sua noite será como o meio-dia. (Isaías 58:9-10)

**LISTA DE FIGURAS**

Figura 2.2-1 Modelo não-linear de um neurônio, utilizado em MLP. Adaptada (HAYKIN, 2001).....	39
Figura 2.2-2 Representação de uma rede Multi layer perception, com duas camadas ocultas. Elaborado pelo autor.....	40
Figura 2.4-1 Imagens da Comunidade Marujá, região do lagamar, Vale do Ribeira. (MORANTE, 2004). .....	44
Figura 2.4-2 Casa e animais de criação típicos da região da caatinga em Pedra Branca (MORANTE, 2000). .....	45
Figura 2.4-3 Comunidade Vera Cruz (MORANTE, 2004). .....	45
Figura 2.4-4 Ilha Grande, Praia Vermelha (MIRANDA, 2004).....	46
Figura 2.4-5 Correlação entre consumo real por família e a demanda prevista. Elaborada pelo autor. Saída SPSS-IBM.....	52
Figura 2.4-6 Correlação entre valores reais de consumo e demanda prevista, em kWh, por família. Elaborado pelo autor. Saída do SPSS-IBM.....	53



**LISTA DE ILUSTRAÇÕES**

Ilustração A-1 Topologia da RN_1 .....	61
Ilustração A-2 Ajuste Linear de dados de consumo de energia elétrica familiar real e previstos pela RN_1 .....	63
Ilustração A-3 Topologia da RN_2 .....	66
Ilustração A-4 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN_2.....	68
Ilustração A-5 Topologia da RN_3 .....	71
Ilustração A-6 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN_3.....	73
Ilustração A-7 Topologia da RN_4 .....	76
Ilustração A-8 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN_4.....	78
Ilustração A-9 Topologia da RN_5 .....	81
Ilustração A-10 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN_5.....	83
Ilustração A-11 Topologia da RN_6 .....	86
Ilustração A-12 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN_6.....	88
Ilustração A-13 Topologia da RN_7 .....	90
Ilustração A-14 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN_7.....	92

## LISTA DE QUADROS

Quadro 1.1-1 Domicílios particulares premanentes, por existência de energia elétrica, segundo as grandes regiões e unidades da federação (IBGE, 2010).....	19
Quadro 1.6-1 Investimentos em sistemas individuais fotovoltaicos, no âmbito do PRODEEM , de 1998até 2002.....	28
Quadro 1.6-2 Elaborada pelo autor. Fonte: PRODEEM/MME .....	28
Quadro 1.6-3 Padronização dos sistemas individuais pela ANEEL.....	31
Quadro 2.4-1 Estatística dos modelos de regressão com oito, duas e uma variável previsoras. Elaborada pelo autor.....	48
Quadro 2.4-2 Estatística dos modelos de regressão com oito, duas e uma variável previsoras., para amostra de 50 famílias, rurais e urbanas. Elaborada pelo autor. ....	49
Quadro 2.4-3 Estatística dos modelos neurais com oito, duas e uma variável previsoras, para amostra de 50 famílias, rurais e urbanas. Elaborada pelo autor. ....	51
Quadro 2.4-4 Configuração de rede que apresentou os melhores desempenhos para as três amostras submetidas. ....	51
Quadro 2.4-5 Desempenho das redes a partir de amostras particionadas em amostra de treinamento e teste. ....	52

## LISTA DE TABELAS

Tabela A-1 Programação do modelo RN_1 .....	59
Tabela A-2 Particionamento da mostra e configuração da RN_1 .....	60
Tabela A-3 Medidas de desempenho da RN_1 .....	62
Tabela A-4 Pesos sinápticos da RN_1.....	62
Tabela A-5 Importância das variáveis independentes RN_1.....	63
Tabela A-6 Programação do modelo RN_2 .....	64
Tabela A-7 Particionamento da mostra e configuração da RN_2 .....	65
Tabela A-8 Medidas de desempenho da RN_2 .....	67
Tabela A-9 Pesos sinápticos da RN_2.....	67
Tabela A-10 Importância das variáveis independentes RN_2.....	68
Tabela A-11 Programação do modelo RN_3 .....	69
Tabela A-12 Particionamento da amostra e configuração da RN_3 .....	70
Tabela A-13 Medidas de desempenho da RN_3 .....	72
Tabela A-14 Pesos sinápticos da RN_3.....	72
Tabela A-15 Importância das variáveis independentes RN_3.....	73
Tabela A-16 Programação do modelo RN_4 .....	74
Tabela A-17 Particionamento da amostra e configuração da RN_4 .....	75
Tabela A-18 Medidas de desempenho da RN_4 .....	76
Tabela A-19 Pesos sinápticos da RN_4Parameter Estimates.....	77
Tabela A-20 Importância das variáveis independentes RN_4.....	78
Tabela A-21 Programação do modelo RN_5 .....	79
Tabela A-22 Particionamento da amostra e configuração da RN_5 .....	80
Tabela A-23 Medidas de desempenho da RN_5 .....	82
Tabela A-24 Pesos sinápticos da RN_5.....	82
Tabela A-25 Importância das variáveis independentes RN_5.....	83

Tabela A-26	Programação do modelo RN_6 .....	84
Tabela A-27	Particionamento da amostra e configuração da RN_6 .....	85
Tabela A-28	Medidas de desempenho da RN_6 .....	86
Tabela A-29	Pesos sinápticos da RN_6.....	87
Tabela A-30	Importância das variáveis independentes RN_6.....	88
Tabela A-31	Programação do modelo RN_7 .....	89
Tabela A-32	Particionamento da amostra e configuração da RN_7 .....	90
Tabela A-33	Medidas de desempenho da RN_7 .....	91
Tabela A-34	Pesos sinápticos da RN_7.....	91

## SUMÁRIO

LISTA DE FIGURAS .....	VIII
LISTA DE ILUSTRAÇÕES .....	IX
LISTA DE QUADROS .....	X
LISTA DE TABELAS .....	XI
SUMÁRIO .....	XIII
APRESENTAÇÃO .....	15
1 ARTIGO CIENTÍFICO 1.....	16
ELETRIFICAÇÃO RURAL E POLÍTICAS PÚBLICAS – O ATENDIMENTO A COMUNIDADES ISOLADAS .....	16
RESUMO .....	16
ABSTRACT.....	16
1.1 INTRODUÇÃO .....	17
1.2 A DEMANDA DE ENERGIA ELÉTRICA .....	20
1.3 DESENVOLVIMENTO E POLÍTICAS PÚBLICAS.....	21
1.4 POLÍTICAS PÚBLICAS.....	24
1.5 MATERIAL E MÉTODO .....	26
1.6 RESULTADOS E DISCUSSÕES .....	27
1.7 CONCLUSÃO .....	31
REFERÊNCIAS BIBLIOGRÁFICAS .....	33
2 ARTIGO CIENTÍFICO 2.....	35
MÉTODOS DE PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA REGRESSÃO E REDES NEURAIIS ARTIFICIAIS .....	35
RESUMO .....	35
ABSTRACT.....	35
2.1 INTRODUÇÃO .....	36
2.2 REDES NEURAIIS ARTIFICIAIS .....	37

2.3	MATERIAL E MÉTODO .....	42
2.4	RESULTADOS E DISCUSSÕES .....	44
2.4.1	APLICAÇÃO DE REGRESSÃO LINEAR MÚLTIPLA .....	47
2.4.2	APLICAÇÃO DE REDES NEURAIIS .....	49
2.5	CONCLUSÃO .....	54
	REFERÊNCIAS BIBLIOGRÁFICAS .....	56
	APÊNDICES .....	58
	APÊNDICE A – REDES NEURAIIS ARTIFICIAIS .....	59
	APÊNDICE B – REGRESSÃO LINEAR MÚLTIPLA E ESTATÍSTICAS .....	95

## APRESENTAÇÃO

O trabalho está organizado de forma que no Artigo 1, foi desenvolvida uma Revisão de Literatura acerca dos temas associados à problemática do consumo de energia elétrica rural, como conceitos de desenvolvimento, sustentabilidade e suas relações com as políticas públicas de eletrificação rural brasileiras, apresentando e analisando resultados de programas públicos de eletrificação rural como o Programa de Desenvolvimento para Estados e Municípios (PRODEEM), vigente de 1994 a 2003 e o Programa Luz para Todos (LpT) iniciado em 2003, tratando especialmente dos projetos que utilizam fontes de energia alternativas e renováveis como a solar. O estudo revela a ineficácia de programas de eletrificação rural brasileiros, apontando como sua causa a falta de planejamento que considere as realidades e potencialidades locais, bem como a participação das populações envolvidas no planejamento e gestão dos programas e o fato de que em todas as iniciativas o consumo de energia elétrica é considerado como uma constante padronizada, entendimento refutado pelo estudo, e que compromete a eficiência dos sistemas instalados por sub ou super dimensionamento destes.

No Artigo 2, a partir do entendimento de que o consumo de energia elétrica é não é uma constante apresentam-se resultados obtidos a partir da aplicação de dois métodos de previsão de consumo de energia elétrica, Regressão Linear Múltipla e Redes Neurais, sobre amostras obtidas em comunidades isoladas do meio urbano e uma amostra mista, formada por famílias rurais e urbanas, estabelecendo-se análise sobre os métodos e diferentes conjuntos de variáveis de entrada dos modelos, os quais têm como principal variável de saída o consumo de energia elétrica. O estudo mostra que as Redes Neurais podem ser utilizadas com alta eficiência na previsão de consumo de energia elétrica, seja para comunidades isoladas, ou para famílias urbanas.

## 1 ARTIGO CIENTÍFICO 1

### ELETRIFICAÇÃO RURAL E POLÍTICAS PÚBLICAS – O ATENDIMENTO A COMUNIDADES ISOLADAS

#### RESUMO

Políticas públicas subsidiam implantação de sistemas individuais de geração de energia no atendimento a populações isoladas, e o consumo de energia elétrica é tratado como uma constante, gerando ineficiência dos sistemas. Neste trabalho, desenvolve-se estudo sobre políticas públicas de eletrificação rural, especialmente, em atendimento às comunidades isoladas geograficamente e seus resultados efetivos em termos de eficiência dos programas e aumento da qualidade de vida das populações. A pesquisa revela que quantias importantes de recursos públicos foram e ainda devem ser investidos no Brasil em programas de eletrificação rural, visando à universalização do acesso à energia elétrica, o que sem o planejamento necessário leva a desperdícios e ineficiência dos sistemas implantados, fatores que implicam em descrédito com relação às tecnologias de geração de energia a partir de fontes alternativas e dos próprios programas governamentais, sendo que a falta de conhecimento e participação das comunidades envolvidas, junto com a ausência de dimensionamento do consumo de energia apontam para o fracasso dos programas.

**Palavras-chave:** Comunidades Isoladas. Consumo de energia elétrica. Programas públicos. Eletrificação rural.

#### ABSTRACT

Public policies subsidize deployment of individual power generation in serving isolated populations, and electricity consumption is treated as a constant, inefficiency systems. This work develops public policy study on rural electrification, especially those geographically isolated and the actual outcomes in terms of program efficiency and increased quality of life of the people. The research reveals that significant amounts of public funds have been and still should be invested in Brazil in rural electrification programs, aiming at universal access to electricity, which without the necessary planning leads to waste, inefficiency of deployed systems, factors that imply into disrepute with regard to the technologies of power generation from alternative sources and governmental programs themselves, and the lack of knowledge and participation of the communities involved, along with the absence of scaling the consumption point to the failure of the programs.

**Keywords:** Isolated Communities. Electric power consumption. Public programs. Rural electrification.



## 1.1 INTRODUÇÃO

A disponibilidade de energia elétrica no meio rural é um importante vetor de desenvolvimento. A partir desse recurso, é possível favorecer a irrigação, automatizar processos, beneficiar produtos e permitir, em alguns casos, atividades produtivas no período noturno. Ocorre que, no Brasil, justamente esse setor apresenta restrições no tocante ao fornecimento de eletricidade.

O modelo adotado para o setor elétrico brasileiro contribuiu para o êxodo rural, quando essa população passa a ocupar os centros urbanos, levando consigo não só a demanda por energia elétrica, mas também todas as demais associadas a uma comunidade, tais como: educação, moradia, saúde e saneamento (CARVALHO & CRAVEIRO, 2006).

O serviço de distribuição de energia elétrica é considerado essencial pela Constituição Federal. Embora não expressa na Lei Maior, da leitura do § 1º do artigo 9º, que trata do direito de greve, vê-se a remissão à legislação infraconstitucional quanto à competência para definir os serviços ou atividades essenciais. A partir da promulgação da lei 7.783, de 28 de junho de 1989, tal definição se tornou imperativo legal. Em seu artigo 10º, inciso I, assim está definido que são considerados serviços ou atividades essenciais o tratamento e abastecimento de água; produção e distribuição de energia elétrica, gás e combustíveis;”

A Constituição Federal de 1988 estabelece, ainda, em seu artigo 1º, incisos II e III, a cidadania e a dignidade da pessoa humana como fundamentos da República Federativa do Brasil. Os direitos de cidadania ali previstos são garantidos a todos, sem distinção, sendo pressupostos da atividade estatal na consecução de seus objetivos. Ao estabelecer o direito de todos os cidadãos brasileiros ao acesso gratuito ao serviço público de eletricidade, a Lei 10.438, de 26 de abril de 2002, encontra abrigo perfeito no texto constitucional. A Resolução 233 da ANEEL, de abril de 2003, estabeleceu metas de atendimento para a consecução da universalização do acesso à energia elétrica, metas a serem alcançadas até o ano de 2015, em alguns casos.

A partir do lançamento do Programa Luz para Todos (LpT), em novembro de 2003, cujo objetivo é levar o acesso à energia elétrica gratuita à população rural, o Governo Federal e as concessionárias de energia devem instalar sistemas que atendam à demanda energética das populações rurais, o que viabilizou muitos empreendimentos, tanto via rede de distribuição elétrica, quanto por meio de sistemas de geração de energia individuais. Com a nova etapa do Programa Luz para Todos, definida de 2011 a 2014, pela Portaria nº. 628 de 03

de novembro de 2011, muitas outras famílias devem ser beneficiadas com acesso à energia elétrica, de forma que novamente impõe-se a questão da demanda em sistemas de geração com fontes alternativas, adequados para locais onde a rede de distribuição elétrica torna-se inviável econômica, geográfica, ou ambientalmente.

O Programa LpT, segundo o Ministério de Minas e Energia (MME), cumpriu a meta de levar energia a dois milhões de famílias rurais, em maio de 2009, o que equivale a, aproximadamente, dez milhões de pessoas dispostas de energia elétrica e da esperança com relação ao desenvolvimento que esta energia pode proporcionar. Segundo Censo 2000 (IBGE, 2001), este era o número de domicílios no país sem acesso à energia elétrica, e serviu de parâmetro para a meta do Governo Federal, entretanto foi identificado um aumento de quase um milhão de domicílios no decorrer das execuções de obras do LpT.

Os números do Censo 2010 não refletem as afirmações do MME, posto que o Censo 2000 registrou dois milhões de domicílios sem energia e o Censo 2010 registra quase 730.000 famílias ainda sem acesso à esse serviço, verifica-se assim, uma discrepância entre os números de atendimento referidos pelo MME e IBGE, considerando que o número a ser atendido pelo LpT até 2010 era de 2.965.908 domicílios, número maior do que o apontado pelo Censo 2000, mas que ainda não contemplou as 728.512 residências sem atendimento contadas pelo IBGE no Censo 2010.

De fato, até o momento do fechamento do Censo 2010, o Brasil apresenta altas taxas de exclusão energética e sofre o impacto de social dessa exclusão, como o êxodo rural e conseqüente perda de produção agrícola. Por outro lado, as taxas migratórias podem ter sido minimizadas também pelos resultados do LpT até 2010, pois a taxa de crescimento do êxodo rural baixou de 1,31% para 0,65%, nos últimos dez anos e uma onda migratória inversa se configurou, trazendo novamente para áreas rurais noventa e seis mil famílias, possivelmente descrentes da vida nas cidades e voltando a sua terra e ao que sabem fazer, em função da perspectiva de melhor condição de vida proporcionada, especialmente, pela energia elétrica.

Ainda assim, o número de famílias sem acesso à energia elétrica é bastante alto, em todas as regiões do país, conforme Quadro 1.1-1.

São alarmantes as diferenças de acesso à energia elétrica entre as Regiões Sul, Sudeste e Centro-oeste se comparadas às Regiões Norte e Nordeste, todavia os mais de sessenta e seis mil domicílios que não contam com energia elétrica na Região Sudeste, podem representar mais de trezentas mil pessoas no escuro, sem possibilidades de comunicação de massa e

refrigeração, população de uma região rica, mas que também reflete as desigualdades sociais notórias no país.

**Quadro 1.1-1 Domicílios particulares premanentes, por existência de energia elétrica, segundo as grandes regiões e unidades da federação (IBGE, 2010)**

Grandes Regiões e Unidades da Federação	Domicílios particulares permanentes	Existência de Eletricidade Total	Existência de energia elétrica Da companhia distribuidora	Existência de energia elétrica De outra fonte	Existência de energia elétrica Não tinham
Brasil	57.324.185	56.595.007	56.044.395	550.612	728.512
Região Norte	3.975.533	3.724.295	3.547.426	176.869	251.207
Região Nordeste	14.022.901	14.583.662	14.460.942	122.720	339.087
Região Sudeste	25.199.799	25.133.234	24.937.720	195.514	66.211
Região Sul	8.891.279	8.859.224	8.829.870	29.354	31.979
Região Centro-Oeste	4.334.673	4.294.592	4.268.437	26.155	40.028

Nas últimas décadas, a privatização de parte do setor elétrico brasileiro acentuou a exclusão energética, pois a eletrificação para o meio rural, em geral, esbarra em aspectos financeiros, como a baixa densidade da carga, ocasionada pela dispersão das instalações, colaborando para o afastamento do investidor privado (WEITZEL et al., 2006). Assim, compete ao poder público arcar com tais investimentos, assumindo um papel social importante, não só absorvendo o ônus dos investimentos iniciais, mas também concedendo subsídios tarifários às comunidades atendidas. Ocorre que tais recursos, via de regra, são mantidos por meio de arrecadação tributária e, ainda que indiretamente, são repassados aos demais contribuintes (NASCIMENTO et al., 2003; NASCIMENTO et al., 2005), assim a otimização na alocação desses recursos torna-se imprescindível, o que passa inerentemente pelo dimensionamento do consumo de energia.

A previsão de demanda de energia elétrica é de fundamental importância para todo setor elétrico desde o momento da contratação da energia até a distribuição da mesma, no sentido de garantir que o sistema de distribuição trabalhe dentro dos limites contratados, evitando assim pesadas multas. Nestes contratos, o erro médio das previsões deve estar abaixo da tolerância prevista em torno de 5% (RUAS et al, 2007) . Por outro lado, no atendimento a comunidades isoladas é preciso satisfazer, totalmente, demanda energética das famílias e por força de lei, de forma intermitente, ou corre-se o risco de perder credibilidade junto à população, de os investimentos tornarem-se ineficazes e ao mesmo tempo não superestimar a demanda sob pena de perda de eficiência em sistemas individuais.

Neste estudo apresenta-se resultados de programas públicos brasileiros de eletrificação rural, como Programa de Desenvolvimento de Estados e Municípios (PRODEEM) e LpT, estabelecendo-se discussão acerca das causas de seus fracassos, tanto econômicos quanto sociais e técnicas, especialmente no que se refere à ausência de previsão de consumo de energia elétrica familiar desde os seus planejamentos.

## 1.2 A DEMANDA DE ENERGIA ELÉTRICA

A complexidade verificada no momento de dimensionar um sistema individual nasce da problemática imposta pela existência de duas variáveis aleatórias: o recurso solar e a demanda. Quanto à previsão da disponibilidade de energia solar, ou eólica há avanços importantes, ao passo que em relação à demanda essa evolução não tem ocorrido, especialmente em termos de modelos de previsão no âmbito residencial rural. Toda a proposta de eletrificação rural no Brasil se embasa em uma previsão de consumo de energia elétrica familiar padronizado.

Segundo MORANTE e ZILLES (2004), em pesquisa realizada junto a comunidades isoladas que possuem sistemas fotovoltaicos domiciliares, e em comunidades ligadas à rede de distribuição de energia, no Brasil e Peru

[...] pode-se constatar que o comportamento do consumo de energia elétrica está sujeito à interação de diversos fatores que atuam de forma simultânea. Desta maneira, o consumo difere ao longo do tempo tanto na mesma família, quanto entre as famílias ou as comunidades. Em tal sentido os resultados exibem seu caráter aleatório o qual, desde nosso ponto de vista, reflete a atuação de um elenco de fatores.(MORANTE;ZILLES, 2004, p.3)

Os fatores de que falam os autores citados se constituem em elementos de ordem técnica, gerencial, psicológica, geográfica, demográfica, econômica e sociocultural; refletindo a tecnologia envolvida e sua difusão, a gestão dos sistemas no contexto da comunidade, as relações sociais e cultura estabelecidas pelos atores sociais – incluindo o agente implantador - o tipo de relevo, a paisagem, o clima e os acidentes geográficos que configuram o cenário e as influências externas sofridas, bem como a estrutura familiar com suas peculiaridades econômicas e socioculturais.

Este estudo considerou os aspectos envolvidos na configuração do consumo de energia elétrica em comunidades isoladas, entendendo que “qualquer estudo do comportamento da

demanda de energia elétrica deve necessariamente centrar-se no microcosmo das famílias [...] pode-se ressaltar que os projetos de eletrificação com tecnologia fotovoltaica, ou com outro tipo de sistema individual, não podem desconsiderar esta importante questão” (MORANTE, 2004, p.234).

### 1.3 DESENVOLVIMENTO E POLÍTICAS PÚBLICAS

O processo de desenvolvimento, entendido como aquele que considera aspectos socioculturais e tecnológicos e, especialmente, a participação dos atores sociais no planejamento e ações que constituem a transformação, implica em mudança de valores, regras, atitudes e compartilha da importância da capacidade de escolha. A cultura ocidental, por exemplo, promove valores que levam o indivíduo a desejar o desenvolvimento com muito mais ansiedade do que naquelas culturas mais tradicionais, onde a introdução de tecnologias enfrenta maiores obstáculos.

Neste sentido, três questões se destacam na composição da problemática referente à introdução de novas tecnologias, associadas aos sistemas de geração de eletricidade, na vida das comunidades, conforme Serpa :

O primeiro aspecto diz respeito ao processo de mudança decorrente do desenvolvimento tecnológico. O segundo, se refere ao planejamento e gestão dos projetos e dos programas de difusão da tecnologia. O terceiro, ao desenvolvimento das ações de difusão<sup>1</sup> da tecnologia. (SERPA, 2001).

Entendendo que o desenvolvimento não é apenas a apropriação de tecnologias, e sim um processo que remete ao engendramento dos aspectos culturais, sociais e tecnológicos, Amartya Sen, (1993) argumenta que a ação política - projetos e programas de

---

<sup>1</sup> O termo difusão é entendido, neste trabalho, como processo pelo qual uma inovação passa a ser socialmente aceita ou, simplesmente, designa todos os processos ordenados, que produzem semelhanças culturais em várias sociedades que não as produzidas pela invenção. Refere-se à propagação de novas idéias ou novas unidades de cultura, de uma pessoa ou grupo a outro. Fazendo-se um paralelo entre inovação e mutação, a difusão pode ser descrita como o processo de seleção, pelo qual um traço ou é adicionado a outros que fazem parte de uma cultura, ou substitui um traço existente. A difusão bem sucedida de um traço é o resultado de um processo de avaliação, através do qual indivíduos e grupos passam a apreciá-lo e aceitá-lo.

desenvolvimento – devem ter o objetivo de ampliar a capacidade das pessoas de serem responsáveis por atividades e estados valiosos e valorizados, ou seja, na linha de uma ética da capacidade. Disso decorre a orientação de que uma boa ação pública não é aquela que distribui bens a receptores passivos, mas aquela que amplia as escolhas das pessoas e promove suas capacidades, principalmente sua capacidade de criação e escolha. O foco em distintas capacidades humanas implica um enfoque pluralista, pois

“as coisas que as pessoas consideram valiosas fazer, ou ser, podem ser muito diversas, e as capacidades valiosas variam desde a liberdade elementar, tais como livrar-se da fome e da desnutrição, até capacidades complexas, tais como a obtenção do auto respeito e a participação social” (SEN, 1993:332).

Essa discussão remete à idéia de que o desenvolvimento tem um significado bem mais amplo que o de crescimento econômico, que significa mero aumento da riqueza tangível ou física. Neste sentido, em 1980, o relatório “Brand Commission”, da ONU, referia que “desenvolvimento é mais do que a passagem da condição de pobre para a de rico, de uma economia tradicional rural para uma sofisticada: carrega ele consigo não apenas a idéia da melhor condição econômica, mas também a de maior dignidade humana, mais segurança, justiça e equidade” (KLERING et al.,1998:6 apud SERPA, 2001, p.11).

Em relação ao planejamento e gestão, trabalhos de avaliação acerca da difusão da tecnologia de geração de energia em áreas rurais de países em desenvolvimento têm indicado problemas na implementação dos projetos, implicando em fracassos, que não estão relacionados diretamente com a tecnologia, mas com a ausência de metodologias que considerem os aspectos socioculturais e a participação dos atores sociais, desde o planejamento até a implementação e operação dos sistemas (MORANTE, 2004).

As avaliações de projetos em desenvolvimento, em várias partes do mundo, feitas pelo Banco Mundial, indicam um rol de problemas que estão relacionados ao planejamento do projeto executivo, tais como a instalação técnica - reparos e manutenção, o treinamento adequado do usuário para a manutenção dos sistemas e gestão local, além da falta de avaliações econômicas e técnicas realistas (SERPA, 2001).

Para além dessas questões, autores destacam que os problemas estão também relacionados ao fato de muitos projetos terem sido planejados sem participação dos atores sociais, sem pesquisa de campo sobre as demandas de energia e as formas de organização das comunidades. Com esse tipo de encaminhamento, os projetos são elaborados em gabinete, conforme modelo padrão, pelo qual estão definidos os consumos genéricos de energia, o

suposto envolvimento da comunidade e sua aceitação como decorrência lógica dos benefícios hipotéticos do programa. Projetos que prevêem as etapas de monitoramento são raros, apesar destas se constituírem em ações necessárias à própria sustentabilidade do projeto. Essas considerações são negligenciadas dados os custos das avaliações sobre o desempenho técnico dos sistemas instalados e sobre a satisfação dos atores sociais para com a tecnologia (SERPA, 2001; MORANTE, 2004)

Sob este cenário, dois eventos são extremamente pertinentes, o primeiro acerca da capacitação do agente da inovação, cuja necessidade de uma formação humanística se torna bastante relevante, visando o conhecimento de formas sociais e culturais, para auxiliá-lo na tarefa de vencer eventuais resistências dos usuários. Porém, essa orientação não resolve todos os problemas, já que nos programas de desenvolvimento estão representados em distintos sistemas culturais. Tais diferenças podem ser grandes, pois a influência produzida pelo sistema de valores da cultura do técnico é um obstáculo para a mudança, tanto quanto as formas culturais do grupo visado. O etnocentrismo inerente às culturas leva o técnico a concluir que os programas e técnicas que obtiveram sucesso em determinada comunidade serão igualmente adequados a outras, daí a necessidade de se reconhecer que a eficácia do técnico está em estabelecer uma relação de empatia com as condições de vida e trabalho na comunidade onde atua (SERPA, 2001).

O segundo evento a ser considerado é a preparação ética para a seguinte questão: existe direito de decidir o que é bom para certa comunidade? Segundo Foster, “não existe resposta fácil para a questão ética e seguramente não existe resposta que possa merecer o adjetivo de científica” (FOSTER, 1964).

Somente o reconhecimento da questão ética e o estabelecimento de um processo que envolva ampla discussão entre agentes implantadores da tecnologia e comunidades é que poderá contribuir para orientação de cada projeto.

O terceiro aspecto a destacar-se envolve o desenvolvimento das ações de difusão da tecnologia. O sucesso da difusão de um traço é resultado de um processo de avaliação, através do qual indivíduos e grupos passam a apreciá-lo e aceitá-lo. Nesse sentido, poucos são os projetos nos quais o planejamento das ações é feito de forma participativa, por vezes, essa atitude sequer é cogitada pelo planejador, imbuído de uma visão desenvolvimentista delimitada pelo etnocentrismo.

A literatura acerca da difusão da tecnologia apresenta exemplos interessantes sobre a atitude do planejador que, na avaliação de resultados da implantação, operação e manutenção



dos sistemas de geração de energia, menospreza os impactos da alienação dos atores sociais, apesar do projeto executivo, algumas vezes, conceber o respeito às tradições culturais da população como ente fundamental.

Poucos são os projetos em que o programa é pensado como um processo de aprendizagem em via de mão dupla, uma constituída pelo usuário, individual ou coletivo, que passa a compreender a tecnologia, sentir-se parte de sua construção e elaborar internamente as mudanças culturais cabíveis; e outra apreendida pelo planejador, que ao entender a cultura local, mobiliza seu conhecimento técnico em função das demandas da comunidade, em todas as suas especificidades, até que essa nova tecnologia seja absorvida pelo sistema cultural da comunidade e gerida por ela (SERPA, 2001).

#### 1.4 POLÍTICAS PÚBLICAS

Após a reforma do setor elétrico, esforços foram direcionados na busca do aumento do atendimento do serviço de energia elétrica no Brasil, com os projetos denominados Programa de Desenvolvimento Energético nos Estados e Municípios (PRODEEM); Programa Nacional de Eletrificação Rural Luz no Campo; Projeto de Centro Comunitários de Produção (CCP); Programa de Universalização Luz para Todos (LpT).

O maior programa de eletrificação rural a nível mundial, segundo o Ministério de Minas e Energia, em relatório parcial publicado em 2009, Luz para Todos – PLpT – foi instituído pelo Decreto nº 4.873, de 11 de novembro de 2003 e alterado pelo Decreto nº 6.442, de 25 de abril de 2008, e destinou-se a suprir, até o ano de 2010, a demanda de energia elétrica de parcela da população do meio rural brasileiro que ainda não tinha acesso a esse serviço público. Na Resolução nº. 83 de 2004 ficam regulamentadas características dos sistemas individuais de geração com fontes intermitentes – SIGFI – no âmbito do LpT, e em 2009, foi publicado o Manual de Projetos Especiais, tratando dos SIGFI como segue

No âmbito do Programa “Luz Para Todos”, foram identificadas diversas situações em que o atendimento está condicionado à execução de projetos com características especiais, uma vez que as localidades a serem atendidas encontram-se distantes das redes de distribuição de energia elétrica existentes, de difícil acesso, especialmente para o transporte de materiais e equipamentos e normalmente com baixa densidade populacional. Desta forma, **torna-se imperativo o atendimento por meio de geração de**



**energia elétrica descentralizada, utilizando fontes renováveis compatíveis com a realidade local,** (...) utilizando-se tecnologias amparadas pela legislação em vigor. Assim sendo, os Projetos Especiais, dizem respeito a projetos de eletrificação rural destinados ao atendimento das situações acima mencionadas, de **forma sustentável, priorizando a utilização de fontes renováveis e mitigando o impacto ambiental.** (MME, 2009, fl. 1 grifo do autor)

Uma nova etapa do PLpT foi instituída a partir do Decreto n.º. 7520 de julho de 2011, para atender demandas ainda não contempladas em anos anteriores, especialmente, nas regiões Norte e Nordeste, onde encontra-se boa parte das populações inseridas no Plano Brasil sem Miséria, do Programa Territórios de Cidadania, e minorias sociais como indígenas, quilombolas, comunidades extrativistas e localizadas em áreas destinadas a empreendimentos do setor elétrico. A previsão é de que o novo PLpT estenda-se até 2014, e segundo seu manual de operacionalização,

O Programa “LUZ PARA TODOS” se integra ao Programa Territórios da Cidadania e ao Plano Brasil Sem Miséria implementados pelo Governo Federal, para assegurar que o esforço de eletrificação do campo resulte em incremento da produção agrícola, proporcionando o crescimento da demanda por energia elétrica, o aumento de renda e a inclusão social da população beneficiada. (MME, 2011)

Nesse sentido, os sistemas fotovoltaicos individuais, eólicos, ou os que se utilizem de biomassa continuam sendo alternativas por sua característica modular, podendo atender até mesmo uma única residência isolada, como é o caso de muitos domicílios no Norte do país, e nesse ponto reside a relevância do adequado dimensionamento do consumo energético familiar para a sustentabilidade dos projetos, inclusive para que atendam alguns critérios de elegibilidade do MME quanto aos SIGFI<sup>2</sup>, registrados no Manual de Projetos Especiais (MME, 2009), como

5.1.2 Que propiciem o atendimento de comunidades isoladas, preferencialmente da Amazônia legal, que não possam ser atendidas por extensão de rede elétrica convencional, devido a limitações financeiras, técnicas ou ambientais;

5.1.3 Que utilizem tecnologia que melhor aproveite os potenciais energéticos locais ou regionais;

5.1.5 **Que disponibilizem a potência mínima capaz de atender às necessidades básicas dos domicílios (iluminação, comunicação e**

---

<sup>2</sup> Sistemas Individuais de Geração de Energia Elétrica com Fonte Intermitente

**refrigeração), em face da análise de demanda energética da comunidade;** (MME, 2009, grifo do autor).

O Manual de Projetos especiais prevê então, que as necessidades inclusive de refrigeração sejam atendidas, o que requer cada vez mais precisão no dimensionamento da demanda, sob o ônus de que uma sub, ou superestimação de demanda possa levar à perda de eficiência e elevados custos com manutenção do equipamento de geração de energia individual. Se a demanda não for suprida, o agente executor do projeto não terá atendido aos critérios de elegibilidade impostos pelo MME, se por outro lado a demanda for superdimensionada podem ocorrer problemas de ordem técnica e difícil manutenção, especialmente com relação aos acumuladores de energia.

O Manual de Projetos Especiais estabelece alguns padrões mínimos, que não foram fruto de estudos de campo, mas de pesquisas em literatura e consulta pública a especialistas, gerando uma padronização que é refutada pelo trabalho de Morante (2004), quando trata de programas anteriores, os quais sempre consideraram a demanda como uma constante

(...) esta padronização não deveria desconsiderar as diferenças de necessidades quando se trata da eletrificação rural dispersa, em particular quando se utilizam sistemas fotovoltaicos. Na prática, essas diferenças têm por conseqüência a utilização dos equipamentos, em sua maioria, fora da condição de operação para a qual foram dimensionados. Muitos deles trabalharão com folga e outros, sobrecarregados. Em ambos os casos surgirão problemas de ordem técnica no sistema como um todo e, de maneira especial, no desempenho dos acumuladores de energia (baterias). (MORANTE, 2004).

Para evitar os problemas levantados por Morante (2004) é conveniente que se repensem os métodos utilizados para definição dos padrões mínimos iniciais para projetos de eletrificação rural, e que no curso de sua implementação se estabeleça um acompanhamento do consumo de energia, permitindo avaliações acerca da potência ideal para cada família, até que o consumo de energia atinja um ponto de maturação.

## **1.5 MATERIAL E MÉTODO**

A pesquisa foi iniciada a partir de revisão de literatura acerca da tecnologia solar fotovoltaica, especialmente em suas aplicações de caráter social, e nesse sentido teve como foco programas vinculados a políticas públicas de eletrificação rural. No percurso do estudo,

foi percebida a problemática referente à demanda de energia elétrica em comunidades isoladas, as variações que esta demanda pode apresentar, e o insucesso que pode impor aos programas quando é tratada como constante no planejamento da eletrificação rural.

Documentos públicos originários dos programas, editados pelo Ministério de Minas e Energia e auditorias realizadas pelo tribunal de Contas da União serviram de base para a coleta de dados e análise de seus efeitos.

Revisão de literatura acerca da questão da demanda de energia elétrica foi desenvolvida a fim de se discutir sua influência sobre os resultados dos programas de eletrificação, já que esta é sempre tratada como uma constante padronizada, o que difere de estudos nessa área como o de Morante (2004).

## 1.6 RESULTADOS E DISCUSSÕES

Programas de eletrificação rural implementados no Brasil, como o PRODEEM<sup>3</sup> e o atual Luz para Todos, o qual desde 2009 é responsável pelos equipamentos remanescentes do PRODEEM, não obtiveram sucesso, tanto com relação à credibilidade das tecnologias de geração de energia baseadas em fontes alternativas junto às comunidades isoladas, quanto a sua sustentabilidade técnica e financeira, as quais foram impactadas pela falta de planejamento que não considerou a participação das populações na implementação dos programas, na gestão participativa, imprescindível para a manutenção do funcionamento dos sistemas, e na difusão da tecnologia de forma que as populações se apropriassem do conhecimento e pudessem escolher formas de utilizar a energia em prol de seu desenvolvimento e qualidade de vida. O Quadro 1.6-1 mostra o montante de investimentos, somente com os sistemas instalados no âmbito do PRODEEM de 1994 até 2002, todos sistemas fotovoltaicos<sup>4</sup>.

Considerou o TCU<sup>5</sup> que até o final de 2003, ano de término do PRODEEM, o montante de investimentos, contando sistemas emergenciais, foi de aproximadamente 45 milhões de dólares. Os sistemas a que se refere o Quadro 1.6-1, em sua maioria não estão mais em funcionamento, alguns foram completamente depreciados, outros poucos retirados dos locais de funcionamento original e ainda sem destino determinado. Resultados de

---

<sup>3</sup> Programa de desenvolvimento de Estados e Municípios.

<sup>4</sup> Sistema de geração de energia que transforma energia solar em energia elétrica.

<sup>5</sup> Tribunal de Contas da União.

auditorias requisitadas em 2009 para o Programa ainda não foram divulgadas em sua totalidade, e mesmo resultados parciais, como o de Furnas, apresentam-se bastante vagos. Portanto, aproximadamente 45 milhões de dólares foram investidos sem que, até o momento, se possa avaliar sua efetiva contribuição para o desenvolvimento das comunidades envolvidas (BRASIL, 2003).

**Quadro 1.6-1 Investimentos em sistemas individuais fotovoltaicos, no âmbito do PRODEEM, de 1998 até 2002.**

	Atendimento até 2001		Atendimento em 2002		Atendimento até 2002	
	Sistemas	US\$	Sistemas	US\$ 1	Sistemas	US\$
TOTAL	4592	\$ 16.990.400,00	3000	\$ 15.801.000,00	7592	\$ 32.791.400,00

O Quadro 1.6-2 apresenta a distribuição regional dos sistemas em suas cinco fases de implantação, as quais são criticadas pela auditoria do TCU<sup>6</sup>, pois há em todo processo descontinuidade, haja vista que a Secretaria Federal de Controle Interno – SFC – detectou que no ano 2000 havia 138 sistemas (13% da amostra de 1029 sistemas) ainda estocados, além de outras inconformidades que impedem o atendimento às comunidades, num total de 77% de inconformidades, e mesmo assim o Programa seguia sem rever suas práticas, sequer de instalação e manutenção (BRASIL, 2003).

**Quadro 1.6-2 Elaborada pelo autor. Fonte: PRODEEM/MME**

SISTEMAS ADQUIRIDOS – DISTRIBUIÇÃO POR REGIÃO							
Fase	I e II	III	Emergencial	IV	Subtotal	V	Total
R Norte	265	299	50	808	1.422	511	1.933
E Nordeste	354	544	750	606	2.254	2.107	4.361
G Centro Oeste	342	26	0	226	594	88	682
I Sudeste	63	100	0	209	372	283	655
Ã Sul	41	4	0	150	195	11	206
O Estoque	88	86	0	901	1.075	0	1.075
Total	1.153	1.059	800	2.900	5.912	3.000	8.912
Valor (US\$ mil)	4.533	4.669	2.220	10.026	21.448	15.810	37.258

Fonte: CEPEL e MME. abril/2002

A incapacidade do Programa de manter os sistemas em operação, fez com que o Ministério de Minas e Energia adotasse na Fase V um novo modelo de aquisição dos equipamentos, os quais passaram a ser mais padronizados, entregues instalados e com garantia de 3 (três) anos de funcionamento, sem ênfase, entretanto, na formação de parcerias locais, no dimensionamento das reais necessidades dos beneficiários e no desenvolvimento de

<sup>6</sup> TC n°. 013.245/2002-7 Relatório de auditoria de natureza operacional programa energia das pequenas comunidades (nome atribuído ao PRODEEM no Plano Plurianual 2000-2003 do Governo federal)

recursos humanos especializados. Novamente impondo-se a questão do mau, ou nenhum, dimensionamento da demanda energética da comunidade a ser atendida.

O Programa deixava o treinamento dos usuários a critério das instaladoras nos Estados, que constituía-se de práticas como ligar e desligar o sistema, sem ensinar operações de conservação e manutenção das instalações, e nem mesmo ocorria orientação dos usuários no sentido de limpar os módulos, ocasionando o baixo desempenho do sistema pelo acúmulo de sujeira sobre os painéis e como obstáculo adicional, existia o fator cultural de os usuários não possuírem familiaridade com a utilização da eletricidade, dificultando a capacitação e manejo dos equipamentos, implicando em reduzido envolvimento da comunidade com o programa.

O descontrole patrimonial foi desastroso, segundo o TCU, pois frustraram os beneficiários e não permitiram que se desenvolvesse o sentido de zelo pelo sistema, ocorrendo a relevante perda patrimonial da União, de aproximadamente US\$ 5.000,00 por sistema, de equipamentos com alto valor unitário e que alcançam um valor total superior a US\$ 20.000.000,00 (vinte milhões de dólares), referentes à proporção dada como inexistente pela SFC e correspondente a 46% do total de sistemas instalados pelo Programa (BRASIL, 2003).

Ainda segundo o TCU, a proximidade de locais onde há energia convencional desestimulou as comunidades beneficiadas de exigir outros serviços potencialmente viáveis com o fornecimento da energia elétrica fotovoltaica. A existência desses serviços em localidades distantes poucos quilômetros, ou seja, acessíveis a custo de um pequeno deslocamento, fez com que as populações se satisfizessem com os serviços lá oferecidos, como: informação e lazer, basicamente por meio da televisão; escola já dotada de energia convencional; cursos e treinamentos oferecidos em centros comunitários e associações já alcançados pela rede de energia elétrica; entre outros. Outra implicação da maior proximidade com localidades alcançadas pela rede de distribuição de energia é a baixa valorização do sistema instalado pela comunidade, acarretando tanto resultados acanhados, quanto pouco comprometimento da população com a integridade física e a manutenção desses sistemas. Essa relação entre resultados e distância pode ser notada nas visitas às comunidades com diferentes distâncias à rede, por meio das entrevistas com os usuários e responsáveis. Os beneficiados mais entusiasmados estavam nas populações de comunidades da Amazônia, onde as distâncias à rede são enormes e não existe perspectiva de outra forma de acesso à energia (BRASIL, 2003).

A demanda por energia das comunidades não foi considerada no dimensionamento dos sistemas a serem nelas instalados. Como exemplo, em Alvorada do Norte/GO uma escola beneficiada com o bombeamento d'água pode melhorar as suas condições de higiene e de seus alunos, no entanto, a bomba só precisava funcionar 5 minutos por dia para garantir o volume de água consumido, evidenciando importante superdimensionamento e desperdício de investimento público, e com o sistema trabalhando muito abaixo de sua capacidade o que pode ocasionar problemas nos acumuladores de energia. As causas dessas situações podem ser entendidas pela demasiada centralização do programa, com aquisição padronizada de sistemas, aliada a sua acanhada estrutura funcional e de parcerias, que não permitiram o adequado levantamento de campo antes da instalação do sistema de geração de energia ou de bombeamento d'água. Não raro, para que fosse colocado em uso um equipamento, por exemplo, uma televisão, é forçoso que outro fosse desligado, ou que não fosse utilizada toda a iluminação possível.

O gestor do Programa afirmou ao TCU que inicialmente o Programa atendia equipamentos comunitários isolados pelo critério da demanda, ou seja, por solicitação de vereadores, prefeitos, deputados entre outros, afirmação que reflete a visão do PRODEEM sobre o conceito de demanda, ou seja, necessidade ou não do sistema energético, sem considerar as distintas necessidades de cada comunidade e família (BRASIL, 2003).

A partir da concepção centralizada é que o Programa desenvolveu seu modelo, baseado na aquisição de grandes lotes padronizados de equipamentos concebidos para atender as mais diversas condições encontradas no país, sem a necessidade de maiores participações das comunidades a serem beneficiadas, ou estudos das características locais. Após a aquisição, o Programa simplesmente escolheu, às vezes com participação das suas coordenações estaduais, os locais que deveriam ser contemplados e solicitou a alguns de seus parceiros, que os instalassem, ou seja, o modelo de atuação foi desenvolvido para atender às características do MME e não aos objetivos do Programa (BRASIL, 2003).

A escolha da tecnologia adotada, a fotovoltaica, se deu a partir da constatação de sua adequação a todo território nacional e sua capacidade de atendimento pontual. Sendo assim, não haveria necessidade de maiores estudos locais sobre a disponibilidade de fontes de energia mais adequadas, a demanda necessária pela comunidade, a forma de manutenção adequada e tudo o mais que fosse importante para se garantir o sucesso da implantação de um microsistema energético (BRASIL, 2003).

A atuação da auditoria realizada pelo TCU, por meio de representação apresentada no decurso dos trabalhos, acarretou a suspensão da aquisição de 6.000 equipamentos previstos para 2002 no valor de U\$ 30.000.000,00 (trinta milhões de dólares) e que correriam elevado risco patrimonial, com conseqüente desperdício de recursos públicos. A adoção das recomendações também garantiria o melhor uso dos, aproximadamente, US\$ 45.000.000,00 (quarenta e cinco milhões de dólares) estimados para a compra dos 9.000 equipamentos prevista para 2003. No mínimo, diminuiria significativamente o percentual de equipamentos não funcionando (36%), acarretando economia de mais de US\$ 16.000.000,00 (dezesseis milhões de dólares), correspondente ao valor dos equipamentos adquiridos que potencialmente viriam a quebrar. No entanto, não há documentos públicos que comprovem medida cautelar que tenha impedido o investimento previsto para o ano de 2003.

Conforme Resolução Normativa n.º. 493/12 as unidades consumidoras (UC) atendidas por meio de SIGFI devem ser enquadradas conforme as disponibilidades mensais de energia explicitadas no Quadro 1.6-3.

**Quadro 1.6-3 Padronização dos sistemas individuais pela ANEEL.**

Disponibilidade mensal garantida (kWh/ mês UC)	Consumo de referência (Wh/dia/UC)	Autonomia mínima (horas)	Potência mínima (W/UC)
13	435	48	250
20	670	48	250
30	1.000	48	500
45	1.500	48	700
60	2.000	48	1.000
80	2.650	48	1.250

Algumas concessionárias não investem em sistemas isolados por já terem constatado que os sistemas que garantem menos de 80 kWh por mês não atendem aos critérios de atendimento da totalidade das necessidades energéticas das famílias, mas atualmente esses são os critérios, igualmente padronizados, impostos pela Resolução da ANEEL de 2004.

## 1.7 CONCLUSÃO

As novas exigências legais impostas pelas revisões do programa Luz para Todos, obrigam os agentes executores de programas públicos de eletrificação rural a suprirem



integralmente a demanda energética das famílias, o que deve balizar uma nova perspectiva sobre a questão da demanda energética familiar, que não é uma constante, mas sim uma variável dependente de vários aspectos e nesse sentido, requer métodos para sua previsão e não padronizações, pois fracassos como o do PRODEEM, no qual se investiu mais de trinta milhões de dólares oriundos de recursos públicos, não podem ser repetidos no âmbito do LpT, sob pena de não se alcançar a universalização do acesso ao serviço de energia elétrica a toda população brasileira, universalização esta instituída por lei.

A participação da comunidade em todo processo de planejamento dos programas, bem como na pesquisa preliminar sobre a cultura e vida cotidiana, na escolha do tipo de fonte energética mais adequada, além da otimização na previsão do consumo de energia elétrica preliminarmente baseada na carga, pode se configurar em um caminho para o melhor aproveitamento dos sistemas de geração de energia individuais e conseqüentes benefícios às populações isoladas geograficamente, pontos dos quais os planejadores não podem prescindir.

O estudo revela que o sucesso dos programas de eletrificação rural pode estar associado a três tipos de questões, a técnica especialmente com relação à previsão do consumo de energia familiar, a social requerendo a participação das comunidades e a preocupação com seu desenvolvimento desde o planejamento dos programas e a econômica que aponta para a otimização dos recursos públicos ao se, efetivamente, considerar os dois aspectos citados anteriormente.



## REFERÊNCIAS BIBLIOGRÁFICAS

Agência Nacional de Energia Elétrica, ANEEL. Resolução Normativa nº 83, de 20 de setembro de 2004. Estabelece os procedimentos e as condições de fornecimento por intermédio de Sistemas Individuais de Geração de Energia Elétrica com Fontes Intermitentes – SIGFI.

BRASIL. Lei 7.783 de 28 de junho de 1989. Dispõe sobre o exercício do direito de greve, define as atividades essenciais, regula o atendimento das necessidades inadiáveis da comunidade, e dá outras providências. Diário Oficial da União. Disponível em [http://www.planalto.gov.br/ccivil\\_03/leis/L7783.htm](http://www.planalto.gov.br/ccivil_03/leis/L7783.htm). Acesso em julho de 2011.

BRASIL. Tribunal de Contas da União. Avaliação do TCU sobre o Programa Energia das Pequenas Comunidades / Tribunal de Contas da União. -- Brasília : TCU, Secretaria de Fiscalização e Avaliação de Programas de Governo, 2003. 22p. (Sumários Executivos/TCU-SEPROG ; 10)

BRASIL. Lei 10.438 de 26 de abril de 2002. Dispõe sobre a expansão da oferta de energia elétrica emergencial, recomposição tarifária extraordinária, cria o Programa de Incentivo às Fontes Alternativas de Energia Elétrica (Proinfa), a Conta de Desenvolvimento Energético (CDE), dispõe sobre a universalização do serviço público de energia elétrica, dá nova redação às Leis nº 9.427, de 26 de dezembro de 1996, nº 9.648, de 27 de maio de 1998, nº 3.890-A, de 25 de abril de 1961, nº 5.655, de 20 de maio de 1971, nº 5.899, de 5 de julho de 1973, nº 9.991, de 24 de julho de 2000, e dá outras providências. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2002/L10438.htm](http://www.planalto.gov.br/ccivil_03/leis/2002/L10438.htm). Acesso em 20 junho de 2011.

BRASIL. Decreto 4.873 de 11 de novembro de novembro de 2003. Institui o Programa Nacional de Universalização do Acesso e Uso da Energia Elétrica - "LUZ PARA TODOS" e dá outras providências. Diário Oficial da União. Disponível em: <http://www.dnpm.gov.br/conteudo.asp?IDSecao=67&IDPagina=84&IDLegislacao=28>. Acesso em: 13.01.2007.

BRASIL. Decreto 7.520 de 08 de julho de 2011. Institui o Programa Nacional de Universalização do Acesso e Uso da Energia Elétrica - "LUZ PARA TODOS", para o período de 2011 a 2014, e dá outras providências. Diário Oficial da União. Disponível em <http://www.lexml.gov.br/urn/urn:lex:br:federal:decreto:2011-07-08;7520>. Acesso em abril de 2011.

BRASIL. Decreto 6442 de 25 de abril de 2008. Institui o Programa Nacional de Universalização do Acesso e Uso da energia Elétrica - "LUZ PARA TODOS", para prorrogar o prazo ali referido. República Federativa do Brasil, Brasília, DF, 8 de março de 1934. Disponível em: <http://www.lexml.gov.br/urn/urn:lex:br:federal:decreto:2008-04-25;6442>. Acesso em abril de 2010.

CARVALHO, P.C.M.; CRAVEIRO, P.M.A. Introdução ao planejamento energético. Fortaleza: IEPRO/SEINFRA, 2006.

FOSTER, G. As culturas tradicionais sob o impacto da tecnologia. Rio de Janeiro: Fundo de Cultura, 1964.

MME. Ministério de Minas e Energia. Manual de Operacionalização do Programa Luz Para Todos. Revisão 6. Anexo à portaria no. 85 de 20 de fevereiro de 2009.

\_\_\_\_\_. Programa nacional de universalização do acesso e uso da energia elétrica manual de operacionalização para o período de 2011 a 2014. Decreto Nº 7.520, De 8/7/2011.

MORANTE, Federico B. T. Demanda de energia elétrica e desenvolvimento socioeconômico: o caso das comunidades rurais eletrificadas com sistemas fotovoltaicos. 311 p., 2004 Tese (Doutorado) – Programa Interunidades de Pós-Graduação em Energia da Universidade de São Paulo.

MORANTE, Federico. ZILLES, Roberto. Principais fatores que influenciam a demanda e o consumo de energia elétrica em sistemas fotovoltaicos domiciliares.. In: Encontro De Energia No Meio Rural, 5., 2004, Campinas. Proceedings online. Disponível em: <[http://www.proceedings.scielo.br/scielo.php?script=sci\\_arttext&pid=MSC0000000022004000200035&lng=en&nrm=abn](http://www.proceedings.scielo.br/scielo.php?script=sci_arttext&pid=MSC0000000022004000200035&lng=en&nrm=abn)>. Acessado em: 09/05/2009.

MONTGOMERY, D. C.. RUNGER, G. C.. Estatística Aplicada e Probabilidade para Engenheiros. 2ª.ed. Rio de Janeiro. LTC, 2003. 466 pg.

NASCIMENTO, M.V.G.; VIEIRA, L.S.; SOARES, G.F. Análise econômica comparativa entre opções para suprimento energético ininterrupto de pequenas comunidades. In: Seminário Nacional De Produção E Transmissão De Energia Elétrica, 17., 2003, Uberlândia. Anais... Uberlândia: Comitê Nacional Brasileiro de Produção e Transmissão de Energia Elétrica, 2003. 1 CD-ROM.

NASCIMENTO, M.V.G.; MELLO, J.C.O.; PIZETA, E.G.; CASTRO, A.L.; SPALDING, E.C.; BIANCO, A. A participação crescente dos encargos setoriais no custo de energia brasileiro In: Seminário Nacional De Produção E Transmissão De Energia Elétrica, 18., 2005, Curitiba. Anais...Curitiba: Comitê Nacional Brasileiro de Produção e Transmissão de Energia Elétrica, 2005. 1 CD-ROM.

SEN, Amartya. O desenvolvimento como expansão de capacidades. Lua Nova, São Paulo, n.28-29, Apr. 1993. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-4451993000100016&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-4451993000100016&lng=en&nrm=iso)>. Acessado em: 29 Junho de 2012.

SERPA, Paulo N.. Eletrificação Fotovoltaica em Comunidades Caiçaras e seus Impactos Socioculturais. Tese de Doutorado. Programa Interunidades de Pós-Graduação em Energia da Universidade de São Paulo, 2001. Disponível em: <http://www.iee.usp.br/biblioteca/producao/2001/teses/2001.htm>. Acessado em maio de 2009.

WEITZEL, L.; ROVERE, E.L.; CUNHA R. As demandas do novo modelo institucional do setor elétrico brasileiro. In: Congresso Brasileiro De Planejamento Energético, 5.,2006, Brasília. Anais... Brasília: Sociedade Brasileira de Planejamento Energético, 2006. 1 CD-ROM.

## 2 ARTIGO CIENTÍFICO 2

### MÉTODOS DE PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA REGRESSÃO E REDES NEURAIAS ARTIFICIAIS

#### RESUMO

Políticas públicas subsidiam implantação de sistemas individuais de geração de energia no atendimento a populações isoladas, e a demanda energética é tratada como uma constante, gerando ineficiência dos sistemas. Neste trabalho, desenvolve-se estudo sobre previsão de demanda energética por regressão linear e redes neurais. O estudo aponta para as redes neurais como método mais eficiente, pois um conjunto de oito variáveis de entrada explicaram 99% do consumo de energia elétrica, considerando uma amostra formada por famílias rurais e urbanas. Redes neurais podem ser utilizadas com alto grau de eficiência para previsão de demanda energética residencial, no meio rural e urbano.

**Palavras-chave:** Comunidades Isoladas. Demanda Energética. Redes Neurais. Regressão Linear.

#### ABSTRACT

Implementation of public policies subsidize individual systems of energy generation in serving isolated populations, and energy demand is treated as a constant, inefficiency of the systems. In this work, to study on energy demand forecast by linear regression and neural networks. The study points to the neural networks as a more efficient method, because a set of eight input variables explained 99% of electricity consumption, whereas a sample made up of rural and urban families. Neural networks can be used with high efficiency for residential energy demand forecasting, in rural and urban areas.

**Keywords:** Isolated Communities. Energy Demand. Neural Networks. Linear Regression.

## 2.1 INTRODUÇÃO

A busca de métodos, baseados em modelos para a previsão de curto, médio e longo prazo, passou a despertar especial interesse, tanto de gestores quanto de pesquisadores nas mais diversas áreas, tais como economia, marketing, finanças e engenharia. Os modelos tradicionais de previsão – baseados em ferramentas estatísticas – ainda hoje são amplamente usados. Entretanto, devido ao acelerado avanço tecnológico observado nas duas últimas décadas, as pesquisas passaram a ser direcionadas para o uso de ferramentas mais modernas para desempenhar a tarefa de fazer previsões.

No trabalho de Carpinteiro e Reis (2004), para a estimativa de demanda de energia elétrica, foi proposto o uso de redes auto-organizáveis em um modelo de estrutura hierárquica. Os autores mencionam que a demanda se configura de forma complexa, sendo necessárias previsões cada vez mais precisas. A relação entre a carga e seus fatores exógenos tem caráter complexo e não-linear, o que torna estas previsões difíceis de serem modeladas por meio de técnicas convencionais, tais como as séries temporais e a regressão linear (CARPINTEIRO; REIS, 2004).

Ao se estudar a demanda e o consumo de energia elétrica algumas particularidades se manifestam, essencialmente, com relação ao tamanho dos sistemas utilizados na produção dessa energia. No que se refere às redes de distribuição de energia elétrica, diversos métodos foram desenvolvidos para prever demanda. Por outro lado, no âmbito da tecnologia solar fotovoltaica, de outros sistemas individuais, e na literatura referente à eletrificação rural, esta é uma questão não resolvida, o que implica em se continuar implantando sistemas projetados a partir de uma demanda padronizada e constante, e assim correndo-se o risco de subdimensionar sistemas, e até mesmo superdimensioná-los conforme mostram estudos em comunidades tradicionais eletrificadas por meio de sistemas individuais de geração de energia (MORANTE, 2004).

Entendendo a demanda como uma variável e buscando métodos de previsão mais adequados poderá se estabelecer uma melhor alocação de recursos públicos e qualidade de vida para as comunidades, bem como a sustentabilidade dos programas de eletrificação rural.

O propósito deste estudo é avaliar se os modelos baseados em redes neurais artificiais podem ser mais eficientes que os tradicionais, fundamentados em métodos paramétricos, obtidos por regressão em séries temporais na previsão de demanda de energia elétrica, no contexto de comunidades rurais, considerando a teia de variáveis que podem estar associadas

a ela. Neste artigo, a análise se dará tendo como amostra dados de comunidades do Vale do Ribeira – SP, Pedra Branca (PE) e Vera Cruz (PA), da comunidade da Praia Vermelha, Ilha Grande – RJ e uma amostra de quinze domicílios urbanos, com fim de avaliar a capacidade de generalização dos modelos.

## 2.2 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNA) são modelos computacionais inspirados nas redes de neurônios biológicos, onde cada elemento de processamento é denominado neurônio artificial, cujas conexões entre estes são constituídas por coeficientes ponderadores representando as sinapses dos neurônios biológicos (HAYKIN, 2001).

As Redes neurais têm sido utilizadas em diversas aplicações, algumas delas, tradicionalmente, dominadas por outras técnicas, e apresentado bons resultados, como em Pedrollo (2006) que modelou variáveis hidrológicas, Paschoallino et al (2007) que realizou estudo sobre previsão de demanda de energia elétrica para o Brasil, utilizando Redes Neurais de Ellman, Catalão (2007) utilizando o método na previsão de preços de energia elétrica, Lugon et al. (2011) aplicando à modelagem de problemas inversos envolvendo adsorção gás-líquido, transferência de massa, transferência radiativa e simultânea de calor e massa. Mais recentemente Fiorin et al. (2011) publicou estudos sobre previsão de disponibilidade de recursos solares utilizando RNA, a qual vinha sendo tratada por meio de processos paramétricos no INPE<sup>7</sup>.

A aplicação de RNA na previsão de demanda de energia elétrica foi implementada para os setores industriais e residenciais urbanos, e de acordo com Teixeira (2009), a combinação de modelos tradicionais e de redes neurais implica em resultados mais precisos, como os apresentados em sua pesquisa sobre métodos de previsão de demanda de energia elétrica residencial urbana; porém não há registros de trabalhos que possam subsidiar políticas públicas voltadas à eletrificação rural fazendo uso de redes neurais, ou mesmo de métodos tradicionais como regressão.

---

<sup>7</sup> Instituto Nacional de Pesquisas Espaciais

A capacidade de aprendizado por meio de treinamento se configura como a mais notável das características nas redes neurais artificiais, proporcionando a adaptabilidade para produzir respostas adequadas a diversos problemas submetidos.

A generalização é a capacidade que a RNA tem de, quando submetida a padrões não utilizados no treinamento, ainda assim responder corretamente a esses padrões, desde que pertencentes ao mesmo domínio, após ter sido treinada para reproduzir um fenômeno a partir de uma amostra. Mais detalhes acerca de treinamento de redes neurais artificiais podem ser verificados em HAYKIN (2001), BRAGA et al. (2007) e CASTRO (1998).

Uma RNA pode ser caracterizada por três aspectos principais: (1) o padrão de conexões entre as unidades (arquitetura), (2) o método de determinação dos pesos das conexões (algoritmo de treinamento ou aprendizado) e (3) sua função de ativação. Os modelos neurais artificiais oferecem um paradigma atrativo, pois “aprendem” a resolver problemas através de exemplos (CASTRO, 1998).

O treinamento de RNA pode ser dividido em supervisionado, quando necessita de um “professor” durante a fase de aprendizagem, que antecede a utilização (execução) da rede; e não-supervisionado que é direcionado por correlações existentes nos dados de entrada e, portanto, não necessita de um “professor”.

A estrutura de uma RNA é constituída por conjuntos de nós, interligados por linhas de conexão, como em um grafo direcionado. Cada nó receptivo é um elemento de processamento, o neurônio artificial, com funções internas de transformação, capaz de receber qualquer número de conexões de entrada e de saída.

Um neurônio é uma unidade de processamento de informação que é fundamental para a operação de uma rede neural. A Figura 2.2-1 mostra o modelo de um neurônio que forma a base para o projeto de redes neurais artificiais, onde o conjunto de  $X_m$  são os dados de entrada da rede,  $W_{jm}$  o conjunto de sinapses, cada uma caracterizada por um peso ou força.; a junção aditiva fazendo o papel de um somador dos sinais de entrada, ponderados pelas respectivas sinapses do neurônio e uma função de ativação para restringir a amplitude da saída de um neurônio a um valor finito. O modelo 2.2-1 inclui também um bias aplicado externamente, representado por  $b_k$ .

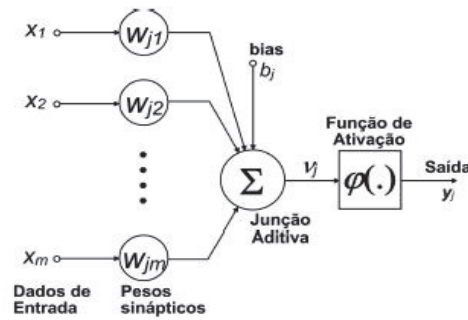


Figura 2.2-1 Modelo não-linear de um neurônio, utilizado em MLP. Adaptada (HAYKIN, 2001).

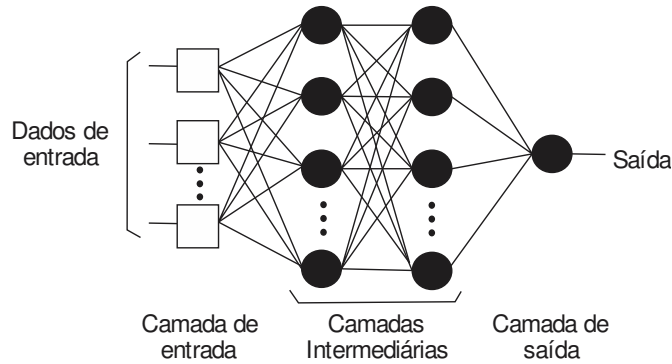
As arquiteturas do tipo perceptron de múltiplas camadas, ou Multilayer Perceptron (MLP), constituem os modelos neurais artificiais mais utilizados e conhecidos. Tipicamente, esta arquitetura consiste de um conjunto de unidades sensoriais que formam uma camada de entrada, uma ou mais camadas intermediárias (ou escondidas) de unidades computacionais e uma camada de saída. Os sinais de entrada são propagados camada a camada pela rede em uma direção positiva, ou seja, da entrada para a saída.

Nas redes MLP o somatório ponderado dos valores das entradas é combinado ao valor ajustável do bias ( $b_j$ ), que tem a função de expandir ou reduzir a entrada líquida da função de ativação para ajustar o chamado nível de atividade interna do neurônio ( $v_j$ ). A saída do neurônio ( $y_j$ ) é produzida pela aplicação da função de ativação ( $\varphi(\cdot)$ ) ao nível de atividade interna do neurônio. Estas relações são descritas pelas equações 2.1 e 2.2

$$v_j = \sum_{i=1}^m w_{ji} x_i + b_j \quad (\text{eq. 2.1})$$

$$y_i = \varphi(v_j) \quad (\text{eq. 2.2})$$

As redes MLP são muito utilizadas por apresentarem maior simplicidade e facilidade de implementação. Na Fig. 2.2-2 temos um exemplo de MLP com três entradas, duas camadas intermediárias com quatro neurônios e uma camada de saída com um neurônio, produzindo uma única informação de saída.



**Figura 2.2-2 Representação de uma rede Multi layer perception, com duas camadas ocultas. Elaborado pelo autor.**

Há alguns tipos de função de ativação como a linear, arco tangente, sigmóide e tangente hiperbólica. A função sigmóide descrita pela equação 2.3 e cujo gráfico tem a forma de S, é a forma mais utilizada na construção de redes neurais artificiais. Ela é definida como uma função estritamente crescente que exibe um balanceamento adequado entre comportamento linear e não-linear. Esta função se estende somente de 0 a 1 e algumas vezes é desejável que a função de ativação se estenda de -1 a 1, assumindo neste caso uma forma anti-simétrica em relação à origem, como a tangente hiperbólica conforme equação 2.4.

$$f(x) = \frac{e^{px}}{1+e^{px}} = \frac{1}{1+e^{-px}} \quad (\text{eq. 2.3})$$

$$f(x) = \frac{e^{px} - e^{-px}}{e^{px} + e^{-px}} = \tanh(px), \quad (\text{eq. 2.4})$$

O algoritmo backpropagation (retropropagação) é o mais comumente empregado no treinamento supervisionado de redes MLP. Em uma primeira fase ocorre a propagação do sinal funcional (feedforward) mantendo-se os pesos fixos de modo a gerar um valor de saída a partir das entradas fornecidas à RNA. Na segunda fase, as saídas são comparadas com os valores desejados, gerando um sinal de erro, que se propaga da saída para a entrada, ajustando-se os pesos de forma a minimizar o erro (retropropagação do erro, backpropagation).

Originalmente o ajuste dos pesos era baseado no método do gradiente decrescente para minimização do erro, isto é, a correção é aplicada a cada peso do neurônio. O algoritmo possibilita também o ajuste dos pesos das camadas intermediárias da rede. O ajuste dos pesos



sinápticos pode ser realizado a cada iteração (treinamento sequencial) ou após a apresentação à rede de todos os exemplos do conjunto de dados de treinamento (treinamento por lote), o que é mais usado. No treinamento por lote, a apresentação de todos os exemplos de treinamento à RNA é denominada época. Durante cada época, os pesos mantêm-se constantes e define-se uma nova função de custo como a média das funções de custo para cada exemplo. Devido ao cálculo de diferentes gradientes da nova função de custo para cada exemplo, ao final da época ajusta-se os pesos, de acordo com a média dos ajustes calculados para cada exemplo. (CASTRO, 1998)

O algoritmo de redução do gradiente por lote (Batch Gradient Descent) ajusta os pesos e desvios em direção ao gradiente negativo da função custo. O algoritmo Batch Gradient Descent with Momentum apresenta frequentemente melhores resultados ao introduzir o conceito de momento, que leva em conta os ajustes anteriores nos pesos, observando a recente tendência de mudanças na superfície de erro, desprezando pequenas sinuosidades na superfície de erro. Tal comportamento evita que o algoritmo fique preso a um pequeno mínimo local da superfície de erro, podendo mais facilmente chegar ao mínimo global (FIORIN, 2011).

O Gradiente Decrescente utiliza apenas a derivada de primeira ordem para buscar o alvo, mínimo erro. O gradiente é um vetor no ponto  $x$  que indica (localmente) a direção na qual a função  $f(x)$  mais cresce e é ortogonal (perpendicular) à tangente do contorno de  $f(x)$  no ponto  $x$ . A lei de ajuste do método do gradiente é dada por:

$$\theta_{i+1} = \theta_i - \alpha_i \frac{\nabla J(\theta_i)}{\|\nabla J(\theta_i)\|}. \quad (\text{eq. 2.6})$$

Onde  $\theta$  é o vetor de parâmetros,  $\alpha$  é o passo de ajuste e  $\nabla J$  o gradiente.

Os métodos do gradiente conjugado possuem sua estratégia baseada no modelo geral de otimização apresentado no algoritmo do gradiente, mas escolhem a direção de busca  $\mathbf{d}_i$ , o passo  $\alpha_i$  e o coeficiente de momento  $b_i$ . Sua taxa de convergência é quadrática e tem um desempenho superior ao método do gradiente decrescente sem aumentar significativamente o esforço computacional. Moller (1993) introduz uma nova variação no algoritmo de gradiente conjugado (Gradiente Conjugado Escalonado – SCG), que evita a busca unidimensional a cada iteração utilizando uma abordagem de Levenberg-Marquardt cujo objetivo é fazer um escalonamento do passo de ajuste  $\alpha$ . A idéia utilizada por Moller (1993) é estimar o termo

denominado  $s_j = \nabla^2 J(\theta_j) d_j$  do método do gradiente conjugado por uma aproximação da forma  $s_j = \nabla^2 J(\theta_j) d_j \approx [(\nabla J(\theta_j + \sigma_j d_j) - \nabla J(\theta_j))] / \sigma_j$ ,  $0 < \sigma_j \ll 1$ . O algoritmo utilizado neste método consta no Apêndice B.

## 2.3 MATERIAL E MÉTODO

A pesquisa foi iniciada a partir de revisão de literatura acerca da tecnologia solar fotovoltaica, especialmente em suas aplicações de caráter social, e nesse sentido teve como foco programas vinculados a políticas públicas de eletrificação rural. No percurso do estudo, foi percebida a problemática referente à demanda de energia elétrica em comunidades isoladas, as variações que esta demanda pode apresentar, e o insucesso que pode impor aos programas quando é tratada como constante no planejamento da eletrificação rural. A partir desta constatação a pesquisa passou a debruçar-se sobre a questão da demanda energética em comunidades isoladas e formas de se estabelecer um modelo de previsão adequado ao espectro de elementos envolvidos no consumo de energia elétrica, e para isso propõe-se uma análise comparativa entre processos de Regressão Linear Múltipla e Redes Neurais.

A hipótese inicial era a de que o número de cômodos na residência, nível de contato com o meio urbano, renda familiar, existência ou não de renda proveniente de atividade sazonal, escolaridade, número de adultos e número de crianças, bem como o tipo de sistema elétrico seriam fatores que impactariam o consumo de energia elétrica de famílias inseridas em contextos isolados do meio urbano, observando as indicações do trabalho de MORANTE (2004), e assim, estas formariam o conjunto de variáveis previsoras para modelos de regressão, bem como dados de entrada em se tratando de redes neurais.

Os dados acerca das comunidades do Vale do Ribeira - SP, Pedra Branca - PE e Vera Cruz - PA, todas contando com sistemas fotovoltaicos domiciliares, foram coletados a partir do trabalho de MORANTE (2004). As informações sobre a comunidade da Praia Vermelha - RJ, ligada à rede de distribuição de energia por meio de cabo submerso, fazem parte dos resultados da pesquisa de campo realizada no percurso deste estudo. Posteriormente, foram coletados também as mesmas informações sobre famílias urbanas, numa amostra aleatória de 15 famílias.

Os dados passaram por processo de normalização pelo método da faixa (apêndice) para que variáveis com ordens de grandeza diferentes não impusessem maior, ou menor

influência do que outras sobre o modelo, como seria o caso da renda familiar se comparada ao número de membros da família. A opção foi a da normalização através do método da faixa, que tem o objetivo de suprimir o problema de paralisia neural (HAYKIN, 2001). Assim, os dados de entrada são mapeados no intervalo [-1, 1].

Em princípio, foi estudada a correlação linear, por meio do coeficiente linear de Pearson (Apêndice B), entre as oito variáveis independentes constituintes da hipótese inicial, mais a variável dependente - consumo médio de energia familiar mensal (MORANTE, 2004). Quando a correlação entre duas variáveis é alta, presume-se que, mesmo conjuntamente, ocorra multicolinearidade entre variáveis, o que pode fazer com que o acréscimo dessas variáveis ao modelo de regressão não seja significativamente diferente de zero em termos de contribuição para o modelo. Estas avaliações podem ser confirmadas pela razão t (Apêndice B), analisando cada variável independente com relação ao consumo de energia. Nesse sentido, comparou-se a correlação linear, pelo valor de R (Apêndice B), com a estatística t e seu respectivo valor-P, decidindo pela permanência, ou não, da variável independente no conjunto de variáveis previsoras do consumo médio familiar, no modelo de regressão (MONTGOMERY, 2003). Para estas análises foram utilizados o pacote estatístico SPSS – IBM e o suplemento de análise de dados do MS Excel .

O modelo de regressão linear múltipla construído a partir do conjunto das oito variáveis hipoteticamente previsoras foi comparado a modelos em que se retirava variáveis pelos critérios de multicolinearidade e significância, verificando a eficiência do modelo na estimação do consumo de energia pelo coeficiente de determinação,  $R^2$  (Apêndice B).

Num segundo momento, os mesmos dados foram processados no software SPSS, utilizando ferramenta para RNA, o que possibilitou a comparação entre os dois métodos quanto sua eficiência na previsão de demanda energética por família. Em seguida amostras com acréscimo de ruídos foram submetidas a mesma configuração de RNA que a amostra composta pelos dados de campo. Diversas configurações de rede e parâmetros de treinamento foram testados com objetivo de otimizar a função de erro nas saídas da rede - consumo médio por família e seu desvio padrão. As redes que apresentaram os melhores desempenhos são descritas na sessão de resultados, assim como os modelos de regressão.

## 2.4 RESULTADOS E DISCUSSÕES

### Comunidades do Vale do Ribeira (SP) – eletrificada com SFD<sup>8</sup>

A maior parte das famílias desta região mantém o modo de vida característico da cultura caiçara. Apesar das grandes influências externas sofridas nos últimos anos, estas peculiaridades permanecem e se originam de seu grande apego ao mar, como acontece com as comunidades de Sítio Artur, Itapanhapina e Marujá e, além disso, com a terra, tal como se verifica na comunidade de Varadouro. Na Figura 2.4-1 podem ser observados aspectos do dia-a-dia das pessoas destas comunidades.



Figura 2.4-1 Imagens da Comunidade Marujá, região do lagamar, Vale do Ribeira. (MORANTE, 2004).

### Comunidade Pedra Branca – PE eletrificada com SFD

Pedra Branca é caracterizada por caatinga, sofrendo efeitos da falta de chuvas, a comunidade convive com todos os problemas derivados da seca. A maioria dos habitantes subsiste realizando atividades agropecuárias, Figura 2.4-2. Nas temporadas favoráveis, plantam, entre outros produtos, milho, feijão ou mandioca, criam caprinos, o que representa uma das principais fontes de renda e de proteína das pessoas. Além dessas atividades, muitas das famílias dependem também dos rendimentos econômicos derivados das aposentadorias dos mais idosos. Os principais problemas da comunidade são a falta de água para uso produtivo (irrigação e criação), falta de água para consumo doméstico, péssimas condições de ensino, desemprego e falta de condições financeiras, falta de assistência à saúde, dificuldade de comunicação, falta de transporte e falta de trator (MORANTE, 2004).

---

<sup>8</sup> Sistema Fotovoltaico Domiciliar



Figura 2.4-2 Casa e animais de criação típicos da região da caatinga em Pedra Branca (MORANTE, 2000).

### Comunidade De Vera Cruz - PA eletrificada com SFD

A comunidade está inserida no ambiente amazônico, de tal forma que seu estilo de vida está modelado segundo esse entorno. Isto se manifesta na construção das moradias, utilizando materiais do próprio local, nas atividades econômicas de subsistência, nos meios de transporte de maior utilização que são as canoas, e outros, Figura 2.4-3.



Figura 2.4-3 Comunidade Vera Cruz (MORANTE, 2004).

A configuração familiar, em geral, se estabelece sobre relações de solidariedade associada ao parentesco e ao afastamento dentro da floresta e, em certa medida, pela influência dos preceitos religiosos impostos pela Irmandade da Santa Cruz (MORANTE, 2004).

### Praia Vermelha – RJ

A comunidade da Praia Vermelha, Ilha Grande, no município de Angra dos Reis fica a duas horas e meia de barco da costa, Figura 2.4-4. Não conta com transporte regular, mas com



embarcações que destinando-se a outras praias podem aportar, se solicitadas, na Praia Vermelha. A comunidade, atualmente, constitui-se de famílias que em sua maioria tem sua sobrevivência associada a atividades turísticas. Poucas são as casas de pescadores e apesar da história da Ilha, não se entendem como caiçaras.



Figura 2.4-4 Ilha Grande, Praia Vermelha (MIRANDA, 2004).

Por meio de cabos submersos a eletrificação chegou a Praia Vermelha, há mais de cinco anos. A comunidade possui um importante número de homens trabalhando como marinheiros, atividade que os mantém em contato com a zona urbana, constantemente, percebendo-se claramente a influência desse contato sobre o perfil de consumo das famílias pesquisadas. O nível de renda é razoável considerando-se uma comunidade com essas características geográficas, entre dois e três salários mínimos.

. Nesse contexto, sete famílias foram pesquisadas, constituindo uma amostra aleatória do conjunto de moradores, em torno de 15 famílias. Cabe ressaltar que essa comunidade não sofreu processo de pré-eletrificação como algumas comunidades da Baía da Ilha Grande e sua demanda encontra-se de certa forma estabilizada, dado o tempo de que já dispõe da energia elétrica de forma intermitente, sendo latente a distinção com relação a comunidades abastecidas por SFD no que se refere ao nível de consumo de energia elétrica.

## 2.4.1 APLICAÇÃO DE REGRESSÃO LINEAR MÚLTIPLA

### **Estudo sobre amostra formada por famílias de comunidades isoladas (SFD e rede de distribuição elétrica)**

Considerando a hipótese de que pelo menos oito variáveis influenciavam a variação no consumo de energia elétrica familiar, conforme estudo de Morante (2004), fez-se um estudo de correlação para verificar colinearidades. Os coeficientes de correlação linear de Pearson (Apêndice B) entre as oito variáveis predictoras e o consumo de energia elétrica foram analisados para se presumir possíveis colinearidades entre as variáveis hipoteticamente associadas ao consumo de energia familiar, onde as variáveis predictoras são:

- Cômodos: número de cômodos da residência
- Contato urbano: nível de contato com o meio urbano, em escala crescente de 1 a 5.
- Escolaridade: maior escolaridade entre os responsáveis pela família, de zero a 9 (analfabeto – pós-graduação)
- Sazonal: ocorrência ou não de rendimento proveniente de atividade sazonal (0 ou 1)
- Renda Familiar: renda média familiar, em dólares americanos
- Adultos: número de adultos na família, considerados a partir de 13 anos de idade
- Crianças: número de crianças na família, até 12 anos de idade
- Sistema: tipo de sistema de geração de energia elétrica (fotovoltaico = 0, ou rede de distribuição = 1)

Dessa análise observou-se a alta correlação entre renda familiar (renda fam) e o tipo de sistema elétrico (sistema),  $R$  igual a 0,89. Mesmo ciente de que os sistemas não foram instalados às expensas das famílias, a renda pode estar associada, de fato, ao tipo de sistema elétrico, pois a escolha pela implementação da tecnologia pode refletir o maior, ou menor, grau de afastamento do meio urbano e conseqüente tipo de atividade geradora de renda e o próprio nível de renda.

A amostra passou por processo de regressão linear múltipla, contemplando todas as oito variáveis independentes, e em seguida somente com as variáveis renda, sistema, adultos e crianças, as quais não foram eliminadas pelo estudo de colinearidade.

O resumo de seu desempenho consta no Quadro 2.4-1. O objetivo foi o de confirmar por meio de testes, o nível de significância do modelo com oito variáveis predictoras e a significância dos coeficientes parciais, pela razão  $t$  (Apêndice B) e seu correspondente valor- $P$ , definido como a menor escolha que teríamos feito para o nível de significância, de forma que rejeitaríamos  $H_0$  (Apêndice B). Os valores- $P$ , Quadro 2.4-1, foram analisados para a

definição acerca da permanência de cada variável no conjunto de previsoras, e ao se verificar a significância de cada variável com relação a sua contribuição para a estimação do consumo de energia elétrica, a única variável que apresentou valor-P indicando que o parâmetro correspondente é significativamente diferente de zero, ou seja, negando a hipótese nula de que seja igual a zero, é a variável que representa o tipo de sistema elétrico instalado na residência, a nível de confiança de 0,95.

Pelos resultados apresentados no Quadro 2.4-1, pode-se observar que o modelo com as oito variáveis previsoras apresenta o melhor percentual de explicação na variação do consumo de energia elétrica, porém a única variável significativa (valor-P < 0,05) é sistema (fotovoltaico, ou rede de distribuição elétrica). Sistema é uma variável que assume somente dois valores, portanto é preciso que se associe a ela pelo menos uma outra variável significativa, para que o modelo retorne valores suficientemente variados, como cabe a uma previsão de demanda energética. As variáveis renda e sistema apresentam correlação bastante alta entre si, mas quando estas duas variáveis formam o conjunto de variáveis previsoras, o modelo perde somente um ponto percentual no coeficiente de determinação<sup>9</sup> e as duas variáveis tornam-se significativas, fato observado pelo valor-P (Quadro 2.4-1), portanto este modelo mostra-se mais significativo e com menor custo de construção.

**Quadro 2.4-1 Estatística dos modelos de regressão com oito, duas e uma variável previsoras. Elaborada pelo autor.**

MODELO	VARIÁVEIS PREVISORAS	R <sup>2</sup>	Valor-P < 0,05
1	comodos, contato urbano, escolaridade, sazonal, renda, adultos, crianças, sistema	94%	sistema
2	renda, crianças	72%	renda
3	renda, sistema	93%	renda, sistema
4	renda	76%	renda

### **Estudo sobre os dados de comunidades isoladas e famílias urbanas**

Foram acrescentadas à amostra de famílias isoladas, quinze famílias urbanas, com o objetivo de se verificar a possibilidade de se estender o uso dos modelos de regressão e redes neurais a estes dois ambientes.

<sup>9</sup> Coeficiente de determinação, dado por equação a ser verificada no Apêndice B



Percorrendo os mesmos caminhos utilizados com as famílias de comunidades isoladas, submeteu-se uma amostra de 50 famílias entre rurais e urbanas ao processo de regressão linear. Os resultados estão registrados no Quadro 2.4-2. A eficiência, do modelo medida pelo  $R^2$ , perde 9%, considerando o modelo com oito variáveis predictoras, e somente a variável renda apresenta influência significativa sobre a variação do consumo de energia, conforme valor-P (Quadro 2.4-2). Considerando-se o mais eficiente, para essa amostra, Modelo 6 (Quadro 2.4-2), com 81% de coeficiente de determinação e suas duas variáveis predictoras com valores-P indicando significância, se comparado ao modelo com as mesmas variáveis, mas modelado a partir de dados somente de comunidades isoladas (Modelo 3, Quadro 2.4-1), este perde doze pontos percentuais, em sua capacidade de explicação.

Os outros modelos, com quatro e uma variável predictoras, também perdem em sua capacidade de explicação na variação do consumo de energia elétrica, quando os dados são de uma amostra mais heterogênea, o que pode apontar para a dificuldade do método de regressão linear em aproximar dados de conjuntos não homogêneos, adicionado ao fato de que a regressão linear permite apenas uma variável de saída, no caso o consumo, ou demanda energética familiar.

**Quadro 2.4-2 Estatística dos modelos de regressão com oito, duas e uma variável predictoras., para amostra de 50 famílias, rurais e urbanas. Elaborada pelo autor.**

Modelo	VARIÁVEIS PREVISORAS	$R^2$	Valor-P < 0,05
5	comodos, contato urbano, escolaridade, sazonal, renda, adultos, crianças, sistema	85%	renda
6	renda, sistema, adultos, crianças	81%	renda sistema
7	renda, sistema	81%	renda, sistema
8	renda	72%	renda

## 2.4.2 APLICAÇÃO DE REDES NEURAIIS

### Estudo sobre amostra constituída por famílias de comunidades isoladas

O primeiro passo foi o estabelecimento da arquitetura da rede, testando empiricamente o número de variáveis de entrada e de neurônios na camada oculta, pois uma única camada

oculta dá conta da maioria das funções contínuas (HAYKIN, 2001). Os critérios de eficiência adotados foram a soma dos erros quadráticos (Apêndice A) e o coeficiente de determinação entre os consumos de energia elétrica familiar reais e os previstos pela RNA. A melhor arquitetura para a MLP foi estabelecida com oito variáveis de entrada, as mesmas previsoras da regressão linear, cinco neurônios na camada oculta e duas variáveis de saída, consumo de energia elétrica familiar e seu desvio padrão.

Definida a arquitetura foram testadas funções de ativação para a camada oculta e funções para a camada de saída. Nas duas camadas a função que apresentou melhor desempenho foi a sigmóide, a qual costuma ser a mais empregada nas camadas internas de uma MLP por serem contínuas, crescentes, diferenciáveis e não-lineares (CASTRO, 1998).

Nesta etapa, o algoritmo de otimização dos erros gerados na camada oculta tinha sido, em todos os testes, o gradiente decrescente e mesmo alterando-se taxa de aprendizagem, termo de momentum e épocas de treinamento, o coeficiente de determinação entre consumos reais e previstos pela RNA não passava de 80%, ou seja, a rede não era capaz de explicar mais do que 80% da variação no consumo de energia elétrica, considerando amostra de comunidades isoladas e uma amostra heterogênea (famílias de comunidades isoladas e urbanas).

Ao adotar o gradiente conjugado escalonado como algoritmo de otimização, os valores de  $R^2$  passaram a ser superiores a 90%, na maioria das redes testadas, e mesmo quando era considerado um número menor de variáveis de entrada, a eficiência da rede mostrava-se superior a de uma rede com a mesma configuração, porém com o algoritmo de otimização de erros gradiente decrescente. Portanto, a escolha do algoritmo de otimização dos erros na camada oculta foi determinante na definição da melhor configuração de rede.

Para analisar a capacidade da RNA de adaptar-se foram submetidas a mesma configuração de rede, uma amostra somente de famílias que possuíam sistemas fotovoltaicos, e em seguida a uma amostra heterogênea constituída por famílias eletrificadas com SFD, rede de distribuição residentes em comunidade isoladas e por famílias urbanas. Conforme consta no Quadro 2.4-3, o desempenho de redes com a mesma configuração aplicadas às três amostras foi o mesmo, ou seja, as oito variáveis de entrada explicam 99% da variação no consumo de energia elétrica para os três tipos de amostras. A configuração de rede a que foram submetidas as três amostras possui as características descritas no Quadro 2.4-4.

. O bom desempenho das RNA pode revelar que a heterogeneidade das amostras não compromete sua eficiência e ao sinalizar que um menor número de variáveis de entrada é que

pode diminuir sua eficácia (Quadro 2.4-3), corrobora a tese de Morante (2004) que entende o consumo de energia elétrica delineado por um conjunto de variáveis complexas de diversos âmbitos, como social e cultural, psicológico, além do econômico.

**Quadro 2.4-3 Estatística dos modelos neurais com oito, duas e uma variável previsoras, para amostra de 50 famílias, rurais e urbanas. Elaborada pelo autor.**

VARIÁVEIS DE ENTRADA	AMOSTRA	SEQ	R <sup>2</sup>
comodos, contato urbano, escolaridade, sazonal, renda familiar, adultos, crianças, sistema	SFD	0,052	99%
	Isoladas	0,005	99%
	Isoladas e urbana	0,083	99%
renda familiar, sistema	SFD	0,812	46%
	Isoladas	0,281	93%
	Isoladas/urbanas	0,76	93%

**Quadro 2.4-4 Configuração de rede que apresentou os melhores desempenhos para as três amostras submetidas.**

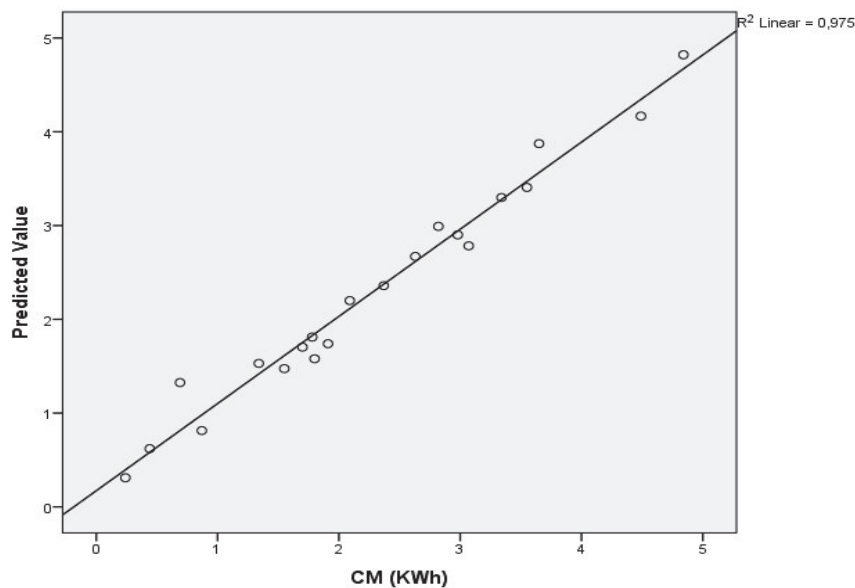
ARQUITETURA	TREINAMENTO	CRITÉRIOS DE PARADA
8 variáveis de entrada, 5 neurônios na camada oculta e duas saídas, função de ativação na camada oculta e função na camada de saída sigmóide.	Apresentação dos dados à rede em lote, algoritmo de otimização de erros gradiente conjugado	Nesta ordem: uma etapa sem redução no erro, 100 etapas de treinamento, alteração relativa mínima no erro de treinamento de 0,0001.

Pela própria natureza do problema, que requer dados de entrada reais, que reflitam especificidades observadas do microcosmo de cada família, houve dificuldade em separar uma amostra de teste, especialmente no que se refere à pequena dimensão das amostras. A escolha foi pela partição a partir de uma semente aleatória, a qual atribui valores a cada elemento da amostra, criando assim partições de treinamento e teste, e mesmo particionando a amostra, os erros mantiveram-se pequenos e a correlação entre dados reais de consumo e os previstos permaneceu bastante alta, conforme Quadro 2.4-5, o que reitera a capacidade de adaptação das redes neurais testadas e aponta para uma boa chance de generalização para amostras maiores e heterogêneas, implicando assim em modelos que podem atender às necessidades de previsão de consumo de energia elétrica tanto no âmbito rural, quanto urbano residencial.

**Quadro 2.4-5 Desempenho das redes a partir de amostras particionadas em amostra de treinamento e teste.**

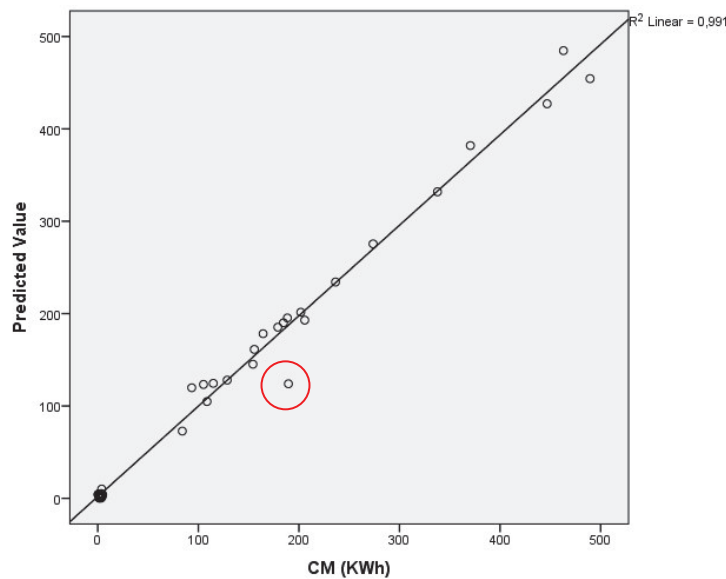
VARIÁVEIS DE ENTRADA	AMOSTRA	SEQ	R <sup>2</sup>
comodos, contato urbano, escolaridade, sazonal, renda familiar, adultos, crianças, sistema	SDF	0,041	98%
	Isoladas	0,003	99%
	Isoladas e urbana	0,057	99%

A amostra mais restrita, que somente inclui famílias de comunidades isoladas e abastecidas por SFD, mesmo com número reduzido de elementos (28 famílias) e particionada em amostra de treinamento e teste obteve resultados bastante semelhantes aos testes com as demais amostras, conforme Figura 2.4-5, com coeficiente de determinação de, aproximadamente, 98%, submetida à idêntica configuração de rede. Essa variação de amostras revelando resultados muito semelhantes pode ser também um indicativo de que a configuração apresentada, nesse estudo, de rede neural é bastante eficiente na previsão do consumo de energia elétrica familiar.



**Figura 2.4-5 Correlação entre consumo real por família e a demanda prevista. Elaborada pelo autor. Saída SPSS-IBM.**

Como a natureza da problemática da demanda energética requer que se atenda, no mínimo, anecessidades básicas da família, é mais adequado que se aproxime essa demanda por valores maiores que os reais, nesse sentido a Figura 2.4-6 exemplifica a correlação entre valores de energia consumida e demanda prevista, pelo coeficiente de determinação, para amostra formadas por famílias de comunidades isoladas e famílias urbanas.



**Figura 2.4-6** Correlação entre valores reais de consumo e demanda prevista, em kWh, por família.  
Elaborado pelo autor. Saída do SPSS-IBM.

Pode ser observado na Figura 2.4-6 que algumas previsões aproximaram-se do consumo real por valores menores, como é o caso do ponto destacado em vermelho. Para resolução desse problema pesquisas posteriores podem investigar formas de se adicionar ao modelo de rede neural um artifício que implique sempre em aproximações por valores maiores que os de consumo reais, talvez utilizando internamente no modelo um desvio padrão médio, considerando-se que a configuração de rede proposta neste trabalho tem como saídas o consumo previsto e seu correspondente desvio-padrão.

### Aplicação de RNA sobre amostras com ruído randômico

Com o objetivo de se verificar se a configuração da RNA treinada com os dados de campo seria adequada para outras amostras, gerou-se duas amostras de 50 exemplos, com ruídos de 1% e 5% com relação ao conjunto de dados reais, aplicados a todas as entradas da RNA, conforme a equação  $Z_i = X_i + n.X_i.r$ , onde  $Z_i$  é o dado ruidoso,  $X_i$  o valor real da entrada,  $n$  um número aleatório entre zero e 1 e  $r$  o percentual de ruído desejado.

As novas saídas da rede, consumos de energia elétrica familiar, foram gerados por meio de equação linear com termo independente igual a zero, cujos coeficientes foram obtidos na aplicação da regressão linear múltipla sobre os dados de campo.

A resposta da rede permanece apresentando soma dos erros quadráticos na casa dos milésimos,  $R^2$  em torno de 99% entre os consumos submetidos à rede e os previstos, tanto na amostra com ruído de 1%, quanto na amostra com 5%, o que pode apontar para eficiência

dessa configuração de RNA para diferentes amostras, revelando sua capacidade de generalização.

## 2.5 CONCLUSÃO

Os estudos mostraram que os melhores resultados referentes à previsão de consumo de energia elétrica familiar foram alcançados quando oito características das famílias foram consideradas como variáveis de entrada dos modelos, tanto de regressão linear, quanto de redes neurais artificiais. A capacidade dos modelos de regressão e RNA de realizarem boas previsões de consumo de energia elétrica familiar, a partir deste conjunto de oito variáveis de entrada, valida a hipótese inicial registrada no trabalho de Morante (2004), de que a demanda energética familiar depende de um espectro de características envolvidas no microcosmo das famílias e não é, portanto, uma constante conforme vem sendo tratada em programas de eletrificação rural.

As RNA mostraram-se mais eficientes que os modelos de regressão linear na previsão do consumo de energia elétrica familiar, especialmente quando a amostra torna-se mais heterogênea, contemplando tanto famílias que contam somente com sistemas fotovoltaicos, quanto amostras constituídas de famílias rurais e urbanas, o que pode estar associado a sua capacidade de generalização, pois mesmo para amostras acrescidas de ruídos a mesma eficiência da configuração de RNA utilizada com os dados de campo foi mantida. Os testes revelaram também o importante impacto do número de variáveis de entrada sobre o desempenho das redes, os quais foram melhores quando o conjunto de oito variáveis de entrada foi implementado. A capacidade de redes submetidas a essas oito variáveis de entrada de explicarem a variação no consumo de energia elétrica, ultrapassando 99% implica em considerar-se as RNA como um método bastante eficaz para previsão do consumo de energia elétrica familiar, seja para comunidades isoladas, ou urbanas, podendo subsidiar tanto o planejamento de projetos vinculados a políticas públicas voltadas a comunidades rurais, quanto projetos que visem incentivar o uso de fontes alternativas de energia no meio urbano.

Como forma de se aperfeiçoar o modelo em trabalhos futuros, este estudo aponta para a inclusão de um artifício matemático que faça com que a rede sempre aproxime a demanda energética por valores maiores dos que são submetidos a ela, indicando a possibilidade de

utilizar-se o desvio padrão médio para este fim. Ainda com relação a estudos subseqüentes, é importante salientar que o aumento da amostra de forma aleatória pode requerer a inclusão de variáveis de entrada como a temperatura média mensal da localidade, pois esta segundo Biondi Neto (2008) e Morante (2004) se configura como importante fator na decisão pelo consumo de energia elétrica.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BIONDI NETO, L. et al. Estudo não paramétrico da relação entre consumo de energia, renda e temperatura. IEEE Latin America Transactions, Vol. 6, n<sup>o</sup>. 2, Junho, 2008.
- BRAGA, Antônio de P., CARVALHO, André P. de L, LUDERMIR, T.B.. Redes Neurais Artificiais. Editora LTC. São Paulo, 2007.
- CARPINTEIRO, O. A. S. ; REIS, A. J. R. . A hierarchical self-organizing map model in short-term load forecasting. In: Congresso Brasileiro De Automática (Cba), 2004, Gramado, RS. Proceedings. v. 1.
- CASTRO, L.N., Análise e Síntese de Estratégias de Aprendizado para Redes Neurais Artificiais. Campinas: FEEC, UNICAMP, Setembro de 1998. Dissertação de Mestrado – Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, pp. 250.
- CATALAO, João, MARIANO, Sílvio, MENDES, Victor et al. Previsão dos Preços da Energia Elétrica através de Redes Neurais Artificiais. Inv. Op., dez. 2007, vol.27, n<sup>o</sup>.2, p.151-163.
- FIORIN, Daniel V. et al . Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares. Rev. Bras. Ensino Fís., São Paulo, v. 33, n. 1, Mar. 2011. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1806-11172011000100009&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-11172011000100009&lng=en&nrm=iso)>. Acesso em: 06/02/2011.
- HAYKIN, Simon. Redes Neurais. Princípios e prática. Porto Alegre, RS: Bookman, 2001.
- LUGON, Jader Jr., et al.. Application of Artificial Neural Networks and Hybrid Methods in the Solution of Inverse Problems. In: HUI, C. L. P.. Artificial Neural Networks - Application. InTech, 2011. p.542-566. Disponível em: <http://www.intechopen.com/articles/show/title/application-of-artificial-neural-networks-and-hybrid-methods-in-the-solution-of-inverse-problems>
- MOLLER M.F., A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning,. Neural Networks, vol. 6, pp. 525-533, 1993. Disponível em <http://sci2s.ugr.es/keel/pdf/algorithm/articulo/moller1990.pdf>. Acessado em fevereiro de 2013.
- MORANTE, Federico B. T. Demanda de energia elétrica e desenvolvimento socioeconômico: o caso das comunidades rurais eletrificadas com sistemas fotovoltaicos. 311 p., 2004 Tese (Doutorado) – Programa Interunidades de Pós-Graduação em Energia da Universidade de São Paulo.
- PASCHOALINO, F. F. et al. Previsão de demanda de energia elétrica no Brasil utilizando redes neurais de Elman. SPOLM 2007 RJ. Disponível em: <[http://www.uff.br/decisao/spolm2007\\_133.pdf](http://www.uff.br/decisao/spolm2007_133.pdf)>. Acesso em: 08/12/2009.



PEDROLLO. Previsão de níveis fluviais com redes neurais – aplicação para Rosário de Sul – RS. In: 1º. Simpósio de Recursos Hídricos do Sul. Curitiba, PR. 2006. Anais Associação Brasileira de recursos Humanos, 2006. CD ROM.

RUAS, Gabriel I. S. et al. Previsão de Demanda de Energia Elétrica Utilizando Redes Neurais Artificiais e Support Vector Regression. 2007. ENIA, Rio de Janeiro, 2007.

TEIXEIRA, Luiz Albino. Combinação geométrica de método preditivos; aplicação à previsão de consumo residencial mensal de energia elétrica. Rio de Janeiro, 2009. 115p. Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

## APÊNDICES

## APÊNDICE A – REDES NEURAS ARTIFICIAIS

### 1. RESULTADOS PARA AMOSTRA DE COMUNIDADES ISOLADAS ABASTECIDAS POR SFD, SOMENTE TREINAMENTO (RN\_1)

**TabelaA-1 Programação do modelo RN\_1**

Notes		
Output Created		12-JUN-2012 19:57:22
Comments		
Input	Data	C:\Users\ROBERTA & EVERTON\Desktop\JADER_11-04-12\DADOS_SDF_DOLAR_8X2.sav
	Active Dataset	Conjunto_de_dados1
	Filter	<none>
	Weight	<none>
	Split File	<none>
Missing Value Handling	N of Rows in Working Data File	28
	Definition of Missing Cases Used	User- and system-missing values are treated as missing.
Weight Handling		Statistics are based on cases with valid data for all variables used by the procedure. not applicable
Syntax		MLP CMKWh (MLEVEL=S) DPKWh (MLEVEL=S) WITH comodos contatourb escolar sazonal RendaFamUS\$ adultos crianças sistema
		/RESCALE COVARIATE=ADJNORMALIZED DEPENDENT=NORMALIZED (CORRECTION=0.02)
		/PARTITION TRAINING=7 TESTING=0 HOLDOUT=0
		/ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=6) HIDDENFUNCTION=SIGMOID OUTPUTFUNCTION=SIGMOID
		/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005 SIGMAININITIAL=0.00005 INTERVALCENTER=0
		INTERVALOFFSET=0.5 MEMSIZE=1000
		/PRINT CPS NETWORKINFO SUMMARY CLASSIFICATION SOLUTION IMPORTANCE
		/PLOT NETWORK PREDICTED
		/SAVE PREDVAL
		/STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=200 ERRORCHANGE=1.0E-4
Resources	Processor Time	ERRORRATIO=0.0010
	Elapsed Time	/MISSING USERMISSING=EXCLUDE .
Variables Created or Modified	Predicted Value	00:00:01,09
		00:00:01,09
	CMKWh	MLP_PredictedValue_1_A
	DPKWh	MLP_PredictedValue_2_A

[Conjunto\_de\_dados1] C:\Users\ROBERTA & EVERTON\Desktop\JADER\_11-04-12\DADOS\_SDF\_DOLAR\_8X2.sav

Tabela A-2 Particionamento da mostra e configuração da RN\_1

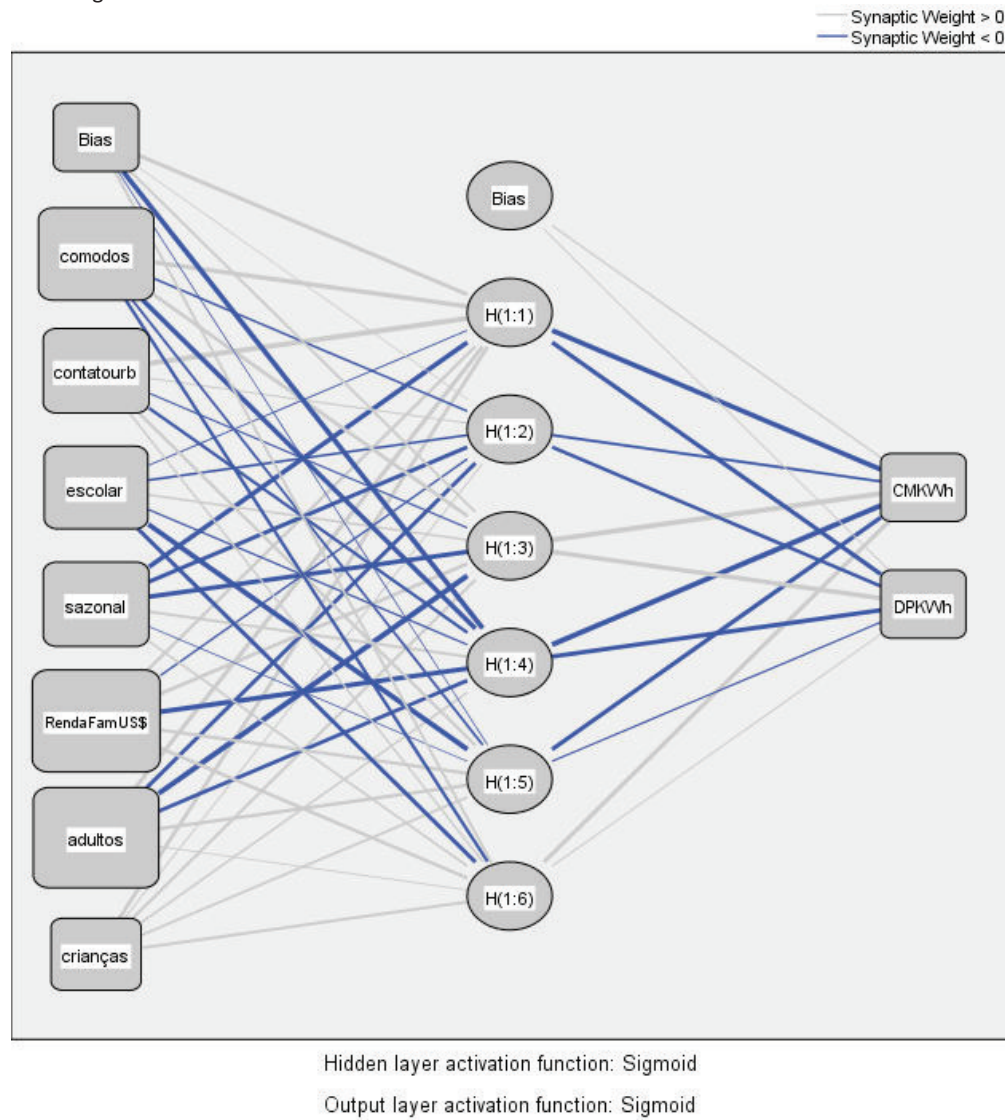
## Case Processing Summary

	N	Percent
Sample Training	28	100,0%
Valid	28	100,0%
Excluded	0	
Total	28	

## Network Information

		1	comodos
		2	contato urb
		3	escolar.
	Covariates	4	sazonal
Input Layer		5	Renda Fam US\$
		6	adultos
		7	crianças
	Number of Units <sup>a</sup>	7	
	Rescaling Method for Covariates		Adjusted normalized
	Number of Hidden Layers	1	
Hidden Layer(s)	Number of Units in Hidden Layer 1 <sup>a</sup>	6	
	Activation Function		Sigmoid
	Dependent Variables	1	CM (KWh)
		2	DP (KWh)
Output Layer	Number of Units	2	
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares

Excluding the bias unit



**Ilustração A-1 Topologia da RN\_1**

Tabela A-3 Medidas de desempenho da RN\_1

## Model Summary

	Sum of Squares Error		,052
	Average Overall Relative Error		,035
Training	Relative Error for Scale Dependents	CM (KWh)	,016
		DP (KWh)	,073
	Stopping Rule Used	Maximum number of epochs (200) exceeded	
	Training Time	0:00:00,01	

Tabela A-4 Pesos sinápticos da RN\_1

## Parameter Estimates

Predictor	Predicted								
	Hidden Layer 1						Output Layer		
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	CMKWh	DPKWh	
Input Layer	(Bias)	2,558	,131	1,083	-4,873	-,092	1,128		
	comodos	4,016	-,896	2,312	-4,377	-1,004	-2,015		
	contatourb	5,954	,455	-,642	-1,801	1,552	,869		
	escolar	-,315	-1,004	,971	-,738	-5,265	-2,595		
	sazonal	-4,348	-2,604	-4,052	1,283	-,036	,978		
	RendaFamUS\$	1,950	-,847	2,268	-4,674	2,030	2,035		
	adultos	2,943	-2,574	-6,746	-2,430	2,095	,286		
	crianças	2,339	,412	1,336	,962	1,060	1,511		
Hidden Layer 1	(Bias)						,798	,317	
	H(1:1)						-4,739	-3,501	
	H(1:2)						-1,196	-2,243	
	H(1:3)						5,923	4,054	
	H(1:4)						-6,150	-3,615	
	H(1:5)						-3,295	-,695	
H(1:6)						3,143	,402		

### Predicted by Observed Charts

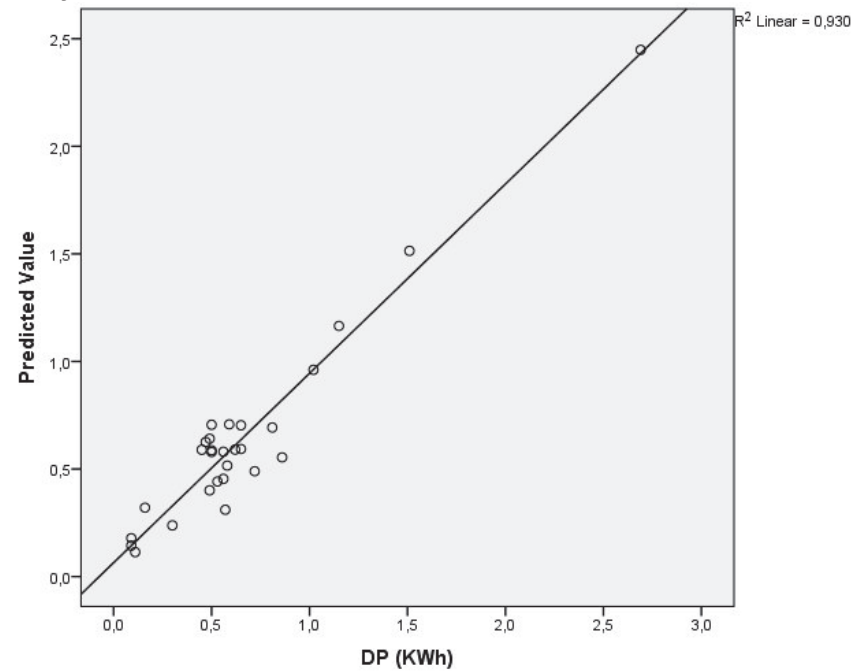


Ilustração A-2 Ajuste Linear de dados de consumo de energia elétrica familiar real e previstos pela RN\_1

Tabela A-5 Importância das variáveis independentes RN\_1

#### Independent Variable Importance

	Importance	Normalized Importance
comodos	,153	85,5%
contato urb	,133	74,3%
escolar.	,128	71,6%
sazonal	,133	74,0%
Renda Fam US\$	,179	100,0%
adultos	,174	97,3%
crianças	,100	55,7%

## 2. RESULTADOS PARA AMOSTRA DE COMUNIDADES ISOLADAS ABASTECIDAS POR SFD, PARTICIONADA EM TREINAMENTO E TESTE (RN\_2)

Tabela A-6 Programação do modelo RN\_2

		Notes	
Output Created			12-JUN-2012 22:24:03
Comments			
Input	Data Active Dataset Filter Weight Split File N of Rows in Working Data File	C:\Users\ROBERTA & EVERTON\Desktop\JADER_11-04-12\DADOS_SDF_DOLAR_8X2.sav Conjunto_de_dados1 <none> <none> <none>	
Missing Value Handling	Definition of Missing Cases Used	User- and system-missing values are treated as missing. Statistics are based on cases with valid data for all variables used by the procedure.	28
Weight Handling		not applicable MLP CMKWh (MLEVEL=S) DPKWh (MLEVEL=S) WITH comodos contatourb escolar sazonal RendaFamUS\$ adultos crianças sistema /RESCALE COVARIATE=ADJNORMALIZED DEPENDENT=NORMALIZED (CORRECTION=0.02) /PARTITION VARIABLE=PARTIÇÃO_0.8 /ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=5) HIDDENFUNCTION=SIGMOID OUTPUTFUNCTION=SIGMOID	
Syntax		/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005 SIGMAINITIAL=0.00005 INTERVALCENTER=0 INTERVALOFFSET=0.5 MEMSIZE=1000 /PRINT CPS NETWORKINFO SUMMARY CLASSIFICATION SOLUTION IMPORTANCE /PLOT NETWORK PREDICTED /STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=200 ERRORCHANGE=1.0E-4 ERRORRATIO=0.0010 /MISSING USERMISSING=EXCLUDE .	
Resources	Processor Time		00:00:01,08
	Elapsed Time		00:00:01,07

[Conjunto\_de\_dados1] C:\Users\ROBERTA & EVERTON\Desktop\JADER\_11-04-12\DADOS\_SDF\_DOLAR\_8X2.sav



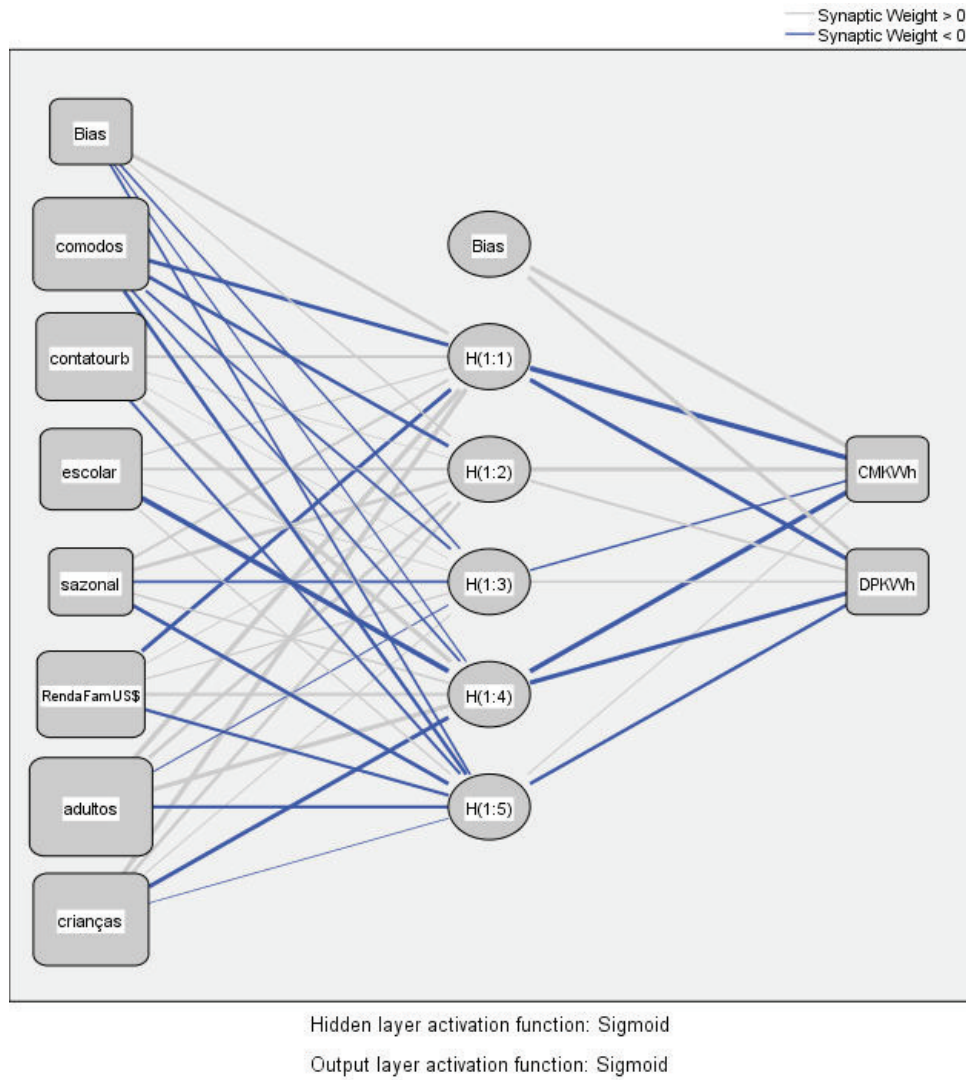
TabelaA-7 Particionamento da mostra e configuração da RN\_2

## Case Processing Summary

		N	Percent
Sample	Training	21	75,0%
	Holdout	7	25,0%
Valid		28	100,0%
Excluded		0	
Total		28	

## Network Information

Input Layer		1	comodos	
		2	contato urb	
		3	escolar.	
	Covariates	4	sazonal	
		5	Renda Fam	US\$
		6	adultos	
		7	crianças	
Hidden Layer(s)	Number of Units			7
	Rescaling Method for Covariates		Adjusted normalized	
	Number of Hidden Layers			1
	Number of Units in Hidden Layer 1			5
Output Layer	Activation Function		Sigmoid	
	Dependent Variables	1	CM (KWh)	
		2	DP (KWh)	
	Number of Units			2
	Rescaling Method for Scale Dependents		Normalized	
	Activation Function		Sigmoid	
	Error Function		Sum of Squares	



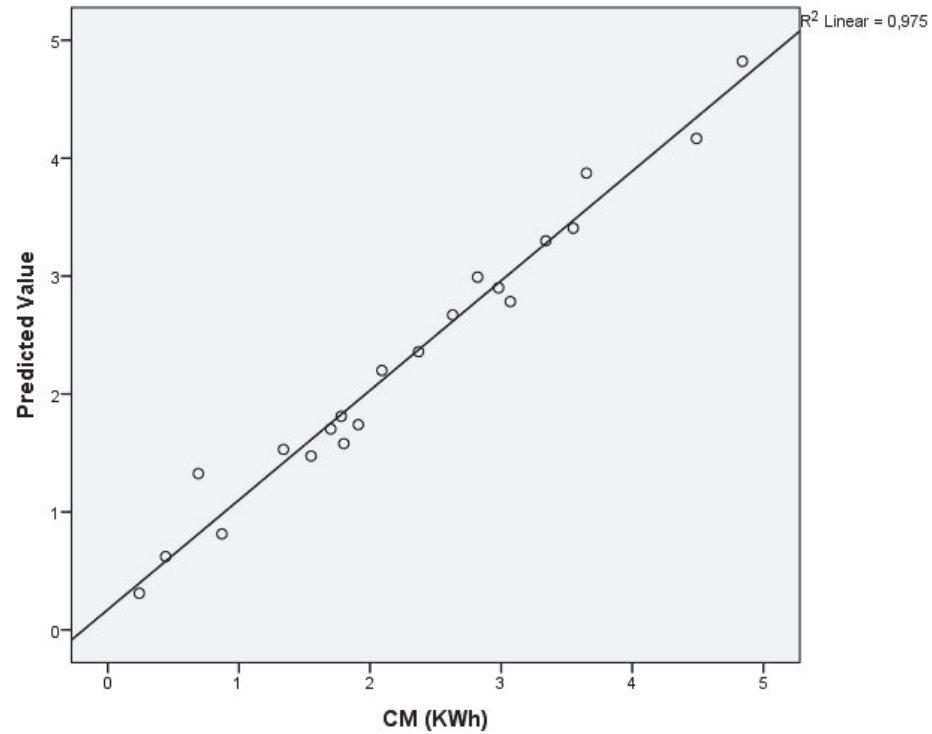
**Ilustração A-3 Topologia da RN\_2**

Tabela A-8 Medidas de desempenho da RN\_2

Model Summary			
Training	Sum of Squares Error		,041
	Average Overall Relative Error		,035
	Relative Error for Scale	CM (KWh)	,027
	Dependents	DP (KWh)	,048
	Stopping Rule Used	Maximum number of epochs (200) exceeded	
	Training Time		0:00:00,01
Holdout	Average Overall Relative Error		,619
	Relative Error for Scale	CM (KWh)	,523
	Dependents	DP (KWh)	,920

Tabela A-9 Pesos sinápticos da RN\_2

Predictor	Predicted						
	Hidden Layer 1					Output Layer	
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	CMKWh	DPKWh
(Bias)	3,728	,991	-1,167	-,668	-1,460		
comodos	-4,527	-3,637	-1,521	-1,449	-3,207		
contatourb	1,999	,293	,253	5,456	-1,659		
escolar	,566	1,784	,397	-6,535	,711		
sazonal	1,851	3,968	-1,562	1,347	-3,923		
RendaFamUS\$	-4,426	,407	,678	2,446	-2,301		
adultos	6,351	3,674	-,426	5,381	-2,509		
crianças	6,098	2,780	,472	-4,535	-,268		
(Bias)						6,082	4,551
H(1:1)						-7,425	-4,883
H(1:2)						4,596	1,625
H(1:3)						-1,227	1,425
H(1:4)						-6,205	-5,181
H(1:5)						,550	-3,026



**Ilustração A-4** Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN\_2

**Tabela A-10** Importância das variáveis independentes RN\_2

Independent Variable Importance		
	Importance	Normalized Importance
comodos	,159	90,9%
contato urb	,146	83,5%
escolar.	,129	74,0%
sazonal	,091	52,4%
Renda Fam US\$	,142	81,4%
adultos	,175	100,0%
crianças	,158	90,6%

### 3. RESULTADOS PARA AMOSTRA DE COMUNIDADES RURAIS ABASTECIDAS POR SFD E REDE ELÉTRICA CONVENCIONAL, SOMENTE TREINAMENTO (RN\_3)

Tabela A-11 Programação do modelo RN\_3

		Notes	
Output Created			08-JUN-2012 18:31:47
Comments			
Input	Data Active Dataset Filter Weight Split File N of Rows in Working Data File	C:\Users\ROBERTA & EVERTON\Desktop\JADER_11-04-12\DADOS_RUR_DOLAR_8X2_NORM.sav Conjunto_de_dados2 <none> <none> <none>	
Missing Value Handling	Definition of Missing Cases Used	User- and system-missing values are treated as missing. Statistics are based on cases with valid data for all variables used by the procedure.	35
Weight Handling		not applicable MLP CMKWh (MLEVEL=S) DPKWh (MLEVEL=S) WITH RendaFamUS\$ sistema comodos contatourb escolar sazonal adultos crianças /RESCALE COVARIATE=NONE DEPENDENT=NORMALIZED (CORRECTION=0.02) /PARTITION TRAINING=7 TESTING=0 HOLDOUT=0 /ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=6) HIDDENFUNCTION=SIGMOID OUTPUTFUNCTION=SIGMOID	
Syntax		/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005 SIGMAINITIAL=0.00005 INTERVALCENTER=0 INTERVALOFFSET=0.5 MEMSIZE=1000 /PRINT CPS NETWORKINFO SUMMARY SOLUTION IMPORTANCE /PLOT NETWORK PREDICTED /STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=200 ERRORCHANGE=1.0E-4 ERRORRATIO=0.0010 /MISSING USERMISSING=EXCLUDE .	
Resources	Processor Time Elapsed Time		00:00:01,01 00:00:01,05

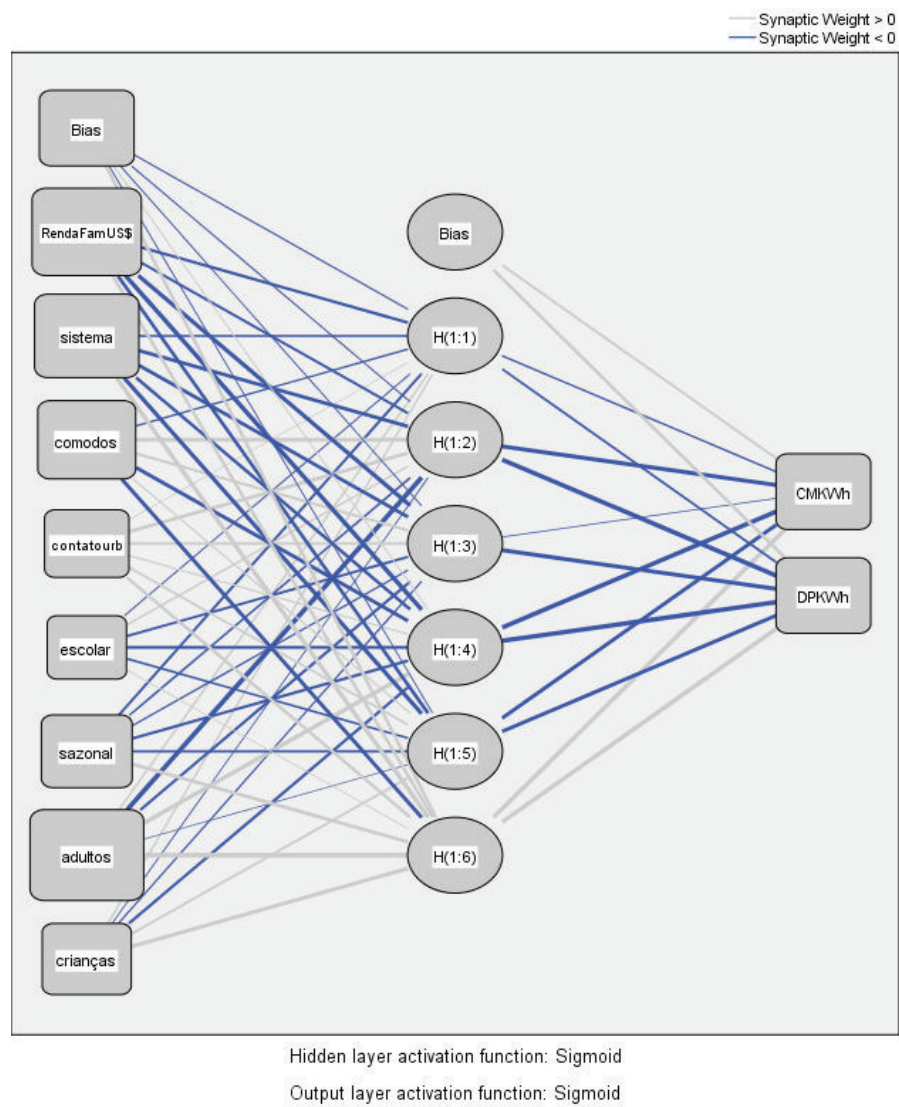
[Conjunto\_de\_dados2] C:\Users\ROBERTA & EVERTON\Desktop\JADER\_11-04-12\DADOS\_RUR\_DOLAR\_8X2\_NORM.sav

Tabela A-12 Particionamento da amostra e configuração da RN\_3

Case Processing Summary				
		N	Percent	
Sample	Training	35	100,0%	
Valid		35	100,0%	
Excluded		0		
Total		35		

Network Information				
Input Layer	Covariates	1	Renda Fam	US\$
		2	sistema	
		3	comodos	
		4	contato urb	
		5	escolar.	
		6	sazonal	
		7	adultos	
		8	crianças	
Hidden Layer(s)	Number of Units			8
	Rescaling Method for Covariates		None	
	Number of Hidden Layers			1
	Number of Units in Hidden Layer 1			6
Output Layer	Activation Function		Sigmoid	
	Dependent Variables	1	CM (KWh)	
		2	DP (KWh)	
	Number of Units			2
	Rescaling Method for Scale Dependents		Normalized	
	Activation Function		Sigmoid	
	Error Function		Sum of Squares	



**Ilustração A-5 Topologia da RN\_3**

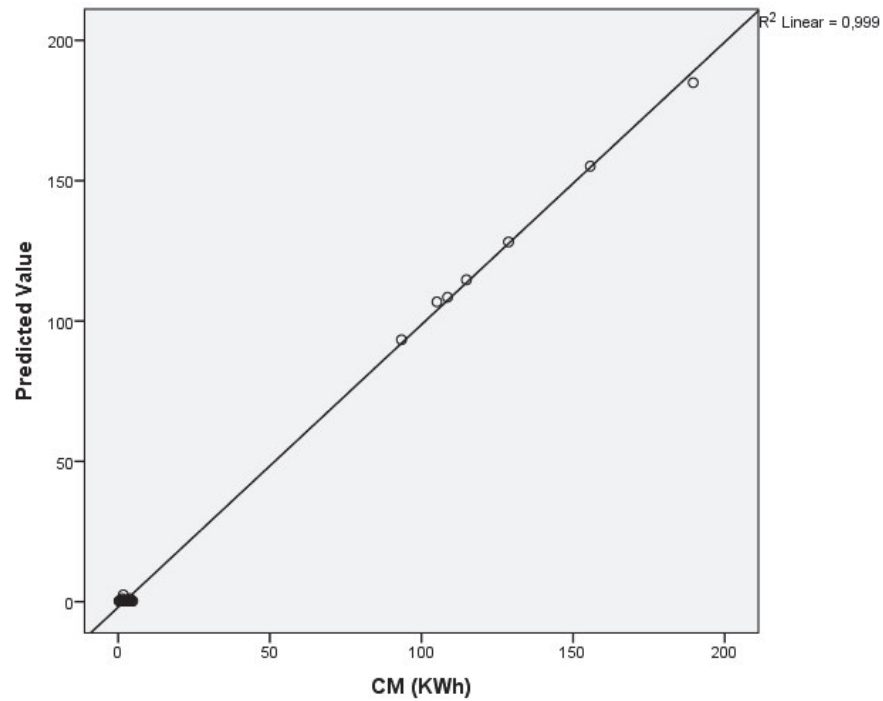
Tabela A-13 Medidas de desempenho da RN\_3

Model Summary			
	Sum of Squares Error		,005
	Average Overall Relative Error		,002
Training	Relative Error for Scale Dependents	CM (KWh)	,002
		DP (KWh)	,003
	Stopping Rule Used	Maximum number of epochs (200) exceeded	
	Training Time		0:00:00,02

Tabela A-14 Pesos sinápticos da RN\_3

Predictor		Parameter Estimates							
		Predicted							
		Hidden Layer 1						Output Layer	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	CMKWh	DPKWh
Input Layer	(Bias)	-,318	-,310	-,169	,247	-,164	1,689		
	RendaFamUS\$	-1,260	-,993	-2,776	-4,154	-1,701	1,556		
	sistema	-,796	-1,834	-1,757	-1,651	-2,602	3,735		
	comodos	-,423	2,515	1,051	-1,842	,349	-1,854		
	contatourb	,069	1,742	1,103	,364	,839	1,503		
	escolar	-,281	,557	-,949	-1,401	-,786	,137		
	sazonal	-,992	-,724	-,408	-1,145	-,905	1,747		
	adultos	,660	-7,554	-1,144	3,689	-,064	7,967		
	crianças	,555	-,137	-,395	-1,248	,866	2,469		
Hidden Layer 1	(Bias)							,947	1,693
	H(1:1)							-,509	-,850
	H(1:2)							-3,210	-5,510
	H(1:3)							-,097	-2,784
	H(1:4)							-5,120	-6,873
	H(1:5)							-1,913	-2,157
H(1:6)							4,514	4,728	





**Ilustração A-6 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN\_3**

**Tabela A-15 Importância das variáveis independentes RN\_3**

Independent Variable Importance		Importance	Normalized Importance
Renda Fam	US\$	,206	88,3%
sistema		,176	75,3%
comodos		,136	58,4%
contato urb		,053	22,9%
escolar.		,021	9,1%
sazonal		,093	39,8%
adultos		,233	100,0%
crianças		,082	35,1%

**4. RESULTADOS PARA AMOSTRA DE COMUNIDADES RURAIS ABASTECIDAS POR SFD E REDE ELÉTRICA CONVENCIONAL, PARTICIONADA EM AMOSTRA DE TREINAMENTO E DE TESTE (RN\_4)**

**Tabela A-16 Programação do modelo RN\_4**

			Notes	
Output Created				12-JUN-2012 19:48:16
Comments				
Input	Data		C:\Users\ROBERTA & EVERTON\Desktop\JADER_11-04-12\DADOS_RUR_DOLAR_8X2.sav	
	Active Dataset		Conjunto_de_dados1	
	Filter		<none>	
	Weight		<none>	
	Split File		<none>	
	N of Rows in Working Data File			50
Missing Value Handling	Definition of Missing		User- and system-missing values are treated as missing.	
	Cases Used		Statistics are based on cases with valid data for all variables used by the procedure.	
Weight Handling			not applicable	
			MLP CMKWh (MLEVEL=S) DPKWh (MLEVEL=S) WITH comodos contatourb escolar sazonal RendaFamUS\$ adultos crianças sistema	
			/RESCALE COVARIATE=ADJNORMALIZED DEPENDENT=NORMALIZED (CORRECTION=0.02)	
			/PARTITION VARIABLE=PARTIÇÃO_0.8	
			/ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=6) HIDDENFUNCTION=SIGMOID	
			OUTPUTFUNCTION=SIGMOID	
Syntax			/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005 SIGMAINITIAL=0.00005	
			INTERVALCENTER=0 INTERVALOFFSET=0.5 MEMSIZE=1000	
			/PRINT CPS NETWORKINFO SUMMARY CLASSIFICATION SOLUTION IMPORTANCE	
			/PLOT NETWORK PREDICTED	
			/SAVE PREDVAL	
			/STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=200	
			ERRORCHANGE=1.0E-4 ERRORRATIO=0.0010	
			/MISSING USERMISSING=EXCLUDE .	
Resources	Processor Time			00:00:01,19
	Elapsed Time			00:00:01,27
Variables Created or Modified	Predicted Value	CMKWh	MLP_PredictedValue_1_A	
		DPKWh	MLP_PredictedValue_2_A	

Conjunto\_de\_dados1] C:\Users\ROBERTA & EVERTON\Desktop\JADER\_11-04-12\DADOS\_RUR\_DOLAR\_8X2.sav

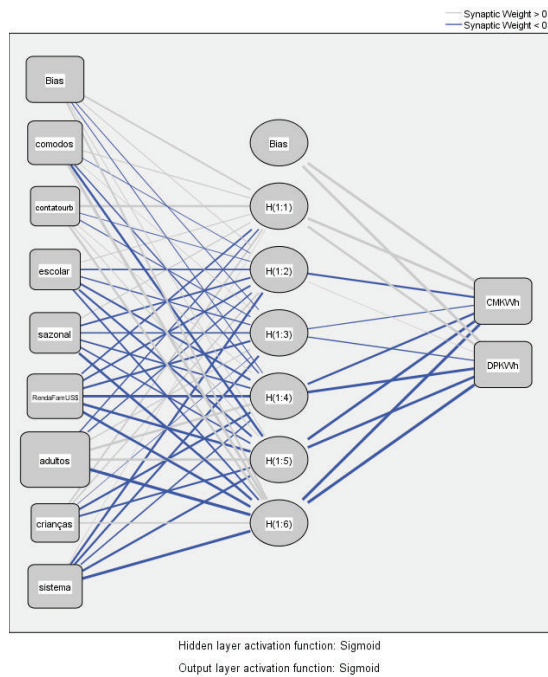
Tabela A-17 Particionamento da amostra e configuração da RN\_4

Case Processing Summary			
		N	Percent
Sample	Training	28	80,0%
	Holdout	7	20,0%
Valid		35	100,0%
Excluded		15	
Total		50	

Network Information			
Input Layer	Covariates	1	comodos
		2	contato urb
		3	escolar.
		4	sazonal
		5	Renda Fam      US\$
		6	adultos
		7	crianças
		8	sistema
Hidden Layer(s)	Number of Units <sup>a</sup>		8
	Rescaling Method for Covariates		Adjusted normalized
	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1 <sup>a</sup>		6
Output Layer	Activation Function		Sigmoid
	Dependent Variables	1	CM (KWh)
		2	DP (KWh)
	Number of Units		2
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares

a. Excluding the bias unit



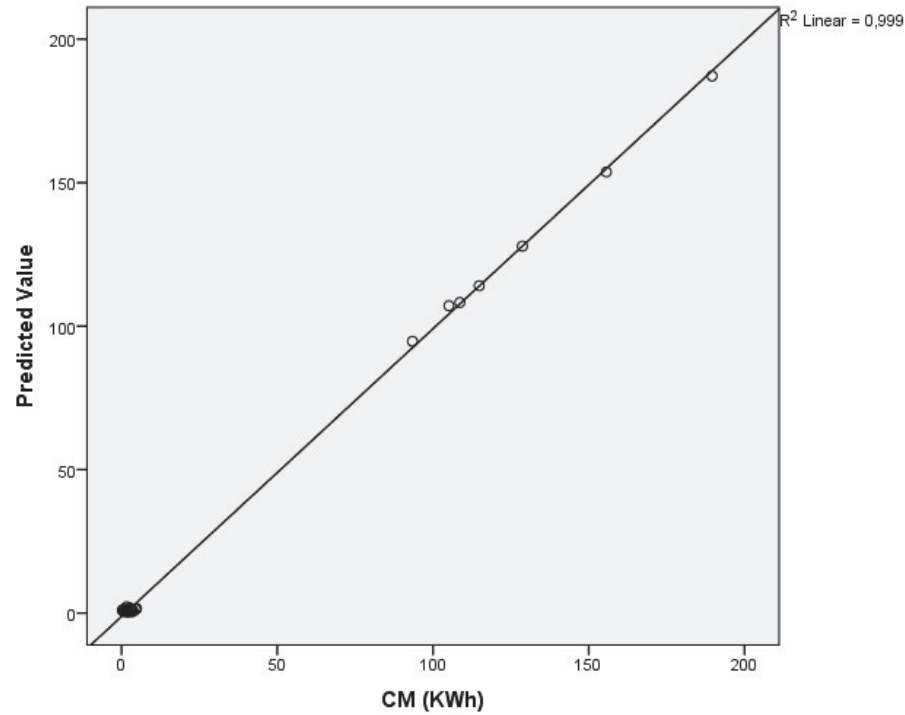
**Ilustração A-7 Topologia da RN\_4**

**Tabela A-18 Medidas de desempenho da RN\_4**

Model Summary			
Training	Sum of Squares Error		,003
	Average Overall Relative Error		,001
	Relative Error for Scale	CM (KWh)	,001
	Dependents	DP (KWh)	,002
	Stopping Rule Used	Maximum number of epochs (200) exceeded	
Holdout	Training Time		0:00:00,02
	Average Overall Relative Error		2,565
	Relative Error for Scale	CM (KWh)	2,512
	Dependents	DP (KWh)	2,631

Tabela A-19 Pesos sinápticos da RN\_4Parameter Estimates

Predictor	Predicted								
	Hidden Layer 1						Output Layer		
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	CMKWh	DPKWh	
Input Layer	(Bias)	1,115	,182	-,174	-,181	,173	1,363		
	comodos	,237	-,164	,120	-,489	-1,916	5,327		
	contatourb	2,001	-,171	-,223	,149	,215	1,847		
	escolar	,212	-,813	-,367	-1,420	-1,521	-1,009		
	sazonal	,272	-1,100	-,702	-,823	-,764	-2,095		
	RendaFamUS\$	-,870	-1,318	-1,592	-2,743	-2,581	-2,712		
	adultos	-,452	-,508	,325	2,184	2,374	-12,121		
	crianças	,918	,753	-,055	-1,670	-1,484	,820		
	sistema	,049	-1,728	-1,660	-,764	-1,792	-3,723		
Hidden Layer 1	(Bias)						4,785	5,205	
	H(1:1)						3,385	2,022	
	H(1:2)						-1,499	,113	
	H(1:3)						-,227	-,292	
	H(1:4)						-1,714	-3,176	
	H(1:5)						-3,329	-3,108	
H(1:6)						-6,558	-7,632		



**Ilustração A-8** Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN\_4

**Tabela A-20** Importância das variáveis independentes RN\_4

Independent Variable Importance		
	Importance	Normalized Importance
comodos	,134	36,6%
contato urb	,049	13,3%
escolar.	,069	18,8%
sazonal	,072	19,6%
Renda Fam US\$	,151	41,3%
adultos	,366	100,0%
crianças	,029	8,0%
sistema	,130	35,4%

## 5. RESULTADOS PARA AMOSTRA COMPOSTA DE FAMÍLIAS DE COMUNIDADES RURAIS ABASTECIDAS POR SFD E REDE ELÉTRICA CONVENCIONAL, E POR FAMILIAS URBANAS, SOMENTE TREINAMENTO (RN\_5)

Tabela A-21 Programação do modelo RN\_5

		Notes	
Output Created			08-JUN-2012 20:18:47
Comments			
Input	Data	C:\Users\ROBERTA & EVERTON\Desktop\JADER_11-04-12\DADOS_TODAS_DOLAR_8X2.sav	
	Active Dataset	Conjunto_de_dados5	
	Filter	<none>	
	Weight	<none>	
	Split File	<none>	
	N of Rows in Working Data File		50
Missing Value Handling	Definition of Missing	User- and system-missing values are treated as missing.	
	Cases Used	Statistics are based on cases with valid data for all variables used by the procedure.	
Weight Handling		not applicable	
		MLP CMKWh (MLEVEL=S) DPKWh (MLEVEL=S) WITH comodos contatourb escolar sazonal RendaFamUS\$ adultos crianças sistema	
		/RESCALE COVARIATE=ADJNORMALIZED DEPENDENT=NORMALIZED (CORRECTION=0.02)	
		/PARTITION TRAINING=7 TESTING=0 HOLDOUT=0	
		/ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=5) HIDDENFUNCTION=SIGMOID	
		OUTPUTFUNCTION=SIGMOID	
Syntax		/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005 SIGMAINITIAL=0.00005	
		INTERVALCENTER=0 INTERVALOFFSET=0.5 MEMSIZE=1000	
		/PRINT CPS NETWORKINFO SUMMARY SOLUTION IMPORTANCE	
		/PLOT NETWORK PREDICTED	
		/STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=200	
		ERRORCHANGE=1.0E-4 ERRORRATIO=0.0010	
		/MISSING USERMISSING=EXCLUDE .	
Resources	Processor Time		00:00:00,98
	Elapsed Time		00:00:01,01

[Conjunto\_de\_dados5] C:\Users\ROBERTA & EVERTON\Desktop\JADER\_11-04-12\DADOS\_TODAS\_DOLAR\_8X2.sav

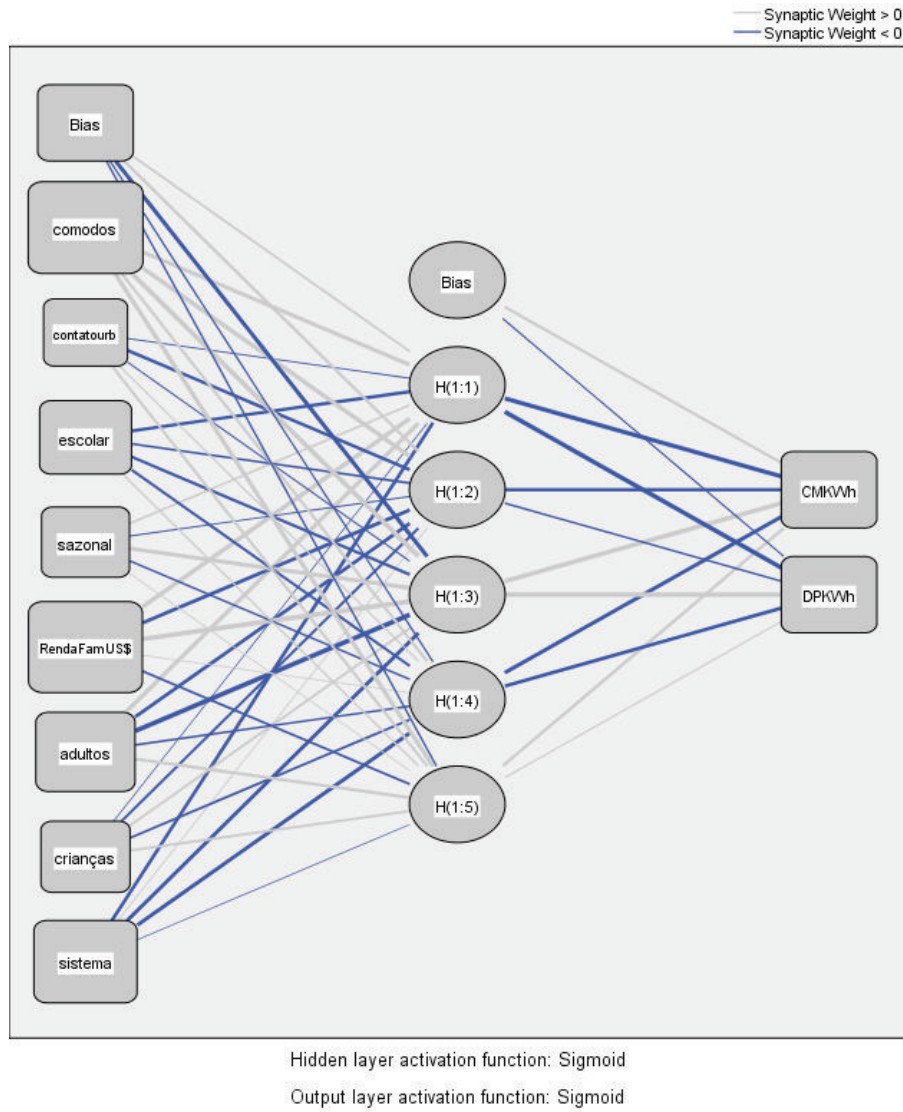
Tabela A-22 Particionamento da amostra e configuração da RN\_5

Case Processing Summary			
		N	Percent
Sample	Training	50	100,0%
Valid		50	100,0%
Excluded		0	
Total		50	

Network Information			
		1	comodos
		2	contato urb
		3	escolar.
	Covariates	4	sazonal
Input Layer		5	Renda Fam US\$
		6	adultos
		7	crianças
		8	sistema
	Number of Units		8
	Rescaling Method for Covariates		Adjusted normalized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1		5
	Activation Function		Sigmoid
	Dependent Variables	1	CM (KWh)
		2	DP (KWh)
Output Layer	Number of Units		2
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares





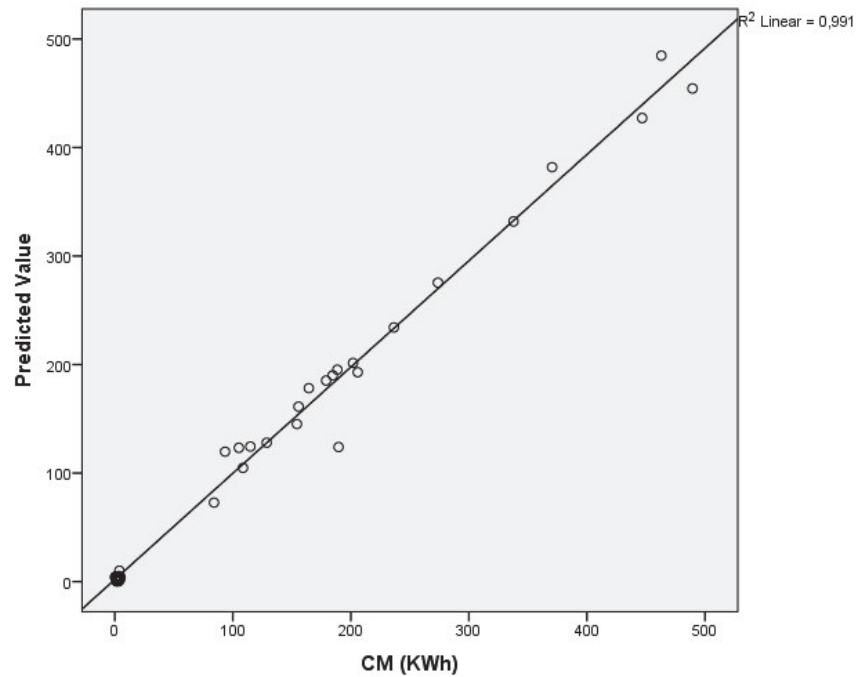
**Ilustração A-9 Topologia da RN\_5**

Tabela A-23 Medidas de desempenho da RN\_5

Model Summary		
	Sum of Squares Error	,083
	Average Overall Relative Error	,029
Training	Relative Error for Scale Dependents	,009
	CM (KWh)	,070
	DP (KWh)	,070
	Stopping Rule Used	Maximum number of epochs (200) exceeded
	Training Time	0:00:00,02

Tabela A-24 Pesos sinápticos da RN\_5

Predictor		Parameter Estimates						
		Hidden Layer 1					Output Layer	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	CMKWh	DPKWh
Input Layer	(Bias)	1,190	2,351	-4,584	-,633	-,884		
	comodos	4,200	5,019	11,046	2,028	4,743		
	contatourb	-,265	-3,010	-,308	,424	2,014		
	escolar	-3,187	-1,184	-2,328	-1,695	,872		
	sazonal	1,117	-,267	4,595	-1,022	,102		
	RendaFamUS\$	5,670	-3,631	11,301	,088	-1,201		
	adultos	4,644	-3,361	-8,674	-1,034	2,874		
	crianças	-,078	-1,339	1,243	-1,239	1,253		
	sistema	-3,359	,125	-3,654	-3,725	-,030		
		(Bias)					2,175	-,379
Hidden Layer 1	H(1:1)						-5,660	-6,357
	H(1:2)						-3,233	-1,004
	H(1:3)						7,078	7,723
	H(1:4)						-3,773	-3,521
	H(1:5)						2,605	,368



**Ilustração A-10** Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN\_5

**Tabela A-25** Importância das variáveis independentes RN\_5

Independent Variable Importance		
	Importance	Normalized Importance
comodos	,223	100,0%
contato urb	,043	19,5%
escolar.	,089	40,0%
sazonal	,066	29,7%
Renda Fam US\$	,218	97,7%
adultos	,132	59,1%
crianças	,073	32,9%
sistema	,155	69,8%

**6. RESULTADOS PARA AMOSTRA COMPOSTA DE FAMÍLIAS DE COMUNIDADES RURAIS ABASTECIDAS POR SFD E REDE ELÉTRICA CONVENCIONAL, E POR FAMILIAS URBANAS, PARTICIONADA EM AMOSTRA DE TREINAMENTO E TESTE (RN\_6)**

**Tabela A-26 Programação do modelo RN\_6**

		Notes	
Output Created			09-JUN-2012 20:57:23
Comments			
Input	Data	C:\Users\ROBERTA & EVERTON\Desktop\JADER_11-04-12\DADOS_TODAS_DOLAR_8X2.sav	
	Active Dataset	Conjunto_de_dados2	
	Filter	<none>	
	Weight	<none>	
	Split File	<none>	
	N of Rows in Working Data File		50
Missing Value Handling	Definition of Missing	User- and system-missing values are treated as missing.	
	Cases Used	Statistics are based on cases with valid data for all variables used by the procedure.	
Weight Handling		not applicable	
		MLP CMKWh (MLEVEL=S) DPKWh (MLEVEL=S) WITH comodos contatourb escolar sazonal RendaFamUS\$ adultos crianças sistema	
		/RESCALE COVARIATE=ADJNORMALIZED DEPENDENT=NORMALIZED (CORRECTION=0.02)	
		/PARTITION VARIABLE=PARTIÇÃO_0.7	
		/ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=5) HIDDENFUNCTION=SIGMOID	
		OUTPUTFUNCTION=SIGMOID	
		/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005	
Syntax		SIGMAINITIAL=0.00005 INTERVALCENTER=0 INTERVALOFFSET=0.5 MEMSIZE=1000	
		/PRINT CPS NETWORKINFO SUMMARY CLASSIFICATION SOLUTION IMPORTANCE	
		/PLOT NETWORK PREDICTED	
		/SAVE PREDVAL	
		/STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=AUTO	
		ERRORCHANGE=1.0E-4 ERRORRATIO=0.0010	
		/MISSING USERMISSING=EXCLUDE .	
Resources	Processor Time		00:00:01,06
	Elapsed Time		00:00:01,06
Variables Created or Modified	Predicted Value	CMKWh	MLP_PredictedValue_1_K
		DPKWh	MLP_PredictedValue_2_K

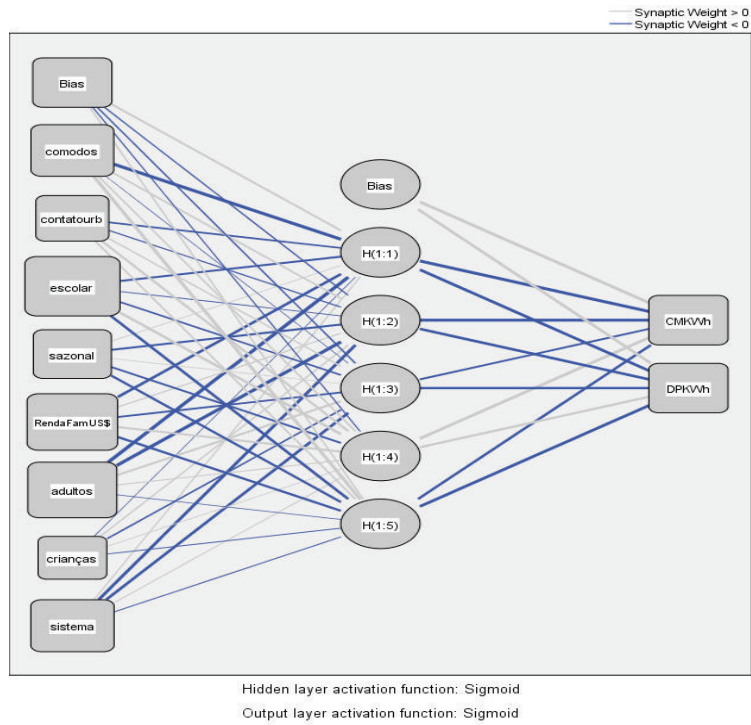
[Conjunto\_de\_dados2] C:\Users\ROBERTA & EVERTON\Desktop\JADER\_11-04-12\DADOS\_TODAS\_DOLAR\_8X2.sav

Tabela A-27 Particionamento da amostra e configuração da RN\_6

Case Processing Summary			
		N	Percent
Sample	Training	34	68,0%
	Holdout	16	32,0%
Valid		50	100,0%
Excluded		0	
Total		50	

Network Information			
Input Layer	Covariates	1	comodos
		2	contato urb
		3	escolar.
		4	sazonal
		5	Renda Fam      US\$
		6	adultos
		7	crianças
		8	sistema
Hidden Layer(s)	Number of Units		8
	Rescaling Method for Covariates		Adjusted normalized
	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1		5
Output Layer	Activation Function		Sigmoid
	Dependent Variables	1	CM (KWh)
		2	DP (KWh)
	Number of Units		2
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
Error Function		Sum of Squares	



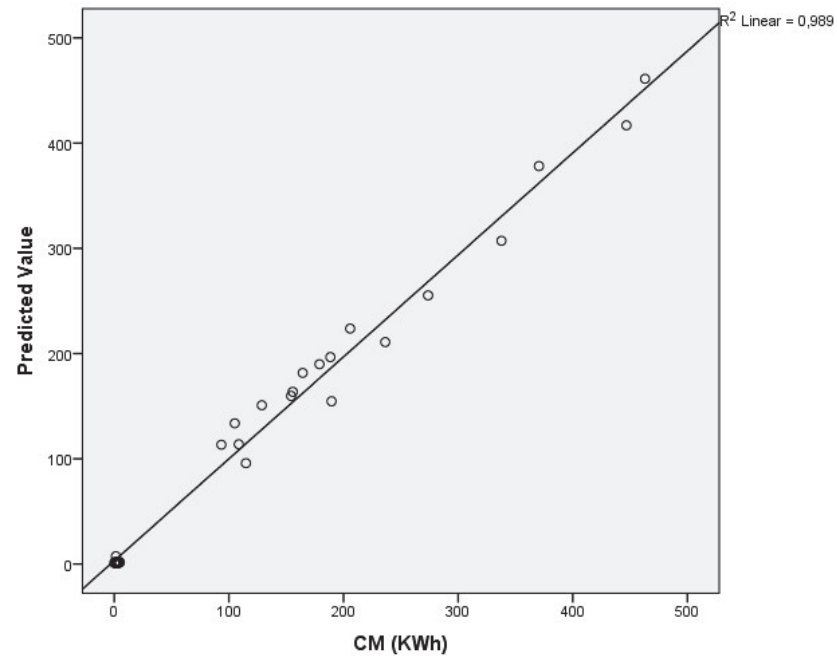
**Ilustração A-11 Topologia da RN\_6**

**Tabela A-28 Medidas de desempenho da RN\_6**

Model Summary			
	Sum of Squares Error		,057
	Average Overall Relative Error		,026
Training	Relative Error for Scale Dependents	CM (KWh)	,012
		DP (KWh)	,053
	Stopping Rule Used	Maximum number of epochs (100) exceeded	
	Training Time		0:00:00,01
Holdout	Average Overall Relative Error		,676
		Relative Error for Scale Dependents	
		CM (KWh)	,416
	DP (KWh)	,839	

Tabela A-29 Pesos sinápticos da RN\_6

Predictor		Parameter Estimates						
		Hidden Layer 1					Output Layer	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	CMKWh	DPKWh
Input Layer	(Bias)	1,364	-,555	-,647	-,439	,693		
	comodos	-3,937	,959	-,057	,754	3,277		
	contatourb	-,839	-,541	,666	1,695	,813		
	escolar	-1,430	-,341	-,775	2,147	-2,078		
	sazonal	,209	-1,456	,120	-1,209	-1,703		
	RendaFamUS\$	-2,164	,594	-1,104	1,565	-2,044		
	adultos	-4,461	-3,862	1,693	,455	-,103		
	crianças	-,073	,657	-,720	,207	-,384		
	sistema	,590	-2,473	-2,197	,419	-,403		
	(Bias)						3,297	2,647
Hidden Layer 1	H(1:1)						-3,365	-2,680
	H(1:2)						-3,413	-2,296
	H(1:3)						-1,275	-2,046
	H(1:4)						3,084	2,049
	H(1:5)						-2,065	-2,944



**Ilustração A-12** Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN\_6

**Tabela A-30** Importância das variáveis independentes RN\_6

Independent Variable Importance

	Importance	Normalized Importance
comodos	,127	61,1%
contato urb	,062	30,0%
escolar.	,207	100,0%
sazonal	,096	46,2%
Renda Fam US\$	,177	85,3%
adultos	,167	80,7%
crianças	,034	16,6%
sistema	,130	62,6%



## 7. RESULTADOS PARA AMOSTRA EM QUE OS DADOS DE CAMPO (RURAIS E URBANAS) FORAM ACRESCIDOS DE RÚIDO RANDOMICO, SOMENTE TREINAMENTO (RN\_7)

Tabela A-31 Programação do modelo RN\_7

		Notes	
Output Created			08-MAR-2013 03:18:32
Comments			
Input	Active Dataset	Conjunto_de_dados1	
	Filter	<none>	
	Weight	<none>	
	Split File	<none>	
	N of Rows in Working Data File		50
Missing Value Handling	Definition of Missing	User- and system-missing values are treated as missing.	
	Cases Used	Statistics are based on cases with valid data for all variables used by the procedure.	
Weight Handling		not applicable	
		MLP CM (MLEVEL=5) WITH comod0s contatourb escolar sazonal RendaFamUS\$ adultos crianças sistema	
		/RESCALE COVARIATE=NORMALIZED DEPENDENT=NORMALIZED (CORRECTION=0.02)	
		/PARTITION TRAINING=7 TESTING=0 HOLDOUT=0	
		/ARCHITECTURE AUTOMATIC=NO HIDDENLAYERS=1 (NUMUNITS=5) HIDDENFUNCTION=SIGMOID OUTPUTFUNCTION=SIGMOID	
		/CRITERIA TRAINING=BATCH OPTIMIZATION=SCALEDCONJUGATE LAMBDAINITIAL=0.0000005 SIGMAINITIAL=0.00005	
Syntax		INTERVALCENTER=0 INTERVALOFFSET=0.5 MEMSIZE=1000	
		/PRINT CPS NETWORKINFO SUMMARY SOLUTION IMPORTANCE	
		/PLOT NETWORK PREDICTED RESIDUAL	
		/SAVE PREDVAL(MLP_PredictedValue_ruido )	
		/STOPPINGRULES ERRORSTEPS= 1 (DATA=AUTO) TRAININGTIMER=ON (MAXTIME=15) MAXEPOCHS=AUTO ERRORCHANGE=1.0E-4	
		ERRORRATIO=0.0010	
		/MISSING USERMISSING=EXCLUDE .	
Resources	Processor Time		00:00:03,57
	Elapsed Time		00:00:03,49
Variables Created or Modified	Predicted Value	MLP_PredictedValue_ruido	

Tabela A-32 Particionamento da amostra e configuração da RN\_7



Ilustração A-13 Topologia da RN\_7

Tabela A-33 Medidas de desempenho da RN\_7

Model Summary		
Training	Sum of Squares Error	,006
	Relative Error	,003
	Stopping Rule Used	Maximum number of epochs (100) exceeded
	Training Time	0:00:00,01

Tabela A-34 Pesos sinápticos da RN\_7

Parameter Estimates							
Predictor		Predicted					
		Hidden Layer 1					Output Layer
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	CM
Input Layer	(Bias)	,183	,602	,070	,120	1,607	
	comodos	-,223	-,675	-1,111	,076	-1,062	
	contatourb	,319	-,096	-,120	-,387	1,448	
	escolar	-1,114	-,878	-,578	1,284	-1,352	
	sazonal	,195	,565	-,475	,406	-,380	
	RendaFamUS\$	-1,505	-1,649	-1,302	1,199	-3,211	
	adultos	-,013	-,359	-,014	,166	,353	
	crianças	,038	,314	,867	-,208	-,240	
	sistema	-,210	-1,092	-1,348	-,176	1,783	
	(Bias)						2,746
Hidden Layer 1	H(1:1)						-1,863
	H(1:2)						-2,379
	H(1:3)						-1,303
	H(1:4)						2,871
	H(1:5)						-4,289

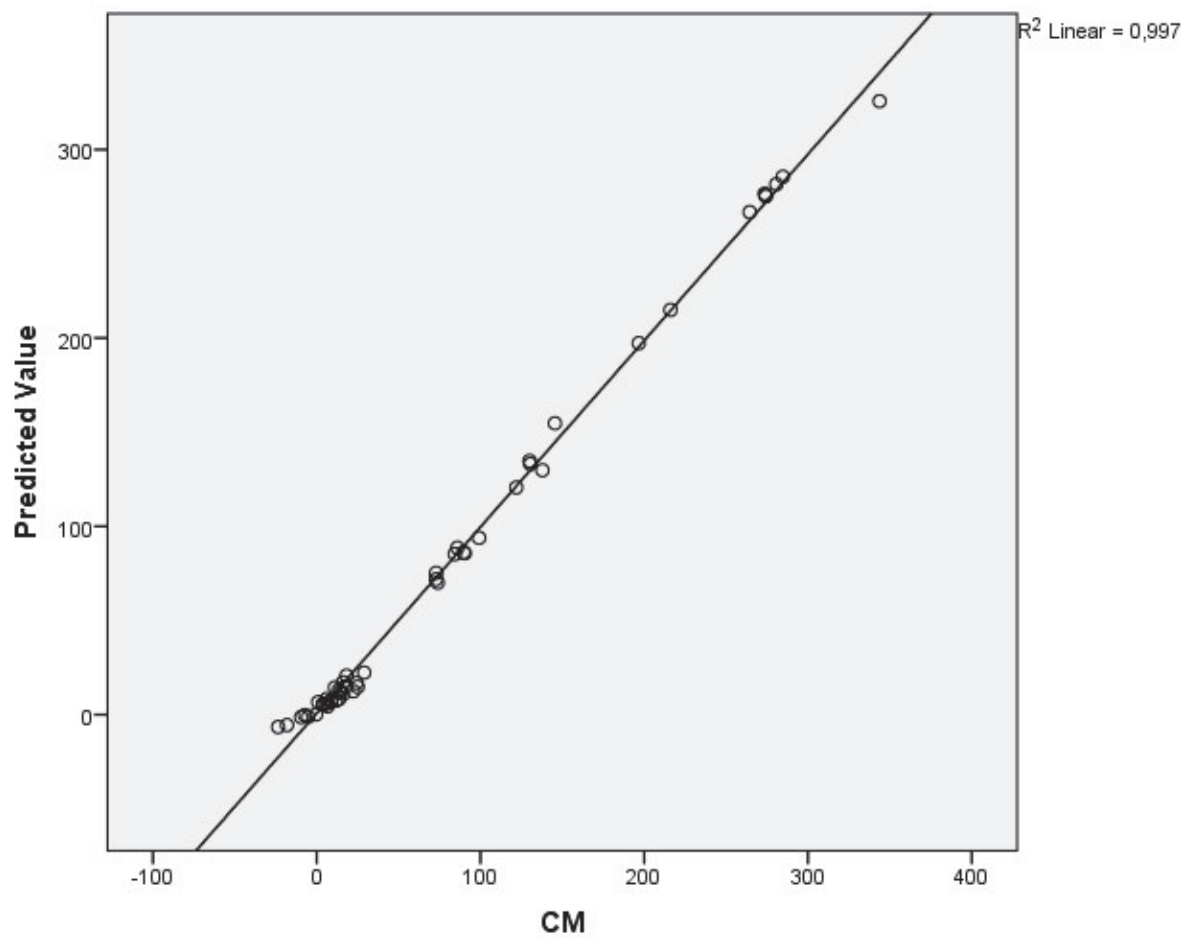


Ilustração A-14 Ajuste linear de consumo de energia elétrica familiar mensal real e previstos pela RN\_7

## 8. SOMA DOS ERROS QUADRÁTICOS

$$\text{SSE} = \sum \left( y - \hat{y} \right)^2$$

## 9. ALGORITMO DO GRADIENTE CONJUGADO ESCALONADO (EXECUTADO NO SPSS-IBM)

1. Escolha um vetor de parâmetros inicial  $\theta_0$  e escalares  $0 < \sigma \ll 1$  e  $\lambda_0 = 0$ ;

2. Faça:  $p_0 = r_0 = -\nabla J(\theta_0); j = 0$ ; sucesso = 1

3. Se sucesso = 1, calcule a informação de 2ª ordem:  $\sigma_j = \frac{\sigma}{\|d_j\|}$

$$s_j = \frac{\nabla J(\theta_j + \sigma_j d_j) - \nabla J(\theta_j)}{\sigma_j} + \lambda_j d_j; \quad \delta_j = p_j^T s_j$$

4. Se  $\delta_j \leq 0$ , então faça a matriz hessiana definida positiva:

$$\lambda_j^N = 2 \left[ \lambda_j - \frac{\delta_j}{\|d_j\|^2} \right]; \quad \delta_j = \delta_j + (\lambda_j^N - \lambda_j) \|d_j\|^2; \quad \lambda_j = \lambda_j^N$$

5. Calcule a taxa de ajuste:

$$\mu_j = p_j^T r_j; \quad \alpha_j = \mu_j / \delta_j$$

6. Calcule o parâmetro de comparação:

$$\Delta_j = \frac{2\delta_j [J(\theta_j) - J(\theta_j + \alpha_j d_j)]}{\mu_j^2}$$

7. Se  $\Delta_j \geq 0$  (o erro pode ser reduzido), então atualize o vetor de pesos:

$$\theta_{j+1} = \theta_j + \alpha_j d_j; \quad r_{j+1} = -\nabla J(\theta_{j+1});$$

8. Se  $(j \bmod P) = 0$ , então reinicialize o algoritmo:

$$p_{j+1} = r_{j+1}.$$

Senão, defina uma nova direção conjugada:

$$\beta_j = \frac{\|r_{j+1}\|^2 - r_{j+1}^T r_j}{\mu_j}; \quad p_{j+1} = r_{j+1} + \beta_j p_j$$

9. Se  $\Delta_j \geq 0.75$ , então faça  $\lambda_j = 0.5\lambda_j$ .

10. Se  $\Delta_j \leq 0.25$ , então faça  $\lambda_j = 4\lambda_j$ .

Senão, uma redução no erro não é possível: sucesso = 0;

11. Se  $r_{j+1} > \varepsilon$ , onde  $\varepsilon \rightarrow 0$ , então faça:  $j = j+1$ ; retorne ao passo 3.

12. Senão, o procedimento de ajuste chegou ao fim e  $\theta_{j+1}$  é o ponto de mínimo.

## APÊNDICE B – REGRESSÃO LINEAR MÚLTIPLA E ESTATÍSTICAS

Fonte: MONTGOMERY (2003)

O modelo de regressão múltipla pode ser expresso como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Os  $\beta$ 's podem ser estimados pelo método dos mínimos quadrados.

Para estimar os  $\beta_k$  pode-se utilizar uma amostra de  $n$  observações de  $y$  associados com os  $x_k$ . O modelo para a  $i$ -ésima observação é

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

As suposições para  $\varepsilon_i$  (ou  $y_i$ ) são análogas àquelas da regressão linear simples a saber:

1.  $E(\varepsilon_i) = 0$  para todo  $i = 1, 2, \dots, n$ , ou, equivalentemente,  $E(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$ .
2.  $var(\varepsilon_i) = \sigma^2$  para todo  $i = 1, 2, \dots, n$ , ou, equivalentemente,  $var(y_i) = \sigma^2$ .
3.  $cov(\varepsilon_i, \varepsilon_j) = 0$  para todo  $i \neq j$ , ou, equivalentemente,  $cov(y_i, y_j) = 0$ .

Escrevendo para cada uma das  $n$  observações temos

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k} + \varepsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2k} + \varepsilon_2$$

⋮

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_k x_{nk} + \varepsilon_n$$

Essas  $n$  equações podem ser escritas na forma matricial como

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

### Regressão Linear Simples

Equação da reta

$$\hat{y} = a + bx$$

### Método dos Mínimos Quadrados

$$b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \quad a = \bar{y} - b \cdot \bar{x}$$

### Coefficiente de Determinação ( $R^2$ )

$$r^2 = \frac{b \cdot \left( \sum xy - \frac{\sum x \sum y}{n} \right)}{\sum y^2 - \frac{(\sum y)^2}{n}}$$

Onde

b : coeficiente angular da reta de regressão linear aplicada a amostra

### Teste de Hipóteses e estatística t

Hipótese nula:  $H_0 : r = 0$

A estatística do teste é dada por:

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

e sob  $H_0$ ,  $t$  tem distribuição t-Student com  $(n - 2)$  graus de liberdade.

### Coefficiente de correlação Linear de Pearson ( $R$ )

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left[ \sum x^2 - \frac{(\sum x)^2}{n} \right] \times \left[ \sum y^2 - \frac{(\sum y)^2}{n} \right]}}$$



