



XXXVII IBERIAN LATIN AMERICAN CONGRESS
ON COMPUTATIONAL METHODS IN ENGINEERING
BRASÍLIA - DF - BRAZIL

REDES NEURAIS NA GERAÇÃO DE MODELOS DE PREVISÃO DE PODER CALORÍFICO DE RESÍDUOS SÓLIDOS

Ricardo Drudi

Kelly Cristina Rosa Drudi

Graziella Colato Antonio

Juliana Tófano de Campos Leite Toneli

ricardo.drudi.br@ieee.org

kellydrudi@gmail.com

graziella.colato@ufabc.edu.br

juliana.toneli@ufabc.edu.br

UFABC – Fundação Universidade Federal do ABC

Avenida dos Estados, 5001, 09210-580, Santo André, SP, Brasil

Abstract. *As crescentes preocupações socioambientais têm incentivado discussões acerca de alternativas que aliem uma diminuição da participação de fontes fósseis na matriz energética e a redução de impactos ambientais negativos decorrentes da atividade humana. O aproveitamento energético de resíduos sólidos urbanos contribui tanto para o incremento da oferta de energia quanto para a redução do volume dos resíduos a serem depositados no solo. Entretanto, para uma implementação adequada de uma planta de recuperação energética é imprescindível que se possa determinar o potencial energético dos resíduos. Este trabalho propôs-se a explorar várias configurações de uma técnica computacional denominada redes neurais artificiais para determinação da melhor abordagem no desenvolvimento de modelos de previsão de potencial energético de resíduos sólidos. Dentre as 432 configurações testadas, 80 (18,5%) obtiveram um índice de previsão considerado excelente, e a rede de melhor desempenho atingiu um índice de erro médio percentual absoluto de apenas 2,9%, superando os modelos matemáticos utilizados para comparação. Os resultados mostram que o algoritmo de treinamento e as funções de ativação são os parâmetros mais importantes de uma rede neural na geração de modelo de previsão de potencial energético de resíduos sólidos urbanos.*

Keywords: *Resíduos sólidos urbanos, Poder calorífico superior, Redes neurais artificiais, Modelos de previsão, Aproveitamento energético de resíduos*

1 INTRODUÇÃO

A crescente geração de resíduos nas grandes metrópoles tem se constituído em um problema para as administrações públicas. De um lado, tem-se o aumento populacional associado à concentração dessa população em áreas urbanas, e, de outro, o crescente consumo de produtos industrializados que geram, como subproduto de sua cadeia de produção-consumo, uma quantidade de resíduos de difícil manejo.

Esse problema tem sido ainda mais significativo em países em desenvolvimento, como o Brasil, pois nesses países a necessidade de crescimento econômico como meio de melhoria da qualidade de vida de sua população se contrapõe às preocupações ambientais, que limitam as formas como esse crescimento deve ser atingido. O adequado equilíbrio entre crescimento econômico, atendimento das necessidades sociais e preservação do meio ambiente é uma tarefa bastante complexa e que dificilmente terá resposta única.

Nesse sentido, dar um tratamento adequado à questão dos resíduos sólidos urbanos (RSU) é parte importante nesse processo, e ações que aliem a diminuição de seu impacto ambiental negativo a um aumento da oferta de energia utilizando-se dos resíduos vêm ao encontro dessa necessidade.

Segundo a ABRELPE – Associação Brasileira das Empresas de Limpeza Pública e Resíduos Especiais, a geração de RSU no Brasil tem aumentado ano a ano (Fig. 1). Em 2014, o país gerou um total de mais de 78 milhões de toneladas de RSU, com um índice *per capita* de 387,63 kg/hab/ano, ou seja, mais de 1 kg de resíduo gerado por pessoa por dia.



Figura 1. Crescimento da geração de RSU no Brasil total e *per capita* (ABRELPE, 2015).

O gerenciamento dos resíduos sólidos urbanos tem se mostrado um desafio considerável para o país. Mesmo após a promulgação da PNRS - Política Nacional de Resíduos Sólidos (lei nº 12.305/2010), o Brasil ainda se encontra distante de dar um tratamento satisfatório a essa questão. Ainda segundo a ABRELPE (2015), 41,6% de todo o resíduo coletado no Brasil tem um destino inadequado, sendo que, em 2014, 58,4% foram destinados a Aterros Sanitários (locais com adequado tratamento dos resíduos, evitando contaminação de ar, solo e meios aquáticos), 17,4% foram depositados em Lixões (vazadouros a céu aberto, sem nenhum tipo de tratamento, com contaminação do meio ambiente e proliferação de animais e vetores de doenças), e 24,2% foram destinados a Aterros Controlados (geralmente um Lixão “remediado”,

com algum tipo de controle, administrativo ou ambiental, mas que ainda permite a contaminação do meio ambiente) (Fig. 2).

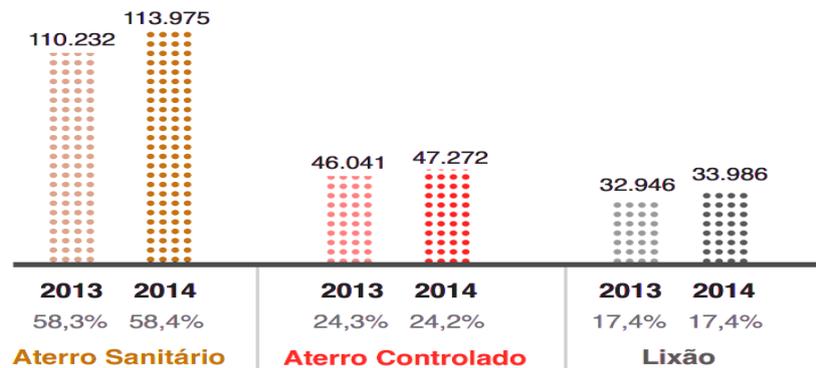


Figura 2. Destinação final do RSU no Brasil (t/dia) (ABRELPE, 2015).

Conforme preconizado pela PNRS em suas diretrizes (Capítulo I, art. 9º e Capítulo II, art. 7º) (Brasil, 2010), a recuperação energética dos resíduos sólidos é uma alternativa à sua deposição no solo, quer seja em aterros sanitários ou em locais inadequados. Segundo relatório elaborado pela Federação Estadual do Meio Ambiente (FEAM) sobre o aproveitamento energético do RSU no Estado de Minas Gerais (FEAM, 2012), “o aproveitamento energético de RSU, desde que utilize rotas tecnológicas apropriadas e devidamente analisadas quanto aos riscos de implementação, é uma alternativa ambientalmente correta de tratamento desses resíduos e uma oportunidade de negócios”.

O aproveitamento energético do RSU pode ser feito através de rotas térmicas ou de rotas biológicas. A Figura 3, adaptada de DDMA (2010), resume os principais processos para o aproveitamento do RSU, bem como seus produtos finais.

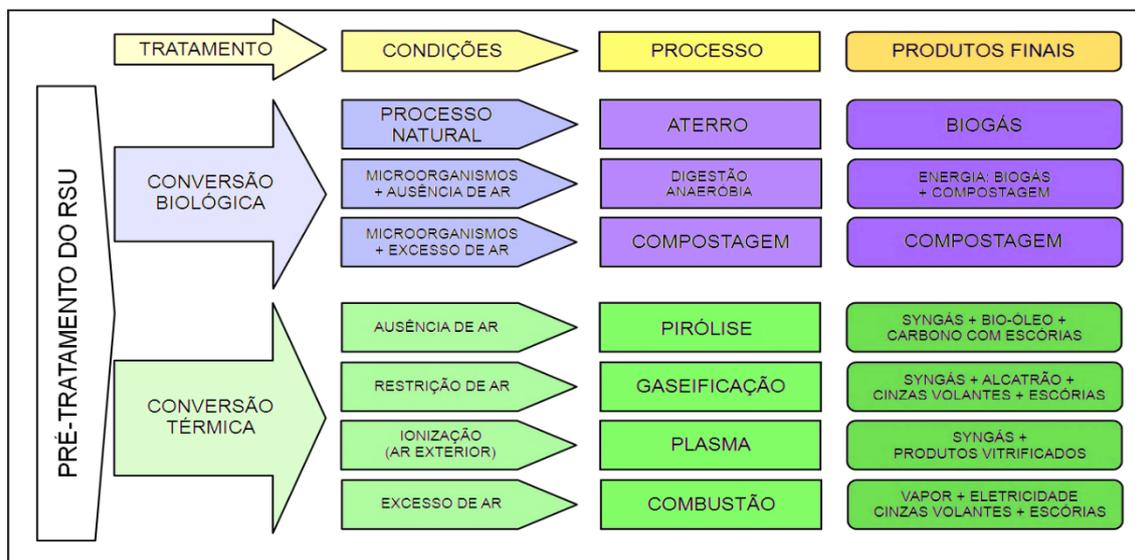


Figura 3. Rotas de aproveitamento do RSU. Adaptado de DDMA (2010).

As rotas de conversão biológica do RSU podem ser definidas como de processo natural, que ocorrem espontaneamente em locais de depósito de resíduos, ou controladas, onde geralmente são utilizados biodigestores próprios para essa tarefa. Já as rotas de conversão térmica dividem-se em: combustão, pirólise, gaseificação e plasma.

A tecnologia de plasma faz uso de um arco elétrico, ionizando gás e gerando altíssimas temperaturas (tipicamente na ordem de 15.000°C). Já a tecnologias de pirólise, gaseificação e combustão têm diferentes produtos finais, priorizando a geração de bio-óleo, gás de síntese e energia, respectivamente.

1.1 Incineração como processo de recuperação energética do RSU

Segundo o relatório da FEAM, a incineração dos resíduos sólidos urbanos traz como benefícios (FEAM, 2012):

- Redução do volume e massa dos resíduos;
- Destruição completa da maioria dos resíduos orgânicos perigosos;
- Recuperação energética (eletricidade e/ou vapor).

Sendo assim, a incineração mostra-se como uma alternativa interessante para a disposição final ambientalmente adequada dos rejeitos, pois além de ser uma fonte alternativa para geração de energia, também reduz o volume dos rejeitos a serem aterrados e mitiga sua toxicidade.

Entretanto, para avaliar a viabilidade da recuperação energética do RSU, bem como permitir o adequado dimensionamento de uma planta de incineração com esse fim, é imprescindível prever seu poder calorífico (Ryu e Shin, 2013; Zhou *et al.*, 2014; Shi, 2016), que é denominado superior (PCS) ou inferior (PCI), conforme se considere a fase da água nos produtos (PCS se a água estiver na fase líquida, e PCI se a água estiver na fase vapor). Nesse sentido, muitos trabalhos têm sido desenvolvidos com o objetivo de encontrar modelos de previsão de PCS que tenham uma melhor adequação local, principalmente em países onde o processo de urbanização da população faz com que o tratamento de resíduos se torne não somente uma questão energética, mas também ambiental e social.

A Figura 4 apresenta um esquemático do funcionamento de uma planta de conversão térmica de RSU. Nela destacam-se: 1) a chegada e deposição dos resíduos; 2) a vala onde os resíduos são mantidos até sua utilização; 3) a grelha onde ocorre a incineração; 4) a caldeira onde é gerado o vapor; 5) a turbina que converte vapor em eletricidade; 6) os exaustores dos gases de combustão e 7) a retirada de cinzas inertizadas para deposição em aterro.

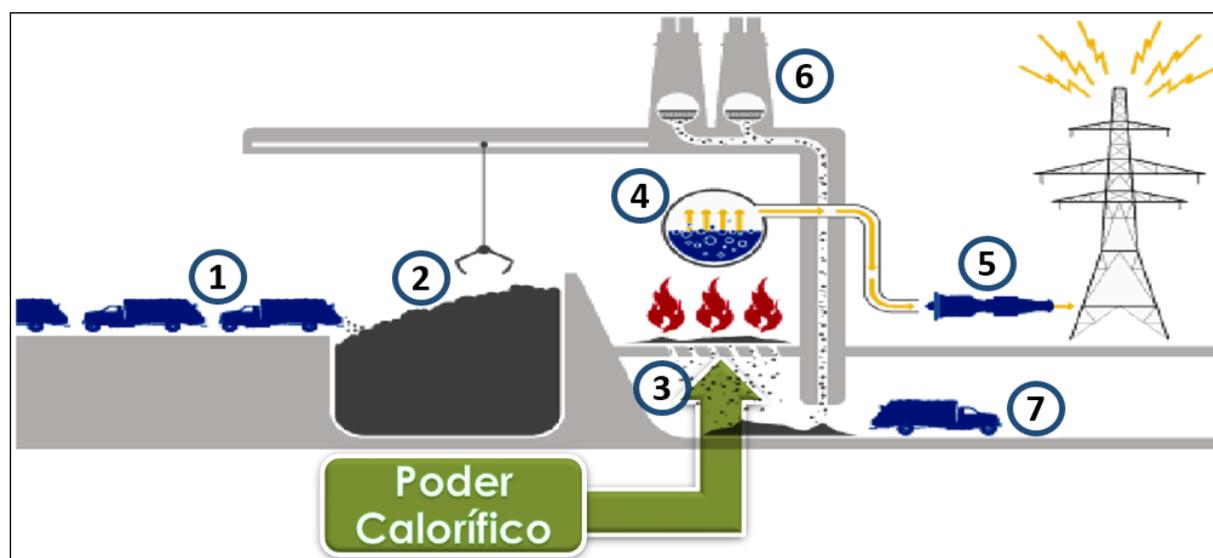


Figura 4. Esquemático de uma planta de conversão térmica de RSU.

A necessidade de se determinar o poder calorífico dos resíduos está no adequado controle da temperatura da câmara de combustão, que não pode ser muito elevada devido a possíveis danos às paredes equipamentos da câmara e não pode estar abaixo de um certo nível para proporcionar uma melhor eficiência da turbina a vapor (Ryu e Shin, 2013).

O desenvolvimento de modelos para a previsão do poder calorífico do RSU tem sido objeto de pesquisa de diversos autores. Chang *et al.* (2007) fizeram uma compilação de 9 equações matemáticas para previsão do PCI, englobando modelos baseados em análise imediata, análise gravimétrica e análise elementar. Kathiravale *et al.* (2003) por sua vez compilaram 18 modelos de previsão de PCS, sendo que 12 somente a partir de análise elementar. Nessa mesma linha, Channiwala e Parikh (2002) listaram em seu trabalho 22 modelos de previsão de PCS baseados em análise elementar, para diversas biomassas. Essa grande quantidade de modelos propostos mostra a relevância de pesquisas nessa área, devido à importância de se obter métodos cada vez mais acurados para a previsão de poder calorífico de biocombustíveis.

2 MATERIAIS E MÉTODOS

A grande maioria dos modelos de previsão de PCS de RSU encontrados na literatura foram desenvolvidos a partir de uma técnica matemática denominada regressão linear múltipla, técnica essa que permite estabelecer as relações lineares entre as variáveis de entrada e a variável de saída. Neste estudo optou-se por uma abordagem computacional para o desenvolvimento do modelo de previsão, através da utilização de redes neurais artificiais (RNA), técnica de inteligência artificial largamente empregada para a aproximação de funções lineares e não lineares (Haykin, 2001).

Redes neurais artificiais já foram utilizadas para a criação de modelos de previsão de PCS de RSU, como pode ser visto nos trabalhos de Dong, Jin e Li (2003) e Ogwueleka e Ogwueleka (2010). Entretanto tais trabalhos pouco exploraram o potencial das RNA como geradoras de modelos de previsão, e limitaram-se a criar um único modelo utilizando uma configuração pré-determinada de rede neural.

A fim de suprir essa deficiência e obter a melhor configuração de RNA para a tarefa, foram testados 5 diferentes parâmetros das redes neurais, que totalizaram 432 diferentes configurações de RNA, cujo desempenho foi medido por 5 diferentes índices de comparação.

Dessa forma foi possível observar o comportamento de diferentes configurações de RNA e a sua influência na capacidade de geração de modelos de previsão de PCS de RSU, e concluir quais configurações são mais adequadas para a tarefa, além de se ter desenvolvido um modelo com um nível de acurácia consideravelmente superior aos encontrado na literatura da área.

2.1 Redes neurais artificiais

RNA são uma técnica de inteligência artificial de inspiração neurofisiológica, que têm a intenção de mimetizar o comportamento do cérebro humano. As RNA são sistemas paralelos distribuídos compostos por unidades de processamentos simples (neurônios artificiais) que processam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos essas conexões estão associadas a pesos, os quais armazenam o conhecimento adquirido pelo modelo e servem para ponderar a entradas recebidas por cada neurônio da rede (Braga, Carvalho e Ludermir, 2007).

Em uma RNA, cada neurônio artificial (Fig. 5) processa o padrão de entrada recebido da seguinte forma:

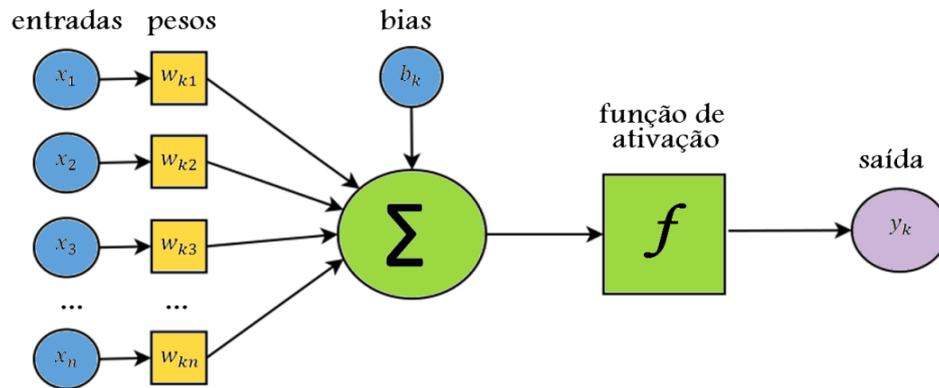


Figura 5. Representação esquemática do neurônio artificial.

1. O neurônio (k) recebe um conjunto de valores de entrada (x_i);
2. Cada valor recebido é multiplicado pelo peso (w_{ki}) associado àquela entrada e àquele neurônio;
3. À soma desses valores é adicionado um coeficiente de deslocamento linear (b_k) associado àquele neurônio;
4. Ao valor obtido é aplicada uma função associada ao neurônio (f_k), chamada função de ativação;
5. O valor de retorno da função, dependendo da função escolhida, é comparado ao seu limiar de ativação, e o resultado (y_k) é enviado ao(s) próximo(s) neurônio(s) da rede;

Matematicamente falando, o funcionamento de um neurônio artificial (k) pode ser descrito pela Eq. (1).

$$y_k = f_k \left(\sum_{i=1}^n w_{ki} x_i + b_k \right) \quad (1)$$

Onde:

y_k : sinal de saída do neurônio;

f_k : função de ativação;

w_{ki} : i -ésimo peso do k -ésimo neurônio;

x_i : i -ésimo valor do vetor de entradas do neurônio, associado ao peso w_{ki} ;

b_k : *bias* de ajuste do valor de entrada da função de ativação;

n : tamanho do vetor de entradas do neurônio.

A função de ativação, a quantidade de neurônios utilizada e a forma de conexão desses neurônios definem a arquitetura, enquanto que os pesos, que geralmente são modificados durante a fase de treinamento, carregam o conhecimento adquirido pela rede.

Existem diversas formas de disposição e conexão dos neurônios em uma rede neural. Essas diferentes configurações são denominadas arquiteturas, e são essas diferentes arquiteturas que tornam as redes neurais aplicáveis a uma grande gama de problemas.

Uma configuração possível em uma RNA é o seu número de camadas. Uma camada é identificada como um conjunto de neurônios que têm a mesma configuração e são processados individual e paralelamente. Uma das categorias de RNA mais utilizadas atualmente são as

chamadas MLP (*Multi Layer Perceptron*) (Fig. 6). Essas redes são formadas por diversos conjuntos de neurônios (*perceptrons*), dispostos geralmente em 2 ou mais camadas.

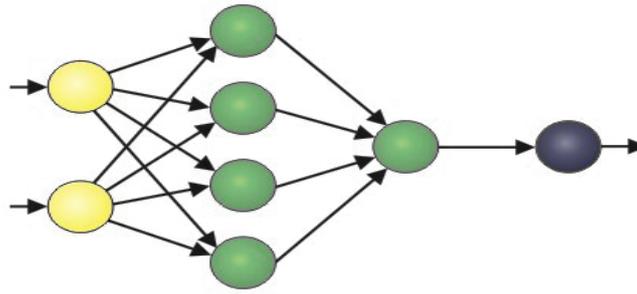


Figura 6. Rede neural MLP de 3 camadas.

O “conhecimento” de uma RNA é obtido através de um processo denominado treinamento. Em redes MLP para aproximação de funções o treinamento é feito a partir da apresentação de um conjunto de valores de entrada e o respectivo valor de saída esperado. A rede calcula então a diferença entre o valor de saída da rede e o valor esperado – o erro da rede – e reajusta seus pesos sinápticos de acordo com um algoritmo previamente configurado. A repetição desse processo é executada até que um critério de parada estabelecido seja atingido.

Esse processo possibilita que a rede desenvolva por si mesma um conhecimento que não estava previamente definido em sua configuração inicial. A capacidade de extrair um conhecimento novo a partir de um conjunto de dados já conhecido é uma das principais características das RNA.

2.2 Configurações de RNA utilizadas

Considerando-se o objetivo da pesquisa, foi definido como escopo do trabalho que seriam utilizadas RNA MLP do tipo *feedforward*, onde a informação na rede segue um fluxo unidirecional da primeira até a última camada, com processamento paralelo e simultâneo a cada camada. Para o treinamento foram selecionados somente algoritmos do tipo *backpropagation*, onde os pesos da rede são ajustados proporcionalmente à sua contribuição para o valor de saída.

Dessa forma foram definidos 5 parâmetros de configuração das redes neurais: a quantidade de neurônios na camada de entrada (4 opções), a quantidade de neurônios na camada intermediária (3 opções), o algoritmo de treinamento (4 opções), a função de ativação na camada intermediária (3 opções) e a função de ativação na camada de saída (3 opções) (Fig. 7). Para cada configuração foram efetuados 30 processamentos, e o melhor resultado de cada rede foi utilizado no comparativo de desempenho.

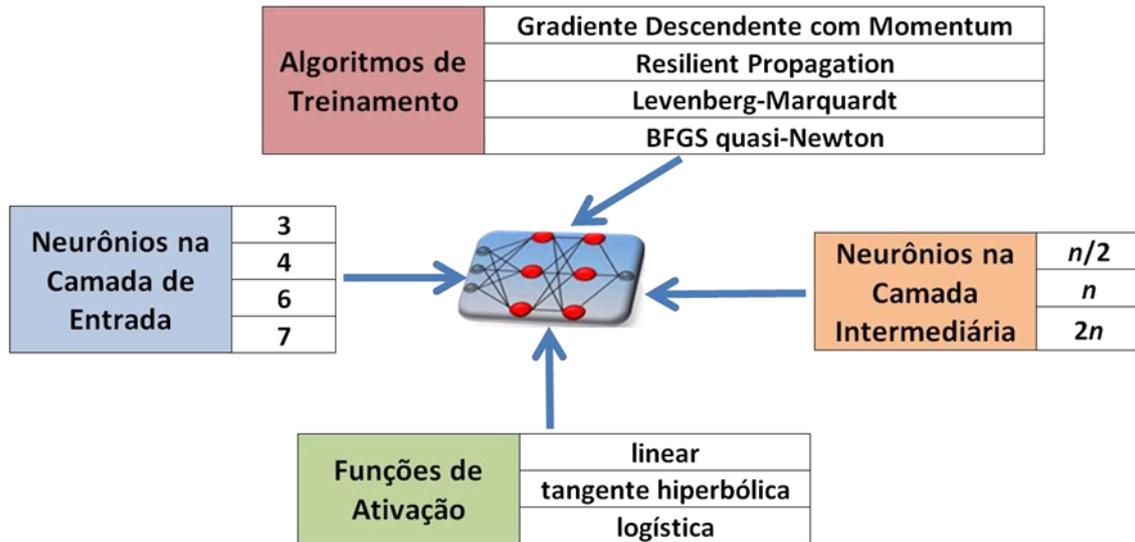


Figura 7. Parâmetros de configuração das redes neurais testadas.

Para o desenvolvimento das redes neurais foi utilizado o aplicativo MatLab, de propriedade da empresa MathWorks. A aplicação foi desenvolvida na versão 7.11.0.584 (R2010b) do MatLab, em sua versão para Windows 64 bits. A opção pelo MatLab deve-se à sua boa velocidade de processamento e disponibilidade dos algoritmos de treinamento selecionados. A programação do MatLab foi feita por codificação direta, sem utilização de assistentes de automatização.

2.3 A base de dados

Como visto, para o treinamento da rede é necessária uma base de dados de padrões de entrada e seus respectivos valores de saída esperados. Nesse trabalho tal base de dados foi obtida a partir de uma compilação feita por Meraz *et al.* (2003), que continha 100 amostras com 7 variáveis independentes de entrada (teores de umidade, C, H, N, O, S e cinzas do RSU) e 1 variável dependente de saída, o PCS, divididas em 7 grupos de resíduos (Fig. 8).

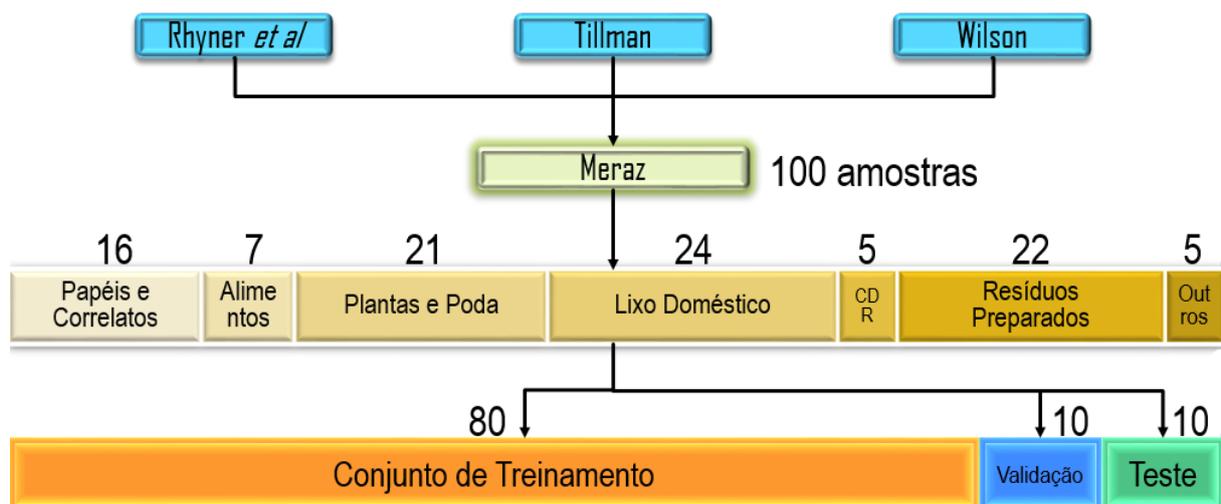


Figura 8. Origem e distribuição da base de dados.

Essa base de 100 amostras foi dividida em 3 conjuntos de dados: o conjunto de treinamento (80 amostras), o conjunto de validação (10 amostras) e o conjunto de testes (10 amostras). O conjunto de treinamento foi utilizado para o ajuste da matriz de pesos da rede de acordo com o algoritmo selecionado. O conjunto de validação serviu como critério de desempenho da rede, e os pesos das conexões dos neurônios da rede utilizados em sua configuração final são aqueles que oferecem o melhor resultado na previsão do conjunto de validação. Por fim, o conjunto de teste foi utilizado para se estabelecer o desempenho da rede em relação a um padrão de dados não conhecido previamente, e sobre esse conjunto foram feitas todas as análises de desempenho das redes e dos modelos matemáticos da literatura.

2.4 Ferramentas estatísticas

Como critério de desempenho das configurações de RNA e também dos modelos matemáticos foram utilizados 3 diferentes índices, todos largamente utilizados na literatura do tema.

O primeiro índice utilizado foi o EMPA (*Erro Médio Percentual Absoluto*) (Eq. (2)). O EMPA é o índice de desempenho direto de acurácia mais utilizado entre os autores de modelos de previsão de poder calorífico, e traz as vantagens de ser adimensional e proporcional ao valor observado, o que torna possível a comparação direta entre resultados obtidos a partir de amostras distintas, ou mesmo entre amostras tomadas em diferentes unidades de medida.

$$EMPA = \frac{1}{n} \sum_{i=1}^n \frac{|PCS_{pi} - PCS_{mi}|}{PCS_{mi}} \times 100 \quad (2)$$

Onde:

$EMPA$: erro médio percentual absoluto;

PCS_{pi} : i -ésimo valor de PCS previsto pelo modelo;

PCS_{mi} : i -ésimo valor de PCS medido experimentalmente;

n : quantidade de amostras de teste.

Segundo Chang *et al.* (2007), os modelos de previsão podem ser classificados em 4 categorias, conforme seu resultado pelo índice EMPA. A Tabela 1 resume essa classificação.

Tabela 1. Classificação de modelos de previsão segundo o índice EMPA (Chang *et al.*, 2007).

Classificação do Modelo	Resultado
Excelente	$EMPA < 10\%$
Bom	$10\% \leq EMPA < 20\%$
Aceitável	$20\% \leq EMPA < 50\%$
Inaceitável	$EMPA \geq 50\%$

Neste trabalho, o EMPA foi utilizado como índice de classificação de desempenho tanto para as configurações de RNA como para os modelos matemáticas.

O segundo índice utilizado foi o DMA (Desvio Médio Absoluto) (Eq. (3)). O DMA determina a diferença média entre os valores previstos pelo modelo e os valores efetivamente medidos. Para evitar o cancelamento da soma de valores positivos e negativos, utilizam-se os valores absolutos das diferenças. Sua principal vantagem é que ele apresenta os resultados na mesma unidade de medida da variável de saída do modelo, permitindo que seja feita uma análise quantitativa da acurácia das previsões.

$$DMA = \frac{1}{n} \sum_{i=1}^n |PCS_{pi} - PCS_{mi}| \quad (3)$$

O terceiro e último índice utilizado foi o EMV (Erro Médio de Viés) (Eq. (4)). Diferentemente dos índices anteriores, o EMV não mede a acurácia do modelo de previsão, mas sim o grau e a direção de seu viés. Quanto maior o valor de EMV do modelo maior seu viés. Se o EMV for positivo, pode-se dizer que o modelo tende a sobrestimar os valores previstos. Caso contrário, o modelo tende a prever valor subestimados.

$$EMV = \frac{1}{n} \sum_{i=1}^n \frac{(PCS_{pi} - PCS_{mi})}{PCS_{mi}} \times 100 \quad (4)$$

3 RESULTADOS E DISCUSSÃO

Dentre as 432 configurações testadas, 80 (18,5%) obtiveram um índice EMPA abaixo de 10%, sendo considerados modelos de excelente desempenho. Outras 72 (16,7%) configurações tiveram um desempenho considerado bom. Com desempenho aceitável ficaram 43 (10,0%) redes, enquanto que 237 (54,9%) redes tiveram um índice EMPA considerado inaceitável, acima de 50% (Fig. 9).

Essas diferenças de resultados entre as configurações de RNA mostram que a escolha dos parâmetros de configuração da rede é extremamente importante para seu desempenho.

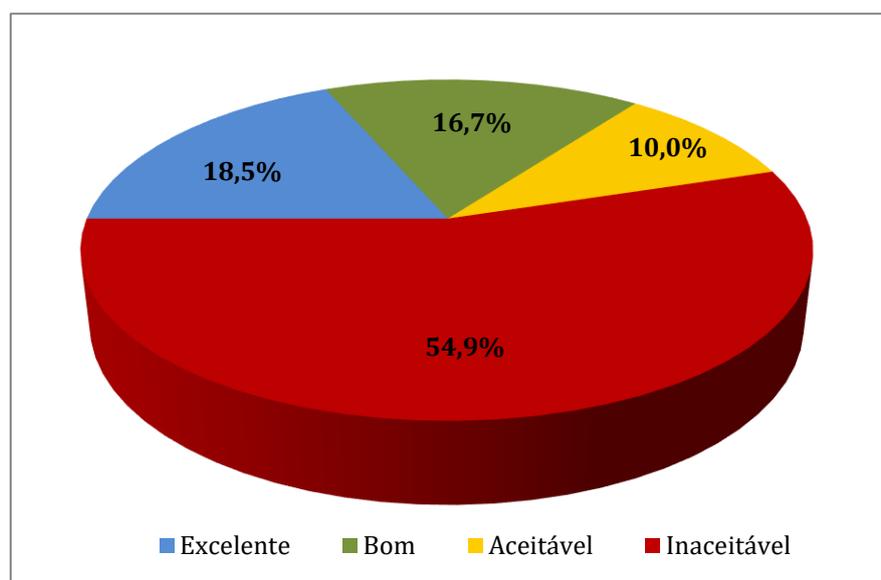


Figura 9. Proporção de adequação das RNA testadas.

Para entender como cada parâmetro utilizado na configuração das RNA afetou o desempenho da rede, foi realizada uma análise detalhada por parâmetro, e os resultados são apresentados a seguir.

3.1 Quantidade de neurônios na camada de entrada

Conforme pode ser observado pela Tabela 2, a quantidade de neurônios na camada de entrada, e, por conseguinte, o número de variáveis utilizadas como padrão de entrada para o treinamento da rede, não tiveram influência significativa no desempenho das redes neurais.

Tabela 2. Desempenho das RNA por neurônios na camada de entrada.

Neurônios na Camada de Entrada		
Quantidade	Excelente	Inaceitável
3	22,2%	25,3%
4	23,5%	24,1%
6	24,7%	25,7%
7	29,6%	24,9%

Embora seja possível observar, entre as redes de desempenho excelente, uma ligeira predominância entre aquelas que utilizam um maior número de neurônios na camada de entrada, a proporção praticamente igualitária entre as redes de desempenho inaceitável mostra que outros parâmetros são muito mais importantes na configuração da rede do que as variáveis utilizadas com padrão de entrada.

3.2 Quantidade de neurônios na camada intermediária

O segundo parâmetro testado foi a quantidade de neurônios da camada intermediária. Após testarem várias configurações, Dong *et al.* (2003) utilizaram 7 neurônios na camada intermediária, a mesma quantidade de neurônios da camada de entrada, pois essa configuração foi a que trouxe o melhor desempenho para a rede por eles desenvolvida. Já Ogwueleka e Ogwueleka (2010) utilizaram 10 neurônios na camada intermediária, e afirmaram em seu trabalho que “a melhor arquitetura de rede neural é aquela com o dobro de neurônios na camada oculta em relação ao número de neurônios na camada de entrada”.

Neste trabalho optou-se por uma abordagem que abrangesse três possibilidades $n_i = \frac{n_e}{2}$, $n_i = n_e$ e $n_i = 2n_e$, respectivamente para quantidade de neurônios na camada intermediária (n_i) menor, igual e maior do que a quantidade de neurônios na camada de entrada (n_e). A Tabela 3 mostra as combinações testadas.

Tabela 3. Quantidade de neurônios na camada intermediária.

Neurônios na Camada de Entrada	Neurônios na Camada Intermediária		
	n/2	n	2n
7	4	7	14
6	3	6	12
4	2	4	8
3	2	3	6

Na análise de desempenho das redes neurais segundo o critério de quantidade de neurônios na camada intermediária (Tabela 4) foi possível observar uma ligeira predominância das redes que mantiveram nessa camada o mesmo número de neurônios da camada de entrada, com 38,3% das redes de melhor desempenho. As redes com o dobro de neurônios foram as que tiveram a menor participação entre as redes com desempenho excelente (29,6%).

Tabela 4. Desempenho das RNA por neurônios na camada intermediária.

Fator	Neurônios na Camada Intermediária	
	Excelente	Inaceitável
n/2	32,1%	34,6%
n	38,3%	32,9%
2n	29,6%	32,5%

3.3 Algoritmos de treinamento

Quatro diferentes algoritmos de treinamento foram testados para as RNA. Considerando trabalhos revistos na área, esperava-se que o algoritmo Levenberg-Marquardt tivesse um desempenho superior aos demais, por se tratar de um algoritmo de rápida convergência e bastante eficaz em tarefas de aproximação de funções. Entretanto, como pode ser observado na Tabela 5, o algoritmo Levenberg-Marquardt não foi o que apresentou os melhores resultados, embora tenha obtido um ótimo desempenho para a tarefa.

Tabela 5. Desempenho das RNA por algoritmo de treinamento.

Algoritmo de Treinamento		
Algoritmo	Excelente	Inaceitável
Gradiente Descendente com <i>Momentum</i>	0,0%	38,8%
Resilient Propagation	21,0%	20,7%
Levenberg-Marquardt	35,8%	20,3%
BFGS quasi-Newton	43,2%	20,3%

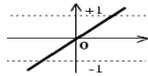
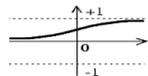
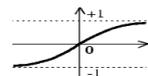
A Tabela 5 mostra que o algoritmo BFGS quasi-Newton foi o que obteve a maior participação entre as redes de desempenho excelente, com 43,2%, enquanto que o algoritmo Levenberg-Marquardt foi utilizado por 35,8% das melhores redes. Entre as redes de desempenho inaceitável os dois algoritmos aparecem com 20,3%, mostrando que outro fator foi responsável pelo desempenho ruim das redes com esse algoritmo.

Já o algoritmo Gradiente Descendente com *Momentum* não esteve presente em nenhuma rede com desempenho excelente, o que mostra que esse algoritmo não é apropriado para a tarefa de previsão de PCS. O algoritmo Resilient Propagation foi utilizado por 21,0% das redes de excelente desempenho e 20,7% das redes de desempenho inaceitável, mostrando que esse algoritmo pode ser utilizado para a tarefa, embora não seja o mais recomendável.

3.4 Função de ativação na camada intermediária

Entre todos os parâmetros testados, os que tiveram maior influência no desempenho da rede neural foram as funções de ativação. Foram utilizadas 2 funções sigmóides, a função tangente hiperbólica e a função logística, além da função linear (Tabela 6).

Tabela 6. Funções de ativação utilizadas nas configurações da rede.

Função de Ativação	Expressão Matemática	Gráfico	Limites
linear	$f(x) = x$		$]-\infty ; \infty[$
logística	$f(x) = \frac{1}{1 + e^{-x}}$		$]0 ; 1[$
tangente hiperbólica	$f(x) = \frac{2}{(1 + e^{-2x}) - 1}$		$] -1 ; 1[$

Conforme pode ser observado na Tabela 7, as funções sigmóides tiveram um desempenho muito superior à função linear, sendo responsáveis por 86,4% das redes de desempenho excelente, com uma ligeira predominância da função logística sobre a função tangente hiperbólica (44,4% e 42,0%, respectivamente). Já a função linear foi utilizada em 43,5% das redes de desempenho inaceitável.

Tabela 7. Desempenho das RNA por função de ativação na camada intermediária.

Função de Ativação na Camada Intermediária		
Função	Excelente	Inaceitável
logística	44,4%	28,7%
tangente	42,0%	27,8%
linear	13,6%	43,5%

3.5 Função de ativação na camada de saída

Para a camada de saída foram utilizadas as mesmas funções da camada intermediária. Entretanto, o resultado observado foi justamente o inverso daquele, com grande predominância da função linear nas redes de melhor desempenho (Tabela 8).

Tabela 8. Desempenho das RNA por função de ativação na camada de saída.

Função de Ativação na Camada Intermediária		
Função	Excelente	Inaceitável
logística	0,0%	60,8%
tangente	12,3%	29,5%
linear	87,7%	9,7%

A função linear foi utilizada em 87,7% das redes de desempenho excelente, e somente 9,7% das redes de desempenho inaceitável a utilizaram, sendo que o desempenho ruim desse grupo de 9,7% das redes pode ser creditado ao algoritmo de treinamento, já que todas utilizavam o algoritmo Gradiente Descendente com *Momentum*. A função logística não aparece entre as redes de desempenho excelente, e foi utilizada em 60,8% das redes de pior desempenho. A função tangente hiperbólica teve um resultado um pouco melhor, aparecendo em 12,3% e 29,5% das redes de melhor e pior desempenho, respectivamente.

3.6 Redes neurais de melhor desempenho

A partir da análise detalhada dos itens de configuração das redes neurais é possível identificar um conjunto de parâmetros que tem maiores possibilidades de gerar modelos de previsão de PCS de RSU com melhor desempenho.

Analisando as 432 configurações testadas, foi possível observar que as redes de melhor desempenho utilizaram, de modo geral, 7 (29,6%) neurônios na camada de entrada, com o mesmo número (38,3%) de neurônios na camada intermediária, algoritmo de treinamento BFGS quasi-Newton (43,2%), com função de ativação logística (44,4%) na camada intermediária e linear (87,7%) na camada de saída, sendo que, efetivamente, a rede configurada com exatamente esses parâmetros (*rna-771-bfg-lp*) foi a que obteve o melhor resultado geral dentre todas as configurações.

A Tabela 9 mostra as 10 configurações de RNA com melhor desempenho na previsão do PCS pelo índice EMPA. A melhor configuração, a *rna-771-bfg-lp*, obteve um índice EMPA de somente 2,9%, mostrando a excelente adequação dessa técnica para a tarefa proposta.

Tabela 9. As 10 configurações de RNA de melhor desempenho.

Pos	Identificação da Rede	Neurônios na Camada de Entrada	Neurônios na Camada Oculta	Algoritmo de Treinamento	Função de Ativação na Camada Oculta	Função de Ativação na Camada de Saída	EMPA
1	<i>rna-771-bfg-lp</i>	7	7	BFGS	Logística	Linear	2,9%
2	<i>rna-631-bfg-tp</i>	6	3	BFGS	TangHip	Linear	3,3%
3	<i>rna-6121-bfg-pp</i>	6	12	BFGS	Linear	Linear	3,3%
4	<i>rna-771-lm-tp</i>	7	7	LM	TangHip	Linear	3,4%
5	<i>rna-441-bfg-lp</i>	4	4	BFGS	Logística	Linear	3,6%
6	<i>rna-741-bfg-lp</i>	7	4	BFGS	Logística	Linear	3,7%
7	<i>rna-661-bfg-tp</i>	6	6	BFGS	TangHip	Linear	3,7%
8	<i>rna-441-lm-tt</i>	4	4	LM	TangHip	TangHip	3,8%
9	<i>rna-441-lm-tp</i>	4	4	LM	TangHip	Linear	3,8%
10	<i>rna-321-lm-tp</i>	3	2	LM	TangHip	Linear	3,9%

Não obstante a melhor configuração de rede obedecer exatamente aos critérios individuais de melhor resultado, é possível observar pela Tabela 9 que os índices de desempenho das 10

melhores configurações são muito próximos, e mostram a grande adaptabilidade das redes neurais à tarefa de previsão de PCS.

Quanto aos algoritmos de treinamento, tanto o BFGS quasi-Newton quanto o Levenberg-Marquardt foram utilizados em redes de excelente desempenho, comprovando a eficácia dos algoritmos de treinamento de segunda ordem para essa tarefa.

As funções sigmóides também são maioria como função de ativação da camada intermediária, embora a rede com o 3º melhor desempenho utilize a função linear nesse parâmetro. E, conforme já observado, a função linear tem larga predominância como função de ativação na camada de saída, embora a rede com função tangente hiperbólica apareça com o 8º melhor desempenho.

3.7 Comparação com modelos matemáticos

Como última etapa do trabalho foi realizada a comparação do índice EMPA obtido pela rede de melhor desempenho (*rna-771-bfg-lp*) rede e modelos matemáticos da literatura, a partir das amostras do conjunto de teste.

Para seleção dos modelos matemáticos utilizados para comparação, 20 diferentes equações encontradas na literatura foram testadas, e somente 8, cujos desempenhos foram considerados excelente pelo índice EMPA, fizeram parte da análise comparativa. A Tabela 10 lista as equações utilizadas e seus autores.

Tabela 10. Modelos matemáticos de previsão de PCS.

Autor	Equação	Ref
Lloyd e Davenport	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3578\%C + 1,1357\%H - 0,0845\%O + 0,0594\%N + 0,1119\%S)$	1
Boie	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3517\%C + 1,1625\%H - 0,1109\%O + 0,0628\%N + 0,1109\%S)$	1
Wilson	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3279\%C + 1,5330\%H - 0,1668\%O + 0,0242\%N + 0,0928\%S)$	1
Meraz et al	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3708\%C + 1,1124\%H - 0,1391\%O + 0,3178\%N + 0,1391\%S)$	1
Strache e Lant	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3406\%C + 1,4324\%H - 0,1532\%O + 0,1047\%S)$	2
Dulong	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3383\%C + 1,443\%H - 0,1804\%O + 0,0942\%S)$	2
Dulong-Berthelot	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,3414\%C + 1,4445\%H - 0,125\%O - 0,125\%N + 0,093\%S)$	2
Shi et al	$HHV = \left(1 - \frac{\%H_2O}{100}\right)(0,349\%C + 1,01\%H - 0,0812\%O - 0,174\%N + 0,886\%S)$	3

Onde:

HHV: poder calorífico superior previsto (do inglês, *higher heating value*);

%*H₂O*: teor de umidade (% de massa);

%*C*: teor de carbono (% de massa);

%*H*: teor de hidrogênio (% de massa);

%*O*: teor de oxigênio (% de massa);

%*N*: teor de nitrogênio (% de massa);

%*S*: teor de enxofre (% de massa);

Referências: (1) Meraz *et al* (2003); (2) Channiwala e Parikh (2002); (3) Shi *et al* (2016).

Uma vez selecionados os modelos matemáticos, procedeu-se com o cálculo dos índices de desempenho para as amostras do conjunto de teste. A partir dos resultados obtidos, foi elaborada a Tabela 11, que sintetiza os resultados dos testes de desempenho dos modelos matemáticos e da rede *rna-771-bfg-lp*.

Tabela 11. Comparativo entre a *rna-771-bfg-lp* e modelos selecionados da literatura.

Modelo	EMPA (%)	DMA (MJ/kg)	EMV (%)
<i>rna-771-bfg-lp</i>	2,9	0,404	2,3
Boie	4,6	0,879	4,5
Wilson	5,1	1,083	1,8
Strache e Lant	5,3	1,070	3,8
Shi <i>et al</i>	6,1	0,581	5,9
Dulong	6,3	1,231	-0,2
Dulong-Berthelot	6,8	1,516	10,8
Meraz <i>et al</i>	7,4	1,197	6,4
Lloyd e Davenport	8,7	1,566	8,7

Como é possível observar, a *rna-771-bfg-lp* obteve um EMPA = 2,9%, com um DMA = 0,404 MJ/kg, e EMV = 2,3%. Isso mostra que a rede tem um excelente acurácia, com um erro médio de previsão de valor de PCS (DMA) de apenas 0,404 MJ/kg. Esse nível de erro de previsão é bastante baixo, considerando-se que o valor médio de PCS das amostras era de aproximadamente 14 MJ/kg. A *rna-771-bfg-lp* ainda obteve um EMV de 2,3%, mostrando que o modelo tem um sutil viés de alta.

O melhor modelo matemático testado foi o de Boie, com um índice EMPA = 4,6%, e um DMA = 0,897 MJ/kg, o que pode ser considerado um bom resultado para um modelo obtido por regressão linear. O índice EMV de 4,5% obtido pelo modelo de Boie mostra que suas previsões também tendem a superestimar os valores de PCS.

O modelo de Meraz, que foi desenvolvido a partir da mesma base de dados, obteve um índice EMPA de 7,4%, dentro da classificação de excelente, mas muito acima do obtido pela *rna-771-bfg-lp*. Acredita-se que esse desempenho consideravelmente superior do modelo gerado por redes neurais deve-se a possíveis relações não lineares existentes entre o PCS e a composição elementar do RSU, que são captadas durante o treinamento da rede neural, mas não estão presentes em modelos matemáticos gerados a partir de regressão linear múltipla.

O modelo de Dulong foi o que demonstrou conter o menor viés, com um EMV de -0,2%. Entretanto, seu índice EMPA de 6,3% mostra que suas previsões, embora pouco enviesadas, têm um acurácia inferior a outros modelos analisados.

4 CONCLUSÕES

Esta pesquisa propôs-se a explorar as potencialidades das redes neurais como ferramenta para o desenvolvimento de modelos de previsão de PCS de RSU. Os resultados obtidos permitem concluir que as RNA são uma excelente técnica para essa tarefa, pois o desempenho obtido pelo modelo de previsão gerado pela RNA obteve índices de desempenho superiores aos modelos matemáticos da literatura.

Entretanto, também foi possível observar que o desempenho de um RNA está intimamente relacionado a algumas de suas configurações. Dentre essas, o algoritmo de treinamento escolhido e as funções de ativação das camadas da rede exercem um papel preponderante na adequação da rede para a tarefa de previsão.

Certamente a metodologia utilizada para o desenvolvimento do modelo também pode ser utilizada com outras bases de dados, a partir de análises gravimétrica ou imediata, e possivelmente também terá bons resultados quando empregada para outras biomassas, além do RSU.

AGRADECIMENTOS

Os autores gostariam de agradecer à UFABC e à Petrobrás S.A., pelo suporte oferecido e disponibilização de recursos para a realização deste trabalho.

REFERÊNCIAS

ABRELPE – Associação Brasileira de Empresas de Limpeza Pública e Resíduos Especiais, 2015. *Panorama dos Resíduos Sólidos no Brasil - 2014*.

Braga, A.P.; Carvalho, A.P.L.F.; e Ludermir, T.B. *Redes Neurais Artificiais – Teoria e Aplicações*. 2ª edição. Rio de Janeiro: Ed. LTC, 2007. 224 páginas.

Brasil, 2010. Ministério do Meio Ambiente. *Plano Nacional de Resíduos Sólidos (PNRS)* - Lei n. 12305, de 2 de Agosto de 2010. Institui a Política Nacional de Resíduos Sólidos; altera a Lei n. 9605, de 12 de fevereiro de 1998; e dá outras providências. Brasília.

- Chang, Y.F.; Lin, C.J.; Chyan, J.M.; Chen, I.M.; Chang, J.E., 2007. *Multiple regression models for the lower heating value of municipal solid waste in Taiwan*. Journal of Environmental Management 85, pp. 891-899. Elsevier Ltd.
- Channiwala, S.A.; Parikh, P.P., 2002. *A unified correlation for estimating HHV of solid, liquid and gaseous fuels*. Fuel 81, pp. 1051-1063. Elsevier Science Ltd.
- DDMA – Doutores do Meio Ambiente, 2010. *Caracterização Gravimétrica, Físico-Química e Elementar dos Resíduos Sólidos Urbanos*.
- Dong, C.; Jin, B.; Li, D., 2003. *Predicting the heating value of MSW with feed forward neural network*. Waste Management 23, pp. 103-106. Elsevier Science Ltd.
- Eurostat, 2015. European Statistics. *Treatment of waste database*. Disponível em: <<http://ec.europa.eu/eurostat/data/database>>. Acesso em: 20/jun/2016.
- FEAM – Fundação Estadual do Meio Ambiente, 2012. *Aproveitamento Energético de Resíduos Sólidos Urbanos: Guia de Orientações para Governos Municipais de Minas Gerais*.
- Haykin, S., 2001. *Redes Neurais – Princípios e prática*. 2ª edição. Porto Alegre: Editora Bookman. 900 páginas.
- Kathiravale, S.; Yunus, M.N.M.; Sopian, K.; Samsuddin, A. H.; Rahman, R.A., 2003. *Modeling the heating value of Municipal Solid Waste*. Fuel 82, pp. 1119–1125. Elsevier Science Ltd.
- Maier, S.; Oliveira, L.B., 2014. *Economic feasibility of energy recovery from solid waste in the light of Brazil's waste policy: The case of Rio de Janeiro*. Renewable and Sustainable Energy Reviews 35, pp. 484–498. Elsevier Ltd.
- Meraz, L.; Dominguez, A.; Kornhauser, I.; Rojas, F., 2003. *A thermochemical concept-based equation to estimate waste combustion enthalpy from elemental composition*. Fuel 82, pp. 1499–1507. Elsevier Science Ltd.
- Ogwueleka, T. C.; Ogwueleka, F. N., 2010. *Modelling Energy Content of Municipal Solid Waste Using Artificial Neural Network*. Iranian Journal of Environmental Health Science & Engineering, Vol. 7, No. 3, pp. 259-266. BioMed Central.
- Ryu, C.; Shin, D., 2013. *Combined heat and power from Municipal Solid Waste: Current status and issues in South Korea*. Energies 6, pp. 45-57. MDPI.
- Shi, H.; Mahinpey, N.; Aqsha, A.; Silbermann, R., 2016. *Characterization, thermochemical conversion studies, and heating value modeling of municipal solid waste*. Waste Management 48, pp. 34-47. Elsevier Ltd.
- Zhou, H.; Meng, A.; Long, Y.; Li, Q.; Zhang, Y., 2014. *An overview of characteristics of municipal solid waste fuel in China: Physical, chemical composition and heating value*. Renewable and Sustainable Energy Reviews 36, pp. 107-122. Elsevier Ltd.