

PREDIÇÃO DE TEMPERATURA DE EBULIÇÃO E FUSÃO DE COMPOSTOS ORGÂNICOS UTILIZANDO CG - RNA

Aline S. LIMA; Hugo L. B. BUARQUE

Centro Federal de Educação Tecnológica do Ceará, Av. 13 de Maio, 2081, Benfica, 60.040-531, Fortaleza, Ceará,
telefone: (85)33073647, fax:33073711 , e-mail: alinesantos_lima@hotmail.com
e-mail: hbruarque@cefetce.br

RESUMO

O desenvolvimento de modelos de predição de propriedades de compostos puros e de misturas é de grande interesse científico e industrial, particularmente, na implementação de sistemas simuladores de processos industriais. Propriedades físicas, químicas e termodinâmicas precisas de compostos orgânicos puros são essenciais para o desenvolvimento e simulação de processos de separação na indústria química. Porém, nem sempre é possível determinar experimentalmente, ou obter na literatura, valores precisos e confiáveis de propriedades de interesse industrial à medida que são necessárias. Por isso, técnicas de estimação são usualmente necessárias. Métodos de contribuição de grupos (CG) e modelos de redes neurais artificiais (RNA) são exemplos dessas técnicas. Neste estudo, foi utilizada uma rede neural artificial progressiva com uma única camada oculta para encontrar correlações entre grupos de contribuição e as temperaturas de fusão e ebulição de, aproximadamente, 900 compostos orgânicos puros. A RNA foi treinada por um algoritmo de retropropagação de erros. O modelo desenvolvido foi comparado a outros propostos na literatura, bem como a modelos lineares, mostrando-se bastante satisfatório e apresentando erros relativos para predição das temperaturas de ebulição e fusão inferiores a 5% e 10%, respectivamente.

Palavras-chave: estimação de propriedades, compostos orgânicos, contribuição de grupos, redes neurais artificiais.

1. INTRODUÇÃO

Os pontos normais de ebulição e fusão têm encontrado amplo uso nos domínios de processos de simulação e desenvolvimento de produtos (WEN e QIANG, 2002). Porém, nem sempre é possível encontrar valores experimentais dessas propriedades na literatura ou obtê-los em laboratório. Por esse motivo, métodos de estimação são normalmente utilizados.

Métodos de contribuição de grupos (CG) têm se mostrado bastante eficazes na predição de propriedades de compostos a partir de informações sobre a sua estrutura molecular (FREDESLUND, JONES, PRAUSNITZ, 1975; JOBACK, REID, 1987; CONSTANTINOU, GANI, 1994; MARRERO-MOREJÓN, PARDILLO-FONTDEVILA, 1999). Nestes métodos, a propriedade de um composto é função de parâmetros estruturalmente dependentes, aqui denominados grupos de contribuição.

Recentemente, redes neurais artificiais (RNA) também têm sido muito utilizadas como técnica de predição de propriedades de misturas usando a composição de grupos ou subgrupos de compostos como informação de entrada (CÔCCO et al., 2005). A arquitetura de rede mais comumente utilizada é a *feedforward*, onde os nós são agrupados em três tipos de camada (i.e. entrada, oculta e saída). Os dados são alimentados na camada de entrada e sendo depois transferidos para as camadas subseqüentes (SOUSA et al., 2007).

Neste trabalho foram utilizados grupos de contribuição como parâmetros de entrada para o desenvolvimento de modelos de redes neurais artificiais *feedforward* (CG-RNA) e de regressão linear múltipla (CG-RLM) para predição dos pontos normais de ebulição e fusão de diversas classes de compostos orgânicos. Os modelos obtidos foram comparados entre si, bem como a outras técnicas de contribuição de grupos. Os modelos CG-RNA desenvolvidos foram os que apresentaram os resultados mais satisfatórios, fornecendo erros relativos aproximadamente iguais a 1% e 10%, respectivamente, para as temperaturas de ebulição e fusão.

2. DESENVOLVIMENTO DOS MODELOS

2.1. Obtenção de propriedades

Foram compilados a partir da literatura (LIDE, 1995; POLING et al., 2001) para regressão e teste dos modelos desenvolvidos, os pontos normais de ebulição e fusão de aproximadamente 900 compostos de diferentes classes orgânicas (hidrocarbonetos, haletos, oxigenados e nitrogenados).

2.2. Determinação dos grupos moleculares

Os grupos moleculares utilizados como variáveis de entrada no desenvolvimento dos modelos CG-RNA e CG-RLM foram os mesmos propostos por Joback e Reid (1987), desde que existentes nos compostos orgânicos compilados. A Figura 1 apresenta os 34 grupos moleculares utilizados como variáveis de entrada nos modelos aqui propostos.

<u>GRUPOS HIDROCARBÔNICOS ACÍCLICOS</u>					<u>GRUPOS OXIGENADOS</u>				
- CH ₃	- CH ₂ -	> CH -	> C <	= CH ₂	- OH	- OH	- O -	- COO -	> CO
G1	G2	G3	G4	G5	G20	G21	G22	G23	G24
= CH -	= C <	= C =	≡ CH	≡ C -	- CHO	- COOH			
G6	G7	G8	G9	G10	G25	G26			
<u>GRUPOS HIDROCARBÔNICOS CÍCLICOS</u>					<u>GRUPOS NITROGENADOS</u>				
- CH ₂ -	> CH -	> C <	= CH -	= C <	- NH ₂	> NH	> N -	- CONH ₂	- CONH -
G11	G12	G13	G14	G15	G27	G28	G29	G30	G31
<u>GRUPOS HALOGENADOS</u>					- CON <	- C≡N	- NO ₂		
- F	- Cl	- Br	- I		G32	G33	G34		
G16	G17	G18	G19						

Figura 1 – Grupos moleculares de contribuição para compostos orgânicos

2.3. Desenvolvimento dos modelos CG-RNA e CG-RLM.

Na estimação dos pontos normais de ebulição e fusão dos compostos orgânicos puros compilados, utilizou-se o método de Buarque (2006) e o pacote computacional MATLAB[®] (THE MATHWORKS, 2002) na obtenção de modelos lineares e no desenvolvimento de redes neurais artificiais progressivas, com uma única camada oculta, alimentadas com a frequência com que os grupos moleculares selecionados ocorrem num dado composto puro e fornecendo a propriedade desejada.

As RNA foram treinadas por um algoritmo de retropropagação de erros usando o método de Levenberg-Marquardt (HAGAN, MENHAJ, 1994). Para o treinamento e teste das RNA foi utilizada a técnica de validação cruzada modificada (BUARQUE, 2006), onde o conjunto de dados disponível foi dividido aleatoriamente em três subconjuntos: treinamento, validação e teste. Esta divisão aleatória foi repetida em cinco distintos ciclos.

Em cada ciclo eram testados três diferentes arranjos de funções de ativação. As funções de ativação usadas nos neurônios da camada oculta eram: a função logística (Equação 1), a função tangente hiperbólica (Equação 2) e uma função linear (Equação 3). Para ativação da camada de saída utilizou-se somente a função logística. Também, as RNA testadas tiveram seus pesos sinápticos aleatoriamente inicializados pelo menos 50 vezes.

$$f(x) = [1 + \exp(x)]^{-1} \quad (1)$$

$$f(x) = \tanh(x) \quad (2)$$

$$f(x) = x \quad (3)$$

As redes neurais foram testadas com os neurônios da camada oculta variando desde um único nó até um número máximo pré-estabelecido. O número máximo de nós ocultos permitidos foi estabelecido com base numa regra proposta por Buarque (2006), segundo a qual o número de padrões disponíveis para o treinamento e validação da rede deveria ser de pelo menos 1,5 vezes o número de pesos sinápticos e polarizações (*biases*) originadas para a RNA.

Os modelos CG-RLM testados consistiam de equações lineares com as mesmas variáveis de entrada dos modelos CG-RNA, usando o método dos mínimos quadrados para a regressão linear múltipla dos seus parâmetros.

Os erros obtidos para os dados correlacionados com estes modelos foram calculados com o intuito de avaliar a linearidade da relação entre os grupos de contribuição selecionados e a propriedade considerada, bem como para compará-los aos modelos CG-RNA obtidos.

2.4. Avaliação da capacidade correlativa e preditiva dos modelos CG-RNA e CG-RLM desenvolvidos e aqueles selecionados a partir da literatura.

A avaliação da habilidade correlativa e do poder preditivo dos modelos de contribuição de grupos selecionados da literatura, bem como dos modelos CG-RNA e CG-RLM desenvolvidos neste estudo foi realizada utilizando-se os seguintes parâmetros estatísticos: média dos erros relativos, δ , erro relativo máximo, δ_m , e erro padrão de estimativa, S , expressos respectivamente nas Equações 4, 5 e 6 ($y_{e,i}$ e $y_{c,i}$ são, respectivamente, os i -ésimos padrões de saída experimental e calculado pelo modelo; N é o número total de padrões de saída). Mais especificamente, com relação aos modelos CG-RNA, os erros padrão de estimativa dos dados de regressão (treinamento e validação) foram utilizados para se determinar o melhor conjunto de pesos sinápticos para o modelo desenvolvido a partir dos diferentes ciclos de validação cruzada.

$$\delta = \frac{\sum_{i=1}^N \frac{|y_{e,i} - y_{c,i}|}{y_{e,i}}}{N} \quad (4)$$

$$\delta_m = \text{máximo} \left(\frac{|y_{e,i} - y_{c,i}|}{y_{e,i}} \right) \quad (5)$$

$$S = \sqrt{\sum_{i=1}^N (y_{e,i} - y_{c,i})^2 / N} \quad (6)$$

3. RESULTADOS E DISCUSSÃO

3.1. Correlação e predição da temperatura normal de ebulição de compostos orgânicos utilizando os modelos de CG-RNA desenvolvidos

Para o desenvolvimento e avaliação dos modelos considerados neste estudo, dos 912 dados de temperaturas normais de ebulição compilados, 775 registros constituíram a série de regressão (593 padrões para o treinamento e 182 para a validação da RNA) e 137 constituíram uma série de teste para verificação do poder preditivo do modelo. Todos os 34 subgrupos moleculares considerados neste estudo podem ser encontrados nos compostos da base de dados compilada, sendo assim utilizados como nós de entrada das redes neurais testadas. Ressalte-se ainda que nestes conjuntos de dados, os valores mínimos e máximos para a temperatura de ebulição foram, respectivamente, 111,65 K e 735,15 K.

O modelo CG-RNA obtido para a predição do ponto normal de ebulição apresenta 16 (dezesseis) neurônios na camada oculta, cada um deles utilizando a função tangente hiperbólica na modulação do sinal de entrada. O neurônio da camada de saída utiliza a função logística. Os pesos sinápticos e os biases obtidos para este modelo estão dispostos nas Tabelas 1 e 2. Os coeficientes ajustados na regressão linear múltipla são apresentados na Tabela 3.

A Tabela 4 mostra os parâmetros estatísticos de avaliação determinados tanto para os modelos desenvolvidos como para os demais modelos testados para predição das temperaturas normais de ebulição. A Figura 2 retrata a qualidade do ajuste e predição das temperaturas normais de ebulição estimadas para o modelo CG-RNA desenvolvido.

O modelo CG-RNA proposto apresentou um erro relativo máximo superior ao erro do modelo linear avaliado, mas demonstra claramente sua superioridade ao compararmos os demais erros avaliados para todos os modelos.

3.2. Correlação e predição da temperatura normal de fusão de compostos orgânicos puros utilizando os modelos CG-RNA desenvolvidos

Dos 754 dados de temperaturas normais de fusão compilados para o treinamento e avaliação dos modelos considerados neste estudo, 641 registros constituíram as séries de regressão (490 dados de treinamento e 151 registros para validação) e os outros 113 registros compuseram a série de teste. Os valores mínimos e máximos de temperaturas normais de fusão para os dados utilizados foram, respectivamente, de 77,15 K e 658,15 K.

O melhor modelo CG-RNA obtido para esta propriedade apresentava treze neurônios ocultos e utiliza a função logística como função de ativação nos neurônios tanto da camada oculta como da camada de saída. Todos os 34 subgrupos moleculares considerados neste estudo podem ser encontrados nos compostos da base de dados compilada, sendo assim utilizados como nós de entrada das redes neurais testadas. Os pesos sinápticos e biases obtidos para este modelo estão dispostos, respectivamente, nas Tabelas 5 e 6. A Tabela 7 dispõe os coeficientes ajustados na regressão linear múltipla (modelo CG-RLM).

Tabela 1 - Pesos sinápticos no modelo CG-RNA para temperatura de ebulição dos compostos orgânicos.

NÓS OCULTOS																	
NÓS DA CAMADA DE ENTRADA		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	G1	-1.750	0.348	0.374	0.871	-2.076	-0.687	-0.274	0.236	0.832	-1.112	-1.852	1.790	-0.981	-1.117	1.400	1.498
	G2	0.808	-1.833	0.509	-0.730	-0.022	0.665	-2.165	-4.992	-0.051	1.337	0.119	2.197	-0.388	-4.811	-0.045	0.052
	G3	1.261	-1.096	0.733	-0.786	1.836	0.078	-1.054	-2.028	-0.102	0.734	0.649	0.060	0.815	-0.958	0.161	-0.454
	G4	-0.238	-0.229	-0.969	-0.294	-0.166	1.295	0.553	-0.637	0.577	-0.786	1.548	1.622	-1.476	0.156	2.519	1.999
	G5	0.706	-0.493	0.110	-0.860	0.953	0.693	1.807	1.446	0.677	-0.663	-0.473	-0.170	0.774	0.781	1.026	0.122
	G6	0.871	-1.186	-0.487	-0.873	-0.524	-1.110	-0.771	0.490	0.559	0.627	-0.749	-0.451	-0.639	-0.739	-0.679	-0.176
	G7	-1.874	0.188	0.309	0.804	0.137	0.837	-0.660	0.475	-0.021	0.439	0.674	-0.045	1.130	-0.647	1.367	2.284
	G8	-0.610	1.102	0.412	0.868	-0.722	-1.439	1.972	0.685	-0.433	0.777	0.575	0.993	0.392	-0.857	0.175	-0.273
	G9	-2.661	1.682	-1.066	1.358	0.490	1.215	-0.239	3.221	0.342	-0.095	-1.875	-2.198	-1.079	0.812	0.883	0.665
	G1	1.758	-0.391	0.701	-3.069	-1.782	-0.834	-1.270	-0.254	-0.059	-0.493	0.721	0.145	0.525	-0.508	0.848	-0.435
	G1	0.064	2.367	1.389	-4.026	-1.034	-0.272	-0.600	0.012	4.848	0.106	-4.313	2.184	0.042	-2.545	-1.900	1.822
	G1	0.822	-1.786	-0.061	0.589	2.395	0.823	-0.632	-0.576	-1.681	-0.327	1.449	-0.977	1.715	-0.044	-0.002	1.294
	G1	1.806	-1.042	0.106	-0.161	0.513	-1.309	0.572	-0.066	-0.715	-0.914	0.949	0.441	0.625	0.439	0.834	1.173
	G1	-1.383	-1.543	-0.772	0.597	-0.495	-0.641	-0.523	-0.039	-0.854	0.380	-0.134	-0.222	1.966	-1.653	-0.203	-1.242
	G1	-0.756	-0.701	0.493	1.330	2.379	0.572	-1.253	0.073	-1.480	-0.458	0.009	-1.552	-1.410	-1.768	1.784	2.076
	G1	1.355	-3.606	-1.445	-0.380	-1.220	0.131	-0.186	-0.833	-0.554	-1.253	0.611	-1.154	0.403	0.286	0.179	-0.200
	G1	0.837	-2.679	1.052	0.141	0.460	-1.202	3.115	0.542	-0.467	-0.953	0.859	0.542	1.469	-0.422	0.356	-2.121
	G1	0.280	-1.552	-0.439	-0.015	0.484	0.179	0.658	-0.768	-0.289	0.639	0.245	0.419	-0.024	-1.057	0.357	-1.788
	G1	0.081	-1.442	1.544	0.803	2.865	0.079	-2.084	-1.701	-3.004	0.536	5.781	-2.287	-0.398	-1.451	2.282	-2.487
	G2	-2.919	-0.843	1.057	2.692	1.534	-0.888	-3.552	-0.915	-2.152	-1.695	1.122	-1.367	-0.576	-4.629	-3.121	-3.668
	G2	-0.288	-0.650	-2.023	0.245	0.458	-1.233	-0.202	1.657	1.183	-1.076	-3.255	-2.727	-1.150	-2.184	-0.725	0.649
	G2	0.017	0.150	0.978	0.216	-0.571	0.751	0.478	1.286	0.633	-0.535	-0.700	1.575	0.163	0.029	-0.010	1.228
	G2	-0.534	0.962	-0.013	-1.089	0.201	0.560	2.500	0.986	2.532	-1.300	1.479	0.113	0.467	-0.895	-1.978	1.074
	G2	0.272	-0.833	-0.872	-0.355	-1.856	-1.145	-0.458	-0.277	0.070	0.559	-0.808	-1.093	-1.144	-1.073	-0.584	0.737
	G2	-1.417	1.201	0.601	2.438	-1.148	-1.058	-1.096	-0.433	-2.531	0.362	1.585	2.408	0.240	-0.969	-0.891	-0.283
	G2	-2.989	0.714	-0.720	1.921	-0.475	-0.870	0.256	1.222	0.738	1.165	-1.529	1.168	0.126	-1.455	-0.631	-1.144
	G2	0.043	-0.870	1.270	-0.884	-1.692	0.182	-0.590	0.767	0.457	-0.721	-0.427	-1.142	-0.792	-0.553	-1.063	-0.281
	G2	-0.158	1.455	1.069	0.555	0.615	1.341	1.585	0.840	-0.926	-0.459	0.515	1.056	0.764	-0.833	-0.433	0.446
	G2	-0.069	0.506	1.405	2.108	-1.203	1.185	1.596	-0.422	-1.988	1.488	-0.245	-0.906	-0.007	-0.362	0.986	0.277
	G3	0.419	-0.477	-1.425	-2.059	0.359	-0.018	-1.191	1.342	1.495	1.256	-0.417	0.460	0.443	-2.812	-0.527	1.273
	G3	-1.034	-0.164	-0.406	0.838	0.860	-0.275	0.052	0.154	0.631	-0.797	-1.260	1.564	1.419	-0.801	-1.003	1.573
	G3	-0.621	-0.287	-0.497	0.479	-0.973	-0.366	1.137	-0.884	1.447	1.843	-2.380	-1.855	-0.719	-2.964	0.782	-0.272
	G3	-0.077	-0.676	-1.253	-5.283	-1.498	2.179	-0.932	3.894	-2.257	-2.332	1.240	-1.426	-2.267	3.871	-0.210	1.702
	G3	1.578	-0.951	1.069	-2.395	-0.779	0.651	1.929	-1.073	2.270	1.878	-2.173	-1.445	1.187	-1.695	0.053	-1.241
NÓ DE SAÍDA		-1.248	-1.109	-1.100	-1.653	1.072	-0.284	0.414	2.027	-1.502	0.746	-0.766	0.525	-0.374	-1.338	-1.402	2.472

Tabela 2 – Biases no modelo CG-RNA para temperatura normal de ebulição dos compostos orgânicos.

CAMADA OCULTA																CAMADA DE SAÍDA
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
2.062	2.354	-3.197	0.466	-1.222	-3.294	0.086	-2.853	0.161	-0.994	1.177	-0.373	-1.603	6.261	1.159	1.033	-0.272

Tabela 3 – Coeficientes lineares do modelo CG-RLM para os pontos normais de ebulição

a ₀	a ₁	a ₂	a ₃	a ₄	a ₅	a ₆	a ₇	a ₈	a ₉	a ₁₀	a ₁₁
201.690	30.839	22.139	7.997	-10.604	14.852	23.028	4.497	28.890	12.597	26.664	26.881
a ₁₂	a ₁₃	a ₁₄	a ₁₅	a ₁₆	a ₁₇	a ₁₈	a ₁₉	a ₂₀	a ₂₁	a ₂₂	a ₂₃
13.810	-5.639	28.766	18.457	4.315	48.967	72.291	99.386	73.939	62.490	16.955	71.566
a ₂₄	a ₂₅	a ₂₆	a ₂₇	a ₂₈	a ₂₉	a ₃₀	a ₃₁	a ₃₂	a ₃₃	a ₃₄	
66.624	83.077	152.490	57.354	37.696	-2.576	219.490	180.520	101.680	130.340	115.180	

Tabela 4 – Parâmetros estatísticos de avaliação dos modelos selecionados para estimação da temperatura normal de ebulição.

MODELO	Temperatura de ebulição		
	δ (%)	δ_m (%)	S (K)
Joback e Reid (1987)	3,99	98,64	25,33
Constantinou e Gani (1994)	11,26	133,2	76,39
CG-RLM	3,28	42,71	0,67
CG-RNA proposto	1,05	42,97	0,36

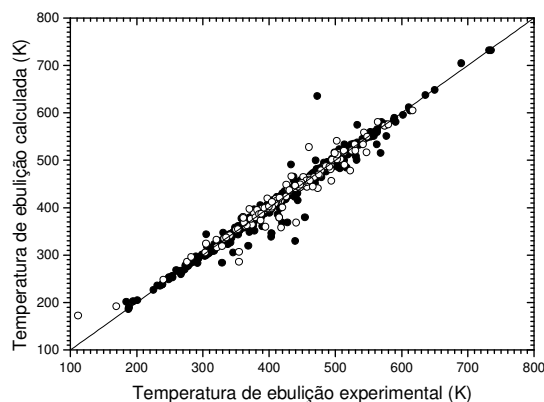


Figura 2 – pontos normais de ebulição estimados pelo CG-RNA versus pontos observados. Pontos cheios são dados da série de regressão; pontos vazados são dados da série de teste.

A partir da Tabela 8, pode-se perceber que o modelo CG-RNA desenvolvido para estimação da temperatura de fusão de compostos orgânicos, quando comparado ao modelo linear obtido mostrou ser um pouco superior àquele. Porém, quando comparado ao modelo de contribuição de grupos avaliado, o modelo CG-RNA desenvolvido apresentou um erro padrão de estimação, aproximadamente, 50 vezes menor que àquele. A Figura 3 mostra o gráfico comparativo entre as propriedades experimentais contra as calculadas pelos modelos CG-RNA.

Os valores dos pontos normais de fusão estimados pelo modelo CG-RNA desenvolvido apresentaram certa dispersão em relação aos valores experimentais. Esse desvio pode ser devido a presença de alguns isômeros presentes no banco de dados utilizados nesse estudo.

Tabela 5 - Pesos sinápticos no modelo CG-RNA para temperatura normal de fusão de compostos orgânicos.

NÓS OCULTOS														
NÓS DA CAMADA DE ENTRADA		1	2	3	4	5	6	7	8	9	10	11	12	13
	G1	-2.142	6.808	1.238	-3.227	-5.269	-4.808	-0.069	-2.839	-0.700	0.273	0.438	3.394	-2.653
	G2	0.243	1.894	-0.922	-0.285	-6.091	0.742	2.655	5.054	5.724	-2.672	4.795	-3.964	1.216
	G3	-1.618	0.378	-0.385	1.990	-3.761	0.845	1.225	-2.699	-2.110	2.789	-2.729	-1.592	0.728
	G4	0.435	1.752	-2.043	2.031	-2.118	1.711	-3.156	-0.614	-6.474	-2.111	-2.124	-2.349	0.837
	G5	-0.842	2.843	2.590	0.629	-1.796	1.839	0.959	-1.989	0.554	4.058	-1.333	-1.502	-0.697
	G6	-0.642	-0.181	-3.535	2.602	2.510	-1.326	2.232	-0.870	-2.948	-1.850	-0.615	1.800	-0.434
	G7	0.272	0.825	4.682	-1.577	1.311	2.416	2.059	-2.558	0.829	0.446	2.639	2.407	2.594
	G8	1.757	1.020	1.100	2.057	-2.097	-2.918	1.403	3.098	-0.400	-1.525	1.304	-4.424	-2.779
	G9	1.596	-0.076	1.783	-0.579	2.676	0.282	2.346	0.465	-0.049	0.109	-1.098	-2.253	0.979
	G10	0.069	-1.955	2.444	-0.894	-1.792	1.292	-0.523	-1.549	2.631	3.668	-0.006	-3.305	-0.048
	G11	0.192	0.738	-1.451	-1.167	3.100	-1.761	-2.213	5.402	2.831	-2.388	0.248	-1.878	-2.959
	G12	2.324	-2.669	-0.898	-4.137	-1.465	-2.555	0.143	0.914	4.185	-4.022	-1.789	3.298	-0.428
	G13	-0.844	-2.175	-1.962	-1.816	2.082	-2.391	2.537	0.801	0.228	1.163	1.690	1.115	3.419
	G14	3.157	-0.723	0.439	0.014	6.383	3.462	1.125	-0.143	3.340	-0.780	-0.745	-0.349	-1.546
	G15	1.271	0.412	-1.940	0.670	0.998	2.032	0.436	-3.377	1.309	-2.591	1.340	4.162	-4.192
	G16	1.765	0.756	-3.396	-0.243	4.335	1.043	-0.427	3.677	-0.801	2.773	1.739	0.875	-0.940
	G17	2.075	-3.402	-3.096	-0.124	-1.419	-2.413	-0.174	-1.624	0.964	2.053	1.750	-0.648	-1.876
	G18	1.465	-1.513	-3.794	-4.220	0.808	-0.822	-0.604	0.358	-0.340	-0.362	0.131	-0.243	-1.657
	G19	0.746	-1.665	-3.323	-0.587	0.439	-2.268	-0.774	0.792	-0.323	-1.184	0.363	-2.997	5.649
	G20	-2.430	-0.521	-1.101	-0.487	0.156	-0.489	0.009	-4.383	0.348	2.116	1.897	1.100	0.877
	G21	-2.682	-1.792	-4.094	2.692	0.132	0.880	-2.995	-0.261	2.034	-3.029	1.987	-5.576	0.636
	G22	1.287	2.612	1.736	0.467	-0.875	-2.952	2.877	-0.114	1.156	-0.658	-2.804	0.220	1.046
	G23	-1.294	2.895	0.184	0.848	-0.270	-1.258	-0.937	0.448	0.615	-3.211	1.093	1.470	2.536
	G24	0.737	-2.885	-3.912	1.444	-0.638	-0.766	0.252	-2.049	-0.683	0.156	2.074	1.507	3.504
	G25	-3.279	-1.773	-2.096	1.417	-1.586	2.992	1.491	-1.029	-0.102	-0.466	1.694	0.264	1.491
	G26	0.217	-3.319	0.298	1.696	2.317	-1.331	1.323	-4.971	1.937	1.229	-0.890	3.274	-0.443
	G27	-1.295	-2.351	-2.601	2.303	-1.085	-1.151	0.809	-1.600	0.847	-1.646	-1.330	2.522	2.807
	G28	1.392	0.184	-2.166	1.802	1.524	1.694	-1.834	-0.166	-4.110	-1.243	2.943	1.388	2.152
	G29	-4.862	2.948	0.604	0.358	3.937	2.493	-0.814	0.525	-1.208	0.041	0.201	-0.653	3.463
	G30	2.288	1.628	-1.560	-1.807	-0.319	2.420	0.924	-1.853	2.053	-3.358	1.700	0.230	-2.691
	G31	2.196	2.330	0.011	-3.060	-2.739	-1.505	-2.742	-4.207	-1.192	-0.222	-1.096	3.913	1.916
	G32	0.024	1.657	2.441	-2.069	1.489	0.597	1.754	-2.881	2.955	1.439	0.310	2.986	-1.129
	G33	4.949	-3.104	-0.125	5.174	4.290	0.058	-1.226	-1.175	-2.931	0.421	0.617	-3.191	0.956
	G34	1.446	-0.012	-0.722	-1.902	-1.227	0.515	2.460	-0.674	1.428	-0.519	-1.424	-3.028	-1.982
NÓ DE SAÍDA		0.614	0.639	-1.308	-0.312	0.704	0.091	0.411	-1.402	0.857	-0.960	1.899	-1.157	1.311

Tabela 6 – Biases no modelo CG-RNA para temperatura normal de fusão dos compostos orgânicos

CAMADA OCULTA													CAMADA DE SAÍDA
1	2	3	4	5	6	7	8	9	10	11	12	13	
0.485	-1.562	5.196	2.072	-2.176	-0.278	-10.069	5.899	-2.700	-0.260	-10.434	1.525	-4.187	0.570

Tabela 7 – Coeficientes lineares do modelo CG-RLM para os pontos normais de fusão.

a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}
112.040	4.783	9.710	4.282	29.890	5.865	8.314	0.711	3.746	28.365	12.491	12.682
a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}	a_{22}	a_{23}
3.006	40.010	12.515	22.990	7.406	25.902	38.700	74.458	70.086	51.312	9.769	48.153
a_{24}	a_{25}	a_{26}	a_{27}	a_{28}	a_{29}	a_{30}	a_{31}	a_{32}	a_{33}	a_{34}	
61.487	64.607	149.730	73.998	35.733	19.715	220.610	152.430	93.964	60.052	80.747	

Tabela 8 – Parâmetros estatísticos de avaliação dos modelos selecionados para estimação da temperatura normal de fusão.

MODELO	Temperatura de fusão		
	δ (%)	δ_m (%)	S (K)
Joback e Reid (1987)	14,52	131,47	52,13
CG-RLM	11,75	114,55	1,42
CG-RNA proposto	9,56	96,72	1,18

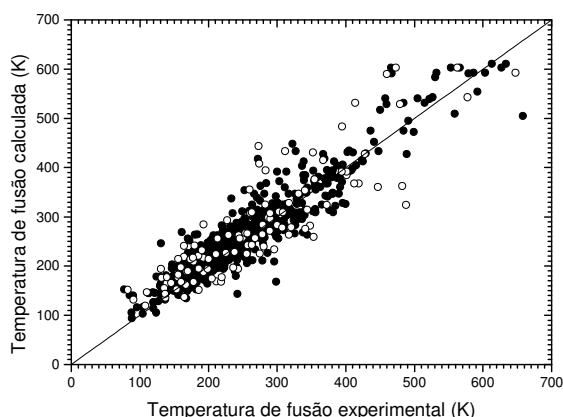


Figura 3– pontos normais de fusão estimados pelo CG-RNA versus observados. Pontos cheios são dados da série de regressão; pontos abertos são dados da série de teste.

4. CONCLUSÃO

Os modelos CG-RNA desenvolvidos neste estudo para predição das temperaturas normais de ebulição e fusão se mostraram bastante superiores a modelos de contribuição de grupos selecionados a partir da literatura para a predição das mesmas propriedades, resultando em erros padrão de estimação inferiores a 1,5 K. Assim sendo, os modelos CG-RNA podem ser satisfatoriamente utilizados na predição das propriedades consideradas.

Os valores calculados para os pontos normais de fusão, utilizando o modelo CG-RNA desenvolvido, apresentaram uma maior dispersão com relação aos valores experimentais. Esses desvios, provavelmente, são devidos a forte dependência da conformação da molécula sobre a temperatura de fusão. Vale ressaltar que os modelos desenvolvidos neste estudo não diferenciam esses isômeros.

REFERÊNCIAS

BUARQUE, H.L.B. **Predição de propriedades de gasolinas a partir das suas composições. Tese de doutoramento.** Fortaleza: Universidade Federal do Ceará, 2006. 206p.

CÔCCO, L.C.; YAMAMOTO, C.I.; VON MEIEN, O.F. **Study of correlations for physicochemical properties of Brazilian gasoline.** Chemometrics and Intelligent Laboratory Systems, v. 76, p. 55-63, 2005.

CONSTANTINOU, L.; GANI, R. **New group contribution method for estimating properties of pure compounds.** AIChE Journal, v. 40, p. 1697-1710, 1994.

FREDENSLUND, Aa.; JONES, R.L.; PRAUSNITZ, J.M. **Group-Contribution Estimation of Activity Coefficients in Nonideal Mixtures.** AIChE Journal, v. 21, p.1086-1099, 1975.

HAGAN, M.T.; MENHAJ, M.B. **Training feedforward networks with the Marquardt algorithm.** IEEE Transactions on Neural Networks, v. 5 (6): 989-993, 1994.

JOBACK, K.G.; REID, R.C. **Estimation of pure-component properties from group-contributions.** Chemical Engineering Communications, v. 57, p. 233-243, 1987.

LIDE, D.R. **Handbook of Chemistry and Physics.** 76th edition. Boca Raton: CRC Press, 1995.

MARRERO-MOREJÓN, J.; PARDILLO-FONTDEVILA, E. **Estimation of Pure Compound Properties Using Group-Interaction Contributions.** AIChE Journal, v. 45, p. 615-621, 1999.

POLING, B.E.; PRAUSNITZ, J.M.; O'CONNELL, J.P. **The properties of gases and liquids.** Fifth edition. New York: McGraw Hill, 2001.

SOUSA, S.I.V.; MARTINS, F.G.; ALVIM-FERRAZ, M.C.M.; PEREIRA, M.C., **Multiple Linear Regression and Artificial Neural Networks Based on Principal Components to Predict Ozone Concentrations.** Environmental Modelling & Software, v.22, p. 97-103, 2007.

THE MATHWORKS, INC. MATLAB® - **The Language of Technical Computing**, version 6.5.0.196271, Release 13.0.1., PDF DOCUMENTATION in CD-ROM, 2002.

WEN, X.; QIANG, Y. **Group vector space method for estimation melting and boiling points of organic compounds.** Ind. Eng. Chem. Res., v. 41, p. 5534-5537, 2002.

AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo fomento financeiro através de bolsa de iniciação científica PIBIC/CNPq/CEFETCE.