

OBTENÇÃO DE CORRELAÇÕES PARA PREVISÃO DE PROPRIEDADES FÍSICO-QUÍMICAS DA GASOLINA A PARTIR DE SUA COMPOSIÇÃO QUÍMICA UTILIZANDO REDES NEURONAS ARTIFICIAIS

Lílian C. Côcco¹, Carlos I. Yamamoto², Oscar F. von Meien³

¹ Universidade Federal do Paraná, Usinas Piloto A, Centro Politécnico s/n, Jardim das Américas, CEP 81.531- 970, Curitiba, Paraná, Brasil, lili@engquim.ufpr.br.

² Universidade Federal do Paraná, Usinas Piloto B, Centro Politécnico s/n, Jardim das Américas, CEP 81.531- 970, Curitiba, Paraná, Brasil, cijama@engquim.ufpr.br.

³ Universidade Federal do Paraná, Usinas Piloto B, Centro Politécnico s/n, Jardim das Américas, CEP 81.531- 970, Curitiba, Paraná, Brasil, oscar@engquim.ufpr.br.

Resumo – As propriedades físico-químicas da gasolina são importantes referências de sua qualidade, devendo ser monitoradas a fim de mantê-la, respeitando-se leis ambientais e parâmetros técnicos. Cada propriedade é, em geral, uma complicada função da composição química da gasolina, podendo ser representada por diversos tipos de correlações matemáticas, as quais podem ser utilizadas para previsão de suas propriedades como uma tentativa para reduzir custos laboratoriais. Entretanto, as correlações existentes não estão ajustadas à gasolina brasileira, cuja composição química é modificada com a adição de álcool etílico anidro, fornecendo resultados não condizentes com a realidade. Por este motivo existe a necessidade de se encontrar correlações apropriadas para a gasolina nacional e aplicá-las para obtenção de previsões mais acuradas, com valores próximos às análises experimentais padrão ASTM ou NBR. Este trabalho descreve a utilização de redes neuronais artificiais como um método de correlação para previsão de algumas propriedades importantes da gasolina, tais como: curva de destilação, densidade relativa a 20°C e pressão de vapor Reid. Uma rede com treinamento supervisionado por retropropagação é usada neste trabalho. Após estudos preliminares e planejamento experimental, 35 amostras de gasolina coletadas no estado do Paraná, a partir de um universo de 1000 amostras, foram escolhidas e submetidas à análise cromatográfica e ensaios laboratoriais padronizados. Um programa específico para classificação de hidrocarbonetos denominado DHA (Detailed Hydrocarbon Analysis) gerou dados de entrada para a rede neuronal, enquanto que os ensaios físico-químicos padronizados forneceram os dados de saída. A capacidade preditiva das redes obtidas foi muito boa para as propriedades estudadas, compatível com a precisão das medidas experimentais. Os desvios obtidos ficaram em média próximos de 1%. Os resultados alcançados estão próximos das faixas de incerteza de medição dos ensaios físico-químicos e comprovam que a ferramenta utilizada consegue correlacionar a composição química e as propriedades físico-químicas estudadas.

Palavras-chave: gasolina; rede neuronal; retropropagação, cromatografia gasosa.

Abstract – The fuel physico-chemical properties are very important references of its quality. They must be monitored in order to maintain and/or improve production quality and match environmental standards as well. Each property is a complex function of the gasoline composition. Different mathematical correlations can be used to predict the fuel properties as an attempt to reduce laboratorial costs. Correlations developed overseas cannot be applied direct to the Brazilian gasoline due to the addition of significant amounts of ethylic alcohol. In the following work gas chromatography (GC) and backpropagation artificial neural networks (ANN) were used to predict important properties of gasoline such as: ASTM distillation curve, relative density at 20°C and Reid vapor pressure. After an experimental planning 35 samples of brazilian gasoline, from a universe of 1000, were submitted to chromatographic analysis and laboratory tests, supplying inputs and outputs to the ANN. The obtained networks predictive capacity proved to be effective for all studied properties, agreeable to experimental precision.

Keywords: gasoline, neural networks, backpropagation, gas chromatography.

1. Introdução

As propriedades físico-químicas da gasolina são importantes referências de sua qualidade, devendo ser monitoradas a fim de mantê-la, respeitando-se leis ambientais e parâmetros técnicos. Cada propriedade é, em geral, uma complicada função da composição química da gasolina, podendo ser representada por diversos tipos de correlações matemáticas, as quais podem ser utilizadas para previsão de suas propriedades como uma tentativa para reduzir custos laboratoriais. Isso se deve ao fato de não estarem ajustadas à gasolina nacional, que recebe uma quantidade apreciável de álcool etílico anidro, alterando significativamente a composição química e conseqüentemente suas propriedades. Por este motivo observou-se a necessidade de encontrar novas correlações matemáticas, ajustadas à gasolina nacional, para que a previsão de suas propriedades físico-químicas esteja próxima a valores obtidos em ensaios experimentais padrão ASTM ou NBR.

Para a montagem de correlações matemáticas, utiliza-se dados gerados por qualquer técnica analítica de identificação, tais como espectrometria por infravermelho, ultravioleta ou de massa, espectrometria de ressonância magnética nuclear de hidrogênio (RMN ^1H) ou carbono 13 (RMN ^{13}C) e cromatografia gasosa, como dados de entrada. Os ensaios padrão ASTM, como densidade à 20°C, curva de destilação da gasolina e pressão de vapor Reid, fornecem os dados de saída para posterior aplicação em um método de regressão de dados. O método matemático deve ser capaz de encontrar a relação existente entre as entradas e saídas, obtendo-se desta forma as correlações desejadas. Em resumo, para o desenvolvimento deste tipo de trabalho existe a necessidade da escolha de um método analítico confiável e um tipo de regressão de dados.

Comparando-se com as técnicas analíticas mais avançadas para identificação de compostos, como a ressonância magnética nuclear e espectrometria de massa atômica, a cromatografia gasosa com detector de ionização de chama destaca-se pela facilidade de interpretação dos cromatogramas gerados. É uma técnica extensamente estudada na indústria do petróleo, o que a torna uma técnica analítica clássica para hidrocarbonetos. Concentrações mínimas de compostos pertencentes à gasolina são detectados, a partir de pequenos volumes injetados de amostra, com excelente separação dos picos, não ocorrendo superposição de compostos como normalmente ocorre ao utilizar as técnicas de infravermelho e ultravioleta. Levando-se em conta as vantagens observadas, optou-se pela utilização da cromatografia gasosa com detector FID como técnica analítica para identificação química da gasolina.

Escolhida a técnica de identificação química, passa-se à escolha do método matemático para regressão de dados. Diversos trabalhos comparativos entre as técnicas mais utilizadas foram efetuados por Meusinger, 2001, Andrade, 1999 e Yang, 2001. A regressão linear múltipla com algoritmo PLS ou PCR e redes neuronais com algoritmo de retropropagação ou genético são os mais citados. Os trabalhos comparativos demonstraram que as redes neuronais possuem capacidade igual ou superior no ajuste de dados. Levando-se em conta que as redes neuronais não necessitam de um modelo matemático para iniciar o estudo, apenas de dados de entrada e saída os quais se supõe que exista relação entre eles, optou-se por este método de regressão de dados. Uma rede neuronal por retropropagação foi escolhida por ser a mais utilizada para este tipo de estudo, além de ser a mais simples com relação aos outros tipos de redes neuronais existentes.

Para a aplicação de redes neuronais, bem como qualquer tipo de regressão de dados, os resultados obtidos são influenciados principalmente pela qualidade dos dados retirados para o estudo, já que dados não confiáveis inseridos dentro do conjunto podem atrapalhar o ajuste. Para obter dados de boa qualidade deve-se observar a calibração e verificação periódica de equipamentos, confiabilidade na medição de variáveis, cálculo de incertezas e planejamento experimental para seleção de amostras. Este estudo pode simplificar a determinação das propriedades físico-químicas da gasolina, que com apenas um ensaio cromatográfico pode ser possível condensar os demais ensaios padronizados e reduzir custos laboratoriais.

2. Rede Neuronal por Retropropagação

Redes neuronais artificiais (RNA's) são estruturas matemáticas capazes de implementar padrões de associações, desenvolvendo automaticamente modelos implícitos. Uma RNA por retropropagação é comumente denominada perceptron multicamadas (multilayer perceptron) e é extensamente estudada. Detalhes podem ser encontrados em trabalhos similares (van Leeuwen, 1994) (Elsharkwy, 2001) (Yang, 2001).

Um problema comum que deve ser observado ao se trabalhar com redes por retropropagação é o "overfitting", que é uma memorização dos dados de treinamento. Este problema ocasiona uma perda da capacidade preditiva da rede, pois observa-se pequenos desvios de previsão para os dados usados na fase de treino, mas grandes desvios quando novos dados de entrada são utilizados. O que pode ocasionar o overfitting é a redundância dos dados e excesso de neurônios na camada oculta. Para evitar o overfitting os erros do treino e do teste são monitorados. Assim que o erro de treinamento cair muito ao mesmo tempo em que o erro de validação subir, ficará caracterizado o overfitting, e a última iteração antes do ocorrido é que deverá ser considerada, finalizando-se o treinamento. O oposto do overfitting é o "underfitting" ou sub-ajuste, o qual é resultante de uma rede com poucos neurônios ocultos, que não consegue representar o sistema que se deseja modelar. Para observar tanto o "overfitting" como o "underfitting", aconselha-se além da fase de treino e teste, aplicar-se outro teste com novos exemplos de entrada para a rede e verificar o seu poder preditivo.

Outro problema é a presença de um “outlier” no conjunto de validação ou treino, que pode afetar as conclusões a respeito da performance da rede. Um outlier é um par entrada/saída que fornece à rede informação errada devido a um grande erro experimental. É muito difícil detectar um outlier, a não ser que seja observada alguma anormalidade durante a execução do ensaio. A decisão final pode ser feita comparando-se os resultados da previsão para um modelo de rede com e sem o ponto.

A presença de uma forma distribuída de não linearidade e a alta conectividade da rede faz com que a análise teórica de um perceptron multicamadas seja de difícil entendimento e o uso de neurônios ocultos fazem com que o processo de aprendizado seja difícil de visualizar. Portanto, existe a necessidade de simplificação da rede para um número pequeno de camadas e neurônios ocultos. A redução das variáveis de entrada propicia uma simplificação na arquitetura, reduzindo significativamente o número de parâmetros a serem encontrados, além do tempo destinado para treinamento. Além da minimização do número de variáveis de saída e a simplificação da arquitetura da rede, é importante que diversos tipos de arquitetura sejam testados ao variar pelo menos, o número de neurônios na camada oculta. Esta otimização é necessária para que se consiga afirmar se houve ou não o melhor ajuste de dados no intervalo estudado.

Levando-se em conta os problemas acima citados, foram tomadas as devidas precauções para se evitar o overfitting, com a quantidade de dados para treino e teste criteriosamente escolhidos através de estudos estatísticos e simplificação da arquitetura de rede. A redução das variáveis de entrada e a variação nas arquiteturas foram aplicadas neste trabalho. A presença de “outliers” foi descartada devido ao bom ajuste dos dados. A qualidade dos dados obtidos também foi verificada por testes de repetitividade e reprodutividade nos equipamentos envolvidos, que no geral estão próximos de 1% de incerteza de medição.

Optou-se por uma rede com camada oculta única, para que a arquitetura seja simples, com menor número de parâmetros de rede a serem ajustados. As variáveis de saída, escolhidas de acordo com a disponibilidade de equipamentos em laboratório, são a densidade relativa a 20°C (medida em densímetro automático), a pressão de vapor Reid (medida em aparelho de pressão de vapor) e a curva de destilação (ponto inicial, T_{10} , T_{50} , T_{90} e ponto final de destilação, medidas em destilador automático). Para cada variável de saída, uma arquitetura de rede neuronal é otimizada. Um programa computacional desenvolvido em MATLAB com neural networks toolbox, foi gerado para se obter os conjuntos de parâmetros para cada rede treinada. Este programa proporciona facilidade na escolha da melhor arquitetura, que envolve a variação de parâmetros da rede, tais como o fator de amortecimento variável ou fixo, número máximo de iterações, intervalo de normalização, número de neurônios na camada oculta e funções de transferência. O programa utiliza os três tipos de função de transferência, sigmoideal (logsig), tangencial (tansig) e linear (purelin), combinando duas a duas (camada oculta e de saída) para a otimização da arquitetura. A faixa de parâmetros de rede explorados para as propriedades estudadas é a seguinte:

Número de amostras para treino: 25.

Número de amostras para teste: 10.

Número máximo de iterações: 20.000.

Função de transferência (camada oculta): logsig, tansig e purelin.

Função de transferência (camada de saída): logsig, tansig e purelin.

Faixa de normalização (entrada e saída): 0,2 a 0,8 e 0,3 a 0,7

Fator de amortecimento (η) fixo: 0,1 a 0,9 com incrementos de 0.02.

Fator de amortecimento (η) variável: 0,1 a 0,9 com multiplicador incremental e decremental de 5%. Valores acima de 0,9 desestabilizaram o processo iterativo, com a ocorrência de oscilações nos erros de treinamento.

Número mínimo e máximo de neurônios da camada oculta: 2 a 20.

O programa gera randomicamente os pesos e os bias para início do processo iterativo. O número total estimado de arquiteturas de redes chega a 2.000 para cada variável quando se utiliza o η fixo, enquanto que para o η variável é testado aproximadamente 180 tipos de arquiteturas. Ambos os tipos de fator de amortecimento chegaram a resultados muito próximos, sendo então de grande vantagem o uso do η variável, já que o tempo destinado ao treinamento é muito menor. O total geral de arquiteturas testadas de rede por retropropagação chegou a 15.400.

3. Estudos Preliminares e Planejamento Experimental

A partir de amostras de gasolina comum coletadas entre janeiro e novembro de 2001 no estado do Paraná, realizou-se uma verificação da composição química, com o objetivo de detectar diferenças entre as amostras de uma população e viabilizar a aplicação de rede neuronal. O equipamento utilizado para esta investigação é um infravermelho portátil, específico para a análise de gasolina, denominado IROX 2000 da Grabner Instruments e fornece sua composição química através da porcentagem volumétrica de etanol, aromáticos, olefinicos e saturados. Levando-se em conta um período de 11 meses e uma alíquota de 5500 amostras de gasolina comum coletadas no estado do Paraná, construiu-se gráficos e verificou-se a variabilidade dos dados de entrada e saída. Concluiu-se que as amostras apresentavam diferenças significativas na composição química e nas propriedades permitindo a aplicação em redes neuronais.

Após a detecção de variabilidade nos dados, iniciou-se a escolha das amostras para a rede neuronal. A informação de composição química fornecida pelo IROX 2000 das amostras dos meses de maio e junho do ano de 2002 foi aproveitada para uma prévia tentativa de correlação. Os dados de saída obtidos através dos ensaios rotineiros do

laboratório de ensaios automotivos da Universidade Federal do Paraná também foram aproveitados. Ao realizar a seleção aleatória das amostras, o conjunto de dados iniciais apresentou uma redundância de dados, forçando a previsão das propriedades à média estatística e comprometendo a capacidade preditiva da rede. Portanto, a estratégia de seleção das amostras foi mudada para evitar a repetição de dados. A nova seleção de amostras seguiu os seguintes critérios:

- dividiram-se os dados por faixas de densidade;
- dentro de cada faixa de densidade foram selecionadas aleatoriamente algumas amostras, totalizando 180 amostras de gasolina (amostragem estratificada);
- dentro das 180 selecionadas, reduziu-se para o mínimo possível de amostras, eliminando-se redundância nos dados de saída e composição química;
- chegou-se ao número ideal de 35 amostras representativas da população inicial que era de 1000 amostras que foram então separadas para análise cromatográfica e os ensaios ASTM.

4. Análise Cromatográfica

O cromatógrafo utilizado para a análise detalhada da composição química da gasolina é um CG-FID CP-3800 da Varian, automático e com coluna capilar, que permite fixar condições de trabalho com alta resolução e repetitividade. A análise por DHA (Detailed Hydrocarbon Analysis) permite o reconhecimento das substâncias pertencentes à gasolina através de um procedimento de calibração interna. Esta técnica de identificação fornece diretamente a proporção de cada composto presente na gasolina. As condições cromatográficas que fornecem a melhor resolução para este tipo de análise são:

- gás de arraste: hélio, velocidade linear = 30.0 a 30.2 cm/s a 35°C;
- detector: FID, temperatura: 300°C, sensibilidade: 32×10^{-12} , alimentação: hidrogênio, ar e nitrogênio;
- injetor: 250°C;
- razão de split: 200:1, fluxo: 285-415 cc/min;
- coluna: 100 m x 0.25 mm de sílica fundida com 0.5 micron de metil silicone, especificada como CP-SIL PONA CB FS 100X.25 (0.5) da Chrompack Capillary Column;
- programa de temperatura: 35°C por 15 minutos, rampa de 1 grau/min até 60°C, patamar de 20 minutos, rampa de 2 graus/min até 200°C e patamar de 10 minutos.

5. Resultados e Discussão

5.1. Previsão da Densidade a 20°C

Esta propriedade está relacionada com características de estocagem da gasolina e também pode indicar presença de contaminantes. Os dados de entrada utilizados para a aplicação com a rede neuronal é a composição química com 66 variáveis de entrada. As 66 variáveis de entrada utilizadas são:

- Aromáticos: AC6(aromáticos com 6 cabonos), AC7, AC8, AC9, AC10, AC11, AC12, AC13, AC14 e AC15.
- C14+: C14 e C15.
- Iso parafinas: IC5, IC6, IC7, IC8, IC9, IC10, IC11 e IC12.
- Naftênicos: NC5, NC6, NC7, NC8, NC9, NC10 e NC12. Nota-se que não existe o NC11. O programa DHA não identificou em nenhuma das amostras este subgrupo.
- Olefinas: OLC4, OLC5, OLC6, OLC7, OLC8, OLC9, OLC10, LC11, OLC12, OLC13, OLC14, OLC15 e OLC16.
- Oxigenados: OXC2, OXC3 e OXC4.
- Parafinas: PC3, PC4, PC5, PC6, PC7, PC8, PC9, PC10, PC11, PC12, PC13, PC14, PC15 e PC16.
- Desconhecidos: DC5, DC6, DC7, DC8, DC9, DC10, DC11, DC12 e DC13.

Após encontrar a correlação entre as 66 variáveis de entrada e a densidade, decidiu-se minimizar o número de variáveis de entrada utilizando-se uma composição química mais geral da gasolina, por agrupamentos de substâncias com mesma função química (análise PIANO).

As 7 variáveis de entrada utilizadas são (em % volumétrica):

- | | |
|-----------------|-----------------|
| • Aromáticos | • Oxigenados |
| • Iso-parafinas | • Parafinas |
| • Naftênicos | • Desconhecidos |
| • Olefinas | |

O ajuste de dados foi realizado com sucesso, sendo que os desvios existentes entre a densidade experimental e a calculada pela rede neuronal não ultrapassaram 1%. Deve-se salientar que os desvios obtidos pelas duas redes estão muito próximos. Com isto é possível prever a densidade da gasolina com pouca informação de entrada.

As figuras 1 e 2 mostram a densidade experimental e as densidades calculadas por cada rede para treinamento e validação respectivamente. Os desvios absolutos (diferença percentual entre o ponto experimental e o calculado pela rede) do conjunto de treinamento e validação estão abaixo de 0,50%.

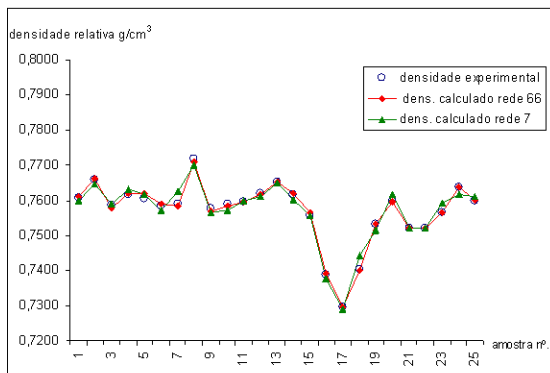


Figura 1 – densidade experimental x calculada para conjunto de treinamento.

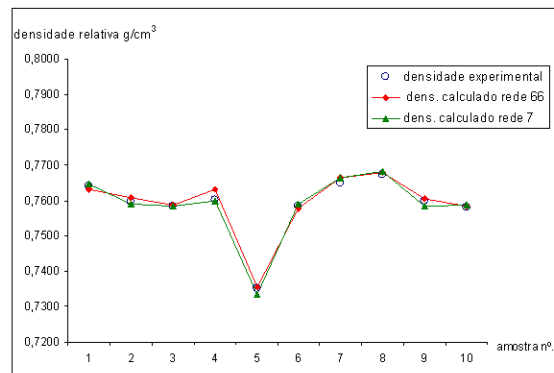


Figura 2 – densidade experimental x calculada para conjunto de validação.

5.2. Previsão da Pressão de Vapor Reid

Esta propriedade indica a volatilidade da gasolina. Pode indicar contaminação por leves e pesados. Uma tentativa de ajuste inicial com as mesmas 66 variáveis de entrada utilizadas na previsão da densidade não foi bem sucedida pois as mesmas representam toda a faixa de composição da gasolina, ao passo que a pressão de vapor está diretamente ligada à quantidade de compostos leves. Com isto, um refinamento dos dados foi necessário. Das 66 variáveis iniciais foram escolhidas 32 que poderiam representar os compostos leves da gasolina.

- Aromáticos: AC6, AC7, AC8 e AC9.
- Iso parafinas: IC5, IC6, IC7, e IC8.
- Naftênicos: NC5, NC6 e NC7.
- Olefinas: OLC4, OLC5, OLC6, OLC7, OLC8 e OLC9.
- Oxigenados: OXC2, OXC3 e OXC4.
- Parafinas: PC3, PC4, PC5, PC6, PC7, PC8 e PC9.
- Desconhecidos: DC5, DC6, DC7, DC8 e DC9.

A redução das variáveis de entrada de 32 para 7 também foi aplicada. Por falta de espaço, os resultados estão apresentados na tabela 1. Nota-se que para a pressão de vapor Reid, existe uma diferença entre a média dos desvios e os erros inseridos pelos equipamentos, que pode ser explicado pela inserção de alguns erros aleatórios nos dados fornecidos para a redes neurais. As possíveis fontes de erros aleatórios são a evaporação de voláteis no manuseio mais demorado da amostra e lavagem insuficiente da célula de medição ao trocar de uma amostra para outra muito diferente.

5.3. Previsão da Curva de Destilação

A curva de destilação está relacionada com a proporção de compostos leves e pesados da gasolina. As informações mais importantes são obtidas através do ponto inicial de ebulição (PIE) e T_{10} (relacionados com a proporção de leves), T_{50} (relacionada com os compostos médios), T_{90} e ponto final de ebulição (PFE) que estão relacionados com a proporção de compostos pesados existentes na gasolina. As variáveis de entrada para o T_{10} e T_{50} foram exatamente as mesmas 66 utilizadas para a previsão da densidade. Um refinamento dos dados de entrada para o PIE, T_{90} e PFE foi necessário, já que o ajuste não ocorreu ao utilizar as 66 variáveis iniciais. Com isto, para o PIE foram utilizadas as mesmas 32 entradas da pressão de vapor, enquanto que para o T_{90} e PFE reduziu-se de 66 para 35 e 24 variáveis respectivamente:

As variáveis selecionadas são (35) para T_{90} :

- Aromáticos: AC10, AC11, AC12, AC13, AC14 e AC15.
- C14+: C14 e C15.
- Iso parafinas: IC9, IC10, IC11 e IC12.
- Naftênicos: NC8, NC9, NC10 e NC12.
- Olefinas: OLC10, LC11, OLC12, OLC13, OLC14, OLC15 e OLC16.
- Oxigenados: OXC2.
- Parafinas: PC10, PC11, PC12, PC13, PC14, PC15 e PC16.
- Desconhecidos: DC10, DC11, DC12 e DC13.

(24) para PFE:

- Aromáticos: AC12, AC13, AC14 e AC15.
- C14+: C14 e C15.
- Iso parafinas: IC10, IC11 e IC12.
- Naftênicos: NC10 e NC12.
- Olefinas: OLC12, OLC13, OLC14, OLC15 e OLC16.
- Parafinas: PC12, PC13, PC14, PC15 e PC16.
- Desconhecidos: DC11, DC12 e DC13.

A redução para as 7 variáveis principais também foi aplicada para todas as temperaturas da curva de destilação. Os resultados estão resumidos na tabela 1. Após o término de todos os treinamentos e testes, observou-se que os ajustes mais difíceis de serem efetuados foram para o ponto final de ebulição. Isto é explicado devido à existência de alterações características do final da destilação, no qual pode ocorrer pirólise ou craqueamento de uma molécula, produzindo moléculas menores com ponto de ebulição mais baixos que a molécula original.

Tabela 1 – Resumo da capacidade preditiva das redes obtidas.

PROPRIEDADE	VAR. DE ENTRADA	DESVIO MÁXIMO - TESTE (%)	DESVIO MÉDIO - TESTE (%)	REPETITIVIDADE DO EQUIPAMENTO (%)
Densidade a 20 °C	66	0,33	0,13	0,015
	7	0,23	0,11	
PV Reid	32	3,11	1,67	0,11
	7	4,94	2,62	
PIE	32	3,53	1,43	1,26
	7	4,01	1,33	
T10	66	1,71	0,97	0,55
	7	1,48	0,86	
T50	66	0,77	0,43	0,21
	7	0,85	0,32	
T90	35	3,92	1,72	0,32
	7	6,37	1,81	
PFE	24	3,75	2,16	1,07
	7	4,41	1,70	

6. Conclusão

De uma forma geral, as redes neuronais obtidas podem ser utilizadas para a previsão de propriedades da gasolina, tais como a densidade relativa a 20°C, pressão de vapor Reid e curva de destilação. A partir de sua composição química obtida por cromatografia gasosa com detector de ionização de chama usada como entrada de variáveis, cada rede neuronal pode prever sua respectiva propriedade, dentro da precisão experimental. As correlações desenvolvidas através da aplicação de redes neuronais artificiais por retropropagação provaram ser efetivas em correlacionar a composição química da gasolina às suas propriedades físico-químicas. A redução de dados para sete variáveis de entrada gerou redes de arquiteturas mais simples, possuindo capacidade preditiva igual ou superior às arquiteturas mais complexas. Um aprofundamento do estudo para inserção de pontos fora do intervalo estudado pode ser efetuado, com o objetivo de adequar as redes obtidas para utilização. O trabalho desenvolvido pode ser estendido para previsão de MON e RON, reduzindo custos laboratoriais.

7. Agradecimentos

Agradecemos à ANP (Agência Nacional do Petróleo) e ao CTPETRO/FINEP pelo suporte financeiro.

8. Referências

- ANDRADE, J. M.; SÁNCHEZ, M.S.; SARABIA, L. A.. Applicability of high-absorbance MIR spectroscopy in industrial quality control of reformed gasolines. *Chemometrics and Intelligent Laboratory Systems*, v. 46, p. 41-55, 1999.
- ELSHARKWY, A. M.; GHARBI, R. B. C.. Comparing classical and neural regression techniques in modeling crude oil viscosity. *Advances in Engineering Software*, v. 32, p. 215 – 224, 2001.
- MEUSINGER, R.; MOROS, R.. Determination of octane numbers of gasoline compounds from their chemical structure by ¹³C NMR spectroscopy and neural networks. *Fuel*, v. 80, p. 613 - 621, 2001.
- VAN LEEUWEN, J. A., JONKER, R. J., GILL, R.. Octane number prediction based on gas chromatographic analysis with non-linear regression techniques. *Chemometrics and Intelligent Laboratory Systems*, v. 25, p. 325 – 390, 1994.
- YANG, H. et al.. Neural network prediction of cetane number and density of diesel fuel from its chemical composition determined by LC and GC-MS. *Fuel*, p. 1-10, 2001.