

**UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIAS
CURSO DE LICENCIATURA EM QUÍMICA**

CARLOS ANDRÉ GOMES BEZERRA

**DETERMINAÇÃO DA CURVA DE DESTILAÇÃO DA GASOLINA UTILIZANDO
FERRAMENTAS ESTATÍSTICAS E COMPUTACIONAIS**

SÃO LUÍS - MA

2016

CARLOS ANDRÉ GOMES BEZERRA

**DETERMINAÇÃO DA CURVA DE DESTILAÇÃO DA GASOLINA UTILIZANDO
FERRAMENTAS ESTATÍSTICAS E COMPUTACIONAIS**

Monografia apresentada ao curso de Química -
Licenciatura da Universidade Federal do
Maranhão, como requisito final para obtenção do
grau de licenciado em Química.

Orientadora: Prof^ª. Dra. Aldaléa Lopes Brandes Marques

Co-orientador: Prof^º. Dr. Alex de Oliveira Barradas Filho

SÃO LUÍS - MA

2016

Gomes Bezerra, Carlos André.

DETERMINAÇÃO DA CURVA DE DESTILAÇÃO DA GASOLINA
UTILIZANDO FERRAMENTAS ESTATÍSTICAS E COMPUTACIONAIS /
Carlos André Gomes Bezerra. - 2016.

72 p.

Coorientador(a): Alex de Oliveira Barradas Filho.

Orientador(a): Aldaléa Lopes Brandes Marques.

Monografia (Graduação) - Curso de Química, Universidade
Federal do Maranhão, São Luís - MA, 2016.

1. Curva de destilação. 2. Gasolina. 3. PLS e MLR.
4. Predição. 5. Redes Neurais Artificiais. I. Barradas
Filho, Alex de Oliveira. II. Brandes Marques, Aldaléa
Lopes. III. Título.

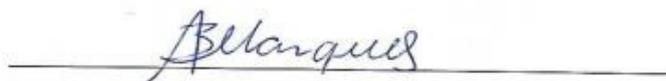
CARLOS ANDRÉ GOMES BEZERRA

**DETERMINAÇÃO DA CURVA DE DESTILAÇÃO DA GASOLINA UTILIZANDO
FERRAMENTAS ESTATÍSTICAS E COMPUTACIONAIS**

Monografia apresentada ao curso de Química -
Licenciatura da Universidade Federal do
Maranhão, como requisito final para obtenção do
grau de licenciado em Química.

Aprovada em 18/08/2016

BANCA EXAMINADORA



Prof. Dra. Aldaléa Lopes Brandes Marques (Orientadora)
(UFMA)



Prof. Dra. Jaciene Jesus Cardoso Rocha
(UFMA)



Prof. M^c. Elmo de Sena Ferreira Júnior
(UFMA)

Dedico essa vitória a Deus, por suas infinitas bênçãos.

Com muito carinho à minha avó Júlia Bezerra, que mesmo lá do céu, sei que torce por mim.

De modo especial, à minha família, com muito amor, à minha mãe, Antonia Gomes, e a meu pai, Raimundo Bezerra, meus dois maiores e melhores professores.

E à Larissa Sousa Pereira, uma pessoa muito especial em minha vida.

AGRADECIMENTOS

O dom da vida é sem dúvidas a mais preciosa virtude que temos, e, portanto, nada mais justo do que demonstrar gratidão àquele que nos concede essa bênção. Por isso, agradeço a **Deus** por tudo o que tem feito em minha vida; por colocar ao meu lado pessoas maravilhosas, que tornam meus dias momentos de pura felicidade.

A minha família, meu bem mais precioso, ao qual atribuo todas as conquistas de minha vida. De modo especial aos meus pais **Antonia Gomes** e **Raimundo Bezerra**, que sempre me apoiaram e incentivaram a nunca desistir de meus planos. Que sempre me ensinaram a trilhar os caminhos da vida sem nunca temer qualquer desafio. Tudo aquilo que tenho e que sou, devo totalmente a vocês.

Aos meus irmãos **Adriano, Andréia, Agostinho, Antonio Francisco e Maria**, que desde muito cedo, tem me ajudado a trilhar os caminhos da vida, e me incentivado a não me desviar dos ensinamentos de nossos pais. Vocês são pessoas muito importantes, às quais desejo muita felicidade e sucesso.

A minha avó **Júlia Bezerra** (*in memoriam*) por ter me ensinado, mesmo sem saber, a ver a vida de um jeito diferente, sem deixar que os desafios se tornem maior que nossas forças.

A minha namorada, amiga e fiel companheira, **Larissa Sousa**, que sempre esteve ao meu lado durante essa caminhada. Agradeço pelo amor, incentivo e ensinamentos que, sem dúvidas, me capacitaram cada vez mais para que eu viesse concretizar mais esse sonho em minha vida. Sou eternamente grato por tudo que vivemos juntos.

A minha extraordinária orientadora **Prof^a. Dra. Aldaléa Lopes Brandes Marques**, por seus ensinamentos e experiências a mim direcionados; pela confiança depositada em mim durante todos esses anos em que tive a honra de participar da família LPQA e LAPQAP; pelas valiosas e construtivas críticas que me fizeram crescer pessoalmente e profissionalmente; por dedicar parte do seu precioso tempo, paciência e conhecimentos que sempre clarearam meus caminhos durante as pesquisas das quais tive a honra de ser seu aluno. Tenho a senhora como referência, seja no campo moral ou profissional, pois sua conduta e sua postura como professora e pesquisadora sempre mostraram a grande pessoa, e excelente profissional que és. Sou

eternamente grato pela oportunidade de ter sido seu aluno de iniciação científica durante minha graduação. Por tudo isso, meu imenso e eterno obrigado.

Ao **Profº. Dr. Alex Barradas** que com sua singular postura de pesquisador e também como pessoa, muito me ensinou durante minha graduação; com certeza suas críticas e conselhos possibilitaram que eu tivesse uma nova visão enquanto profissional e também como pessoa. Por isso, sou muito grato.

Ao **Profº. Dr. Edmar Marques**, pelos ensinamentos e orientações que muito contribuíram na minha formação e aprendizado.

A **Isabelle Viegas**, pelos conhecimentos e críticas que fizeram muita diferença durante a minha formação, possibilitando que eu também pudesse ter uma visão mais crítica sobre os objetos de minhas pesquisas.

Ao **Profº. Dr. João Coelho e família**, por todo o apoio e carinho durante essa etapa de minha vida. Sou muito grato por tudo que fizeram por mim.

A minha turma QM/QB 2012.1, que durante todos esses anos foram como uma família para mim. Em especial a **Mayara Santos, Chirlene Botelho, Liane Miranda, Gierllan Wesley, Adryane Maciel, Vangeles Pires, Lis Conde e Rodrigo Barros**, grandes amigos com os quais tive a honra de conviver. Desejo a todos muito sucesso e felicidade.

A toda equipe do **LAPQAP e LPQA** que desde o início me acolheram e me fizeram crescer dentro dessa instituição, tanto no campo profissional como também social.

Aos honoráveis **Aronilson Aguiar e Luís Cláudio**, grandes amigos que muito me ensinaram ao longo desse período. Agradeço-lhes pelo carinho e dedicação. Desejo a vocês muito sucesso em suas vidas.

A **Residência Estudantil da UFMA (REUFMA)** pelo acolhimento durante esses 4 anos e meio de estudos, onde aprendi importantes lições que levarei para toda a vida. Local que me deu suporte para iniciar e finalizar o meu curso de graduação.

Ao **NEPE, LPQA e LAPQAP** por me darem todo o suporte para que eu desenvolvesse minhas pesquisas e pudesse ascender profissionalmente.

Ao **CNPq** pelo financiamento dos nossos projetos e pelas bolsas concedidas.

A **UFMA**, por ter sido essa estrutura que me acolheu e que permitiu a realização desse sonho, que não é só meu e que por isso tem um valor muito grande pra mim e para toda minha família.

Meu muito obrigado!

“Sejais como os archeiros habilidosos, os quais, parecendo-lhes muito distante o alvo que desejam atingir, e, conhecendo bem as limitações de alcance do seu arco, fazem pontaria para um ponto muito mais alto do que o alvo, não para aí acertar com as suas flechas, mas, sim, para com a ajuda de tão elevada mira, poderem bater o alvo pretendido.”

Nicolau Maquiavel

RESUMO

A gasolina é um dos combustíveis petroquímicos mais consumidos em território nacional, ficando atrás apenas do diesel em termos de volume. Para garantir a qualidade final da gasolina distribuída nos diversos postos de abastecimento, a Agência Nacional do Petróleo Gás Natural e Biocombustível estabelece por meio de uma legislação específica, uma série de testes laboratoriais que, cumulativamente, determinarão se o combustível está apto a ser comercializado pelos postos de abastecimento. No entanto, a realização de alguns dos testes de controle de qualidade demanda uma grande quantidade de tempo, além de ser uma etapa onerosa para laboratório em termos de infraestrutura e consumo de energia. Neste trabalho foi desenvolvido metodologias alternativas para a determinação da curva de destilação da gasolina C por meio de ferramentas de predição de resultados, tendo como parâmetros de referência a massa específica e a pressão de vapor da gasolina. Para tanto, um banco de dados foi obtido da literatura com resultados de massa específica, de pressão de vapor e da curva de destilação de 105 amostras de gasolina. Para o tratamento desse banco de dados foram aplicadas diferentes técnicas quimiométricas e computacionais. As técnicas utilizadas foram a regressão linear múltipla, a regressão por mínimos quadrados parciais e a técnica de redes neurais artificiais. Os pontos da curva de destilação foram tratados separadamente, obtendo-se um modelo preditivo para cada ponto da curva. Foram obtidos bons resultados de predição para os dois primeiros pontos da curva de destilação, sendo que o método de predição por redes neurais artificiais foi o que apresentou os melhores resultados. Já para os dois últimos pontos da curva de destilação, os resultados preditos apresentaram desvios de predição consideráveis, sendo que os melhores resultados também foram obtidos pela aplicação da técnica de redes neurais artificiais. Sendo assim, dentre as técnicas preditivas estudadas, o método por redes neurais artificiais foi o que teve o melhor desempenho de predição. Além disso, verificou-se que as técnicas preditivas por regressão linear múltipla e regressão por mínimos quadrados parciais apresentaram resultados preditivos com grande similaridade, o que foi atribuído ao uso de apenas duas variáveis de entrada para a predição da curva de destilação da gasolina.

Palavras chave: Gasolina. Predição. Curva de destilação. Redes Neurais Artificiais. PLS. MLR.

ABSTRACT

The gasoline is one of petrochemical fuel most consumed in the country, staying behind only of diesel in volume's terms. For to ensure the final quality of distributed gasoline in the various gas stations, the National Agency of Petroleum Natural Gas and Biofuel establishes of through specific legislation, a series of laboratory tests which cumulatively determine if the fuel is apposite for to be commercialized by gas stations. However, the realization of some tests of quality control requires a great deal of time, in addition to be an expensive step to the laboratory in terms of infrastructure and energy consumption. This work developed alternative methods for the determination of C gasoline distillation curve by means of results prediction tools, with reference to density parameters and vapor pressure of the gasoline. Thus, the database were obtained from the literature with results of specific mass, steam pressure and distillation curve of one hundred and five samples. For the treatment of this database were applied different chemometric and computational techniques. The techniques used were multiple linear regression, regression by partial least squares and the technique of artificial neural networks. The points of the distillation curve were treated separately, obtaining a predictive model for each point of the curve. Good results of prediction were obtained for the first two points of the distillation curve, being that the prediction artificial neural networks method showed the best results. For the last two points of the distillation curve, the predicted results showed considerable prediction deviations, being that the best results have been obtained by application the technique of artificial neural networks. Thus, among the predictive techniques studied, the method of artificial neural networks had the best prediction performance. Furthermore, it was found that the predictive techniques by multiple linear regression and regression by partial least squares showed the same predictive results, which was attributed to the use of only two input variables to predict the distillation curve of gasoline.

Keywords: Gasoline; Prediction; Distillation Curve; Artificial Neural Networks; PLS; MLR

LISTA DE ILUSTRAÇÕES

Figura 1 - Consumo de energia pelo setor de transportes, balanço 2013-2014.....	21
Figura 2 - Consumo de gasolina pelo setor de transportes brasileiro.....	22
Figura 3 - Curva de destilação ilustrativa para uma gasolina comercial (HUGO, 2005).....	26
Figura 4 - Diagrama de representação da arquitetura das ANNs treinadas para a predição da curva de destilação da gasolina.	38
Figura 5 - Gráficos preditivos obtidos para os modelos de predição no ponto T10 da curva de destilação.....	43
Figura 6 - Gráficos de resíduos obtidos para os modelos de predição no ponto T10 da curva de destilação.....	44
Figura 7 - Gráficos de predição obtidos para os modelos de predição no ponto T50 da curva de destilação.....	47
Figura 8 - Gráficos de resíduos obtidos para os modelos de predição no ponto T50 da curva de destilação.....	48
Figura 9 - Gráficos de predição obtidos para os modelos de predição no ponto T90 da curva de destilação.....	51
Figura 10 - Gráficos de desvios obtidos para os modelos de predição no ponto T90 da curva de destilação.....	52
Figura 11 - Gráficos de predição obtidos para os modelos de predição no ponto final da curva de destilação.	55
Figura 12 - Gráficos de resíduos obtidos para os modelos de predição para o ponto final da curva de destilação.	57
Figura 13 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para o ponto T10.....	58
Figura 14 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para o ponto T50.....	59
Figura 15 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para o ponto T90.....	60
Figura 16 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para as temperaturas do PFE.	61
Figura 17 - Comportamento preditivo dos modelos de predição em cada ponto da curva de destilação.....	62

LISTA DE TABELAS

Tabela 1 - Parâmetros de qualidade da gasolina exigidos pela ANP segundo a RESOLUÇÃO ANP N° 40, DE 25.10.2013.	23
Tabela 2 - Resultados exploratórios para o banco de dados utilizado no presente trabalho. ..	34
Tabela 3 - Resultados de erros experimentais para os equipamentos usados para a obtenção dos resultados experimentais das propriedades fisico-químicas estudadas.	35
Tabela 4 - Parâmetros de configuração para os modelos de PLS obtidos para a predição da curva de destilação da gasolina.	37
Tabela 5 - Parâmetros de ajuste na construção das ANNs e suas especificações.	38
Tabela 6 - Otimização do banco de dados pelo algoritmo KSXY para a aplicação nos modelos de predição em cada ponto da curva de destilação.	40
Tabela 7 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto T10 da curva de destilação.	41
Tabela 8 - Características estruturais para a melhor ANN treinada para a T10.	42
Tabela 9 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto T50 da curva de destilação.	45
Tabela 10 - Características estruturais para a melhor ANN treinada para a T50.	47
Tabela 11 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto T90 da curva de destilação.	49
Tabela 12 - Características estruturais para a melhor ANN treinada para a T90.	50
Tabela 13 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto final da curva de destilação.	53
Tabela 14 - Características estruturais para a melhor RNA treinada para o PFE.	54
Tabela 15 - Propriedades fisico-químicas das amostras de gasolina usadas na construção dos modelos de predição.	69

LISTA DE ABREVIATURAS E SIGLAS

°C	Graus Celsius
%	Por cento
kg	Quilograma
KPa	Quilo pascal
m³	Metro cúbico
ABNT	Associação Brasileira de Normas Técnicas
AEAC	Álcool Etílico Anidro Combustível
ANN	Redes Neurais Artificiais
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ASTM	<i>American Society for Testing and Materials</i>
GC	Gasolina Comum
LM	Algoritmo de Levenberg–Marquardt
Logsig	Função Logística
MAPE	Mean Absolut Percentage Error
mL	Mililitro
MLP	<i>Multilayer perceptron</i>
MLR	<i>Multiple Linear Regression</i>
NBR	Norma da Associação Brasileira de Normas Técnicas
PFE	Ponto Final de Ebulição
PLS	<i>Partial Least Square</i>
PMQC	Programa de Monitoramento da Qualidade de Combustíveis
PV	Pressão de Vapor
r	Coefficiente de correlação
RMSE	<i>Root Mean Square Error</i>
T10	Temperatura de ebulição de dez por cento da fração destilada
T50	Temperatura de ebulição de cinquenta por cento da fração destilada
T90	Temperatura de ebulição de noventa por cento da fração destilada
Tansig	Função Tangente hiperbólica
UFPR	Universidade Federal do Paraná

SUMÁRIO

1. INTRODUÇÃO	17
1.1. Aspectos da gasolina	19
1.1.1. A Gasolina Brasileira	19
1.1.2. Controle de qualidade e propriedades fisico-químicas da gasolina	23
1.2. Técnicas de predição de resultados	27
1.2.1. Regressão por mínimos quadrados parciais	28
1.2.2. Regressão linear múltipla	29
1.2.3. Redes neurais artificiais	29
1.3. Estado da arte	30
2. OBJETIVOS	32
2.1. Objetivo geral	32
2.2. Objetivos Específicos	32
3. METODOLOGIA	33
3.1. Obtenção do banco de dados	33
3.2. Construção dos modelos de predição	35
3.3. Configurações e parâmetros da técnica de predição por MLR	36
3.4. Configurações e parâmetros da técnica de predição por PLS	37
3.5. Configurações e parâmetros da técnica de predição por ANN	37
3.6. Avaliação dos modelos de predição	39
4. RESULTADOS E DISCUSSÕES	40
4.1. Otimização do banco de dados	40
4.2. Predição do ponto T10 da curva de destilação	41
4.3. Predição do ponto T50 da curva de destilação	45
4.4. Predição do ponto T90 da curva de destilação	49
4.5. Predição do ponto PFE da curva de destilação	53
4.6. Comparativo entre os métodos de predição utilizados	57
4.6.1. Modelos preditivos obtidos para o ponto T10	58
4.6.2. Modelos preditivos obtidos para o ponto T50	59
4.6.3. Modelos preditivos obtidos para o ponto T90	60
4.6.4. Modelos preditivos obtidos para a temperatura do PFE	61
4.6.5. Desempenho geral dos modelos de predição	62

5. CONCLUSÃO	63
6. ETAPAS FUTURAS	64
REFERÊNCIAS	65
ANEXO A	69

1. INTRODUÇÃO

Nas últimas décadas, a demanda energética por apresentar um cenário de constante crescimento tem sido fortemente debatida em diversos setores como universidades, instituições públicas, empresas e indústrias. No Brasil, o estudo intitulado de Demanda de Energia 2050 realizado pela Empresa de Pesquisa Energética (EPE) revela que o consumo total do país, incluindo eletricidade, gasolina e etanol, entre outros, duplicará no período de 2014 até 2050 (EPE, 2014). No ano de 2014, por exemplo, houve um aumento de 11% na produção de petróleo no Brasil em comparação com o ano de 2013 (BRASIL, 2015).

No âmbito do setor automotivo nacional, a gasolina é o segundo combustível mais consumido, ficando atrás apenas do diesel, que é utilizado principalmente na frota dedicada ao transporte de cargas (BRASIL, 2015). No entanto, para que a gasolina seja comercializada é necessário um controle rigoroso de parâmetros de qualidade relacionados ao desempenho e segurança do combustível. Esse monitoramento é realizado com base em uma legislação específica, que estabelece uma série de parâmetros físico-químicos necessários para que o combustível possa ser disponibilizado ao consumidor final (ANP, 2013).

A regulamentação dos parâmetros de qualidade de combustíveis no Brasil é elaborada pela Agência Nacional do Petróleo, Gás Natural e Biocombustível (ANP). No caso da gasolina, a ANP estabelece por meio da resolução nº 40 uma série de testes físico-químicos, em que os valores dos limites máximos e mínimos, as metodologias e as condições para cada propriedade físico-química da gasolina são especificados por normas técnicas da Associação Brasileira de Normas Técnicas (ABNT) e da American Society for Testing and Materials (ASTM) (ANP, 2013).

Dentre os testes exigidos pela ANP, alguns dos principais são a Massa Específica a 20°C, a Pressão de Vapor REID (PV) e a curva de destilação da gasolina expressa por meio dos pontos nos quais 10, 50 e 90% da gasolina destilada é coletada (T10, T50 e T90, respectivamente), e do Ponto Final de Ebulição (PFE).

A curva de destilação é um parâmetro fundamental na caracterização da qualidade da gasolina, pois fornece informações acerca do desempenho do combustível no motor (HUGO, 2005) e pode servir como indicativo para a presença de substâncias adulterantes no combustível (CÔCCO, 2008).

Com relação ao conjunto de testes de controle de qualidade exigido pela ANP, a Massa específica a 20°C e a Pressão de Vapor, são duas propriedades que podem estar relacionados com a curva de destilação da gasolina. Do ponto de vista físico-químico essa associação é feita

considerando que a massa específica é um indicativo do peso molecular, estado de agregação e tipos de interações intermoleculares (BRUICE, 2006), e a pressão de vapor está relacionada com a temperatura na qual a substância passa para o estado de vapor (MAHAN e MEYERS, 1996).

A verificação dos pontos da curva de destilação é feita por meio de procedimentos laboratoriais de destilação segundo as normas da ABNT NBR 9619 e ASTM D86 (ANP, 2013). Esses procedimentos tem a característica de serem demorados e laboriosos, o que é inconveniente do ponto de vista da rotina do laboratório de análises, principalmente por conta do número de amostras que são analisadas diariamente.

Procedimentos laboriosos como este tem servido como ponto de partida para a elaboração de pesquisas voltadas para o desenvolvimento de métodos quimiométricos que tem como objetivo principal contornar as etapas mais complicadas do processo de controle de qualidade da gasolina, através da utilização de técnicas de predição de resultados (CÔCCO, 2008; CÔCCO, 2003; HUGO, 2005).

No que se refere as técnicas de predição de resultados, dentre as principais temos a regressão linear múltipla (do inglês, *Multiple Linear Regression – MLR*), a regressão por mínimos quadrados parciais (do inglês, *Partial Least Square – PLS*) e a técnica de Redes Neurais Artificiais (do inglês, *Artificial Neural Network - ANN*). Na literatura científica essas técnicas de predição são constantemente usadas na determinação de propriedades físico-químicas de diversos tipos de amostras. (BARRADAS FILHO et al., 2015; BRAGA, POPPI, 2004; BUIATTE, 20016; PANDORFI, 2016)

Em diversas situações a aplicação dessas técnicas preditivas exige o uso de microcomputadores para o processamento do banco de dados e obtenção de resultados, o que também tem sido favorecido com o surgimento de equipamentos com o poder de processamento cada vez maior e com um baixo custo de aquisição e manutenção, se comparado ao custo dos equipamentos convencionais usados na rotina de controle de qualidade da gasolina, por exemplo.

Nesse sentido, considerando os aspectos econômicos e infra estruturais, seria vantajoso o desenvolvimento e implementação de um método de obtenção da curva de destilação da gasolina que se baseasse em resultados experimentais relacionados a essa propriedade físico-química e que são obtidos de forma mais rápida e simples. Em função disso, propôs-se a aplicação dos resultados experimentais das propriedades de pressão de vapor e massa específica a 20°C como dados de entrada na predição da curva de destilação da gasolina utilizando as técnicas de predição por MLR, PLS e ANN.

1.1. Aspectos da gasolina

A gasolina é formada por mais de 400 compostos diferentes, sendo constituída por várias espécies de hidrocarbonetos, incluindo as classes parafínica, olefínica, naftênica, aromática e, em menores quantidades, compostos oxigenados e contendo enxofre (BORSATO et al., 2009). As cadeias carbônicas dos hidrocarbonetos da gasolina têm de 4 a 12 átomos de carbono, e a faixa de ebulição desses compostos varia de 30 a 225°C (SILVA et al., 2009).

De modo geral, a obtenção da gasolina envolve uma série de processos físico-químicos de separação, iniciando com a destilação fracionada do petróleo bruto, de onde é extraída a maior parte do componente básico do combustível. Em seguida são aplicados processos de modificação das cadeias carbônicas que compõem o combustível, num processo chamado de craqueamento catalítico, através do qual a gasolina adquire importantes características, como o alto índice de octanagem por exemplo (PETROBRAS, 2016).

Além disso, em diferentes partes do mundo é adicionado determinado volume de Álcool Étilico Anidro Combustível (AEAC) com a finalidade de melhorar as características de combustão e de desempenho da gasolina. No Brasil, por exemplo, é adicionado 26% de AEAC à gasolina, compondo uma mistura conhecida como gasolina comum ou gasolina tipo C (ANP, 2015). Essa mistura tem como objetivo principal melhorar as características do combustível, e no Brasil também cumpri uma finalidade política e econômica, tendo em vista o desenvolvimento do setor alcooleiro nacional (ANP, 2013; GUARNIERI, 1992).

1.1.1. A Gasolina Brasileira

A gasolina brasileira em sua grande parte é produzida em território nacional pela Petrobras e em alguns casos é importada por empresas licenciadas junto à estatal (PETROBRAS, 2016). Do ponto de vista da composição, a gasolina brasileira diferencia-se em dois tipos principais. A gasolina pura, chamada de gasolina tipo A, que é o combustível que sai das refinarias, oriundo dos processos de refino do petróleo, e que não está pronta para ser usada nos veículos automotores. E a Gasolina Comum (GC) que é revendida nos postos de combustíveis. A GC é obtida pela adição de AEAC à gasolina tipo A. Essa mistura é feita segundo as normas da legislação brasileira estabelecidas pela ANP (PETROBRAS, 2016; ANP, 2013).

A adição de AEAC à gasolina tem como premissas duas características muito importantes. Essas características fundamentais são o índice de octanagem e a quantidade de poluentes liberados na atmosfera durante a queima do combustível (GUEDES et al, 2010).

O índice de octanagem é um parâmetro usado para medir o grau de compressibilidade de um combustível até o ponto da detonação dentro da câmara de combustão. Essa propriedade é um indicativo da performance do combustível no motor (FERREIRA, 2006). O desejável é que o combustível apresente altos índices de octanagem, indicando que a energia térmica gerada pela detonação deste terá altos níveis de conversão em energia de movimento (MORAES, 2003).

Inicialmente era adicionado Chumbo Tetraetila à gasolina como forma de melhorar a octanagem desse combustível. No entanto, com o passar do tempo, as pesquisas mostraram que esse composto liberava altos níveis de chumbo no ar, o que estava associado a diversos distúrbios do sistema nervoso observados em pessoas que ficavam expostas a fumaça proveniente da queima da gasolina misturada a esse composto (GUARNIERI, 1992). Com isso, após uma série de pesquisas feitas no sentido de encontrar um substituto ao Chumbo Tetraetila, constatou-se que o AEAC cumpria os requisitos necessários para substituir essa substância nociva (SCHIFER, 2005). Dessa forma, rapidamente o AEAC se popularizou e passou a ser adicionado à gasolina ao redor do mundo.

O AEAC reflete diretamente sobre o índice de octanagem da GC, fazendo com que esse combustível apresente valores de octanagem bem maiores do que aqueles alcançados pela gasolina pura. Além disso, a adição de AEAC à gasolina diminui a quantidade de gases nocivos que são formados durante a queima do combustível no motor, ajudando, dessa forma, a amenizar o problema da poluição atmosférica enfrentado por diversos centros urbanos aqui no Brasil (MALFATTI, 2009).

Em território nacional, a adição de AEAC à gasolina, além de cumprir os requisitos técnicos de composição e desempenho, também possui um papel político e econômico que visa valorizar a indústria nacional de etanol e favorecer, de forma indireta, o sistema de produção de cana de açúcar (GUARNIERI, 1992; ANDRADE, 2009).

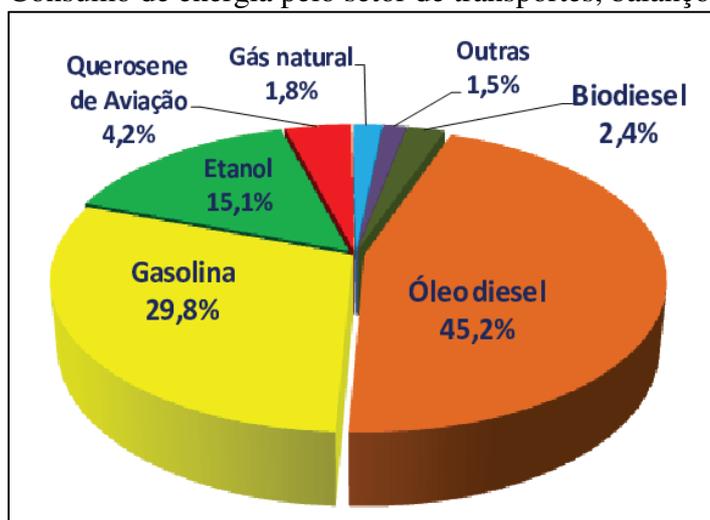
Esse programa de valorização do mercado alcooleiro nacional através da adição de uma quantidade de AEAC acima da percentagem adotada em outros países, teve início na década de 70 com o programa intitulado de PROÁLCOOL. Embora existam argumentos que questionam a validade desse programa energético e político (GUARNIERI, 1992), o PROÁLCOOL possibilitou significativas melhoras no cenário econômico e político nacional na época em que

foi implantado, além de iniciar uma fase de intenso investimento em uma matriz energética de caráter nacional e sustentável (LEME, 2004).

Atualmente, a ANP, por meio da resolução nº 40, estabelece uma série de critérios de qualidade que garantem a qualidade e segurança da gasolina ao consumidor final. Nesse documento, segundo a RESOLUÇÃO ANP Nº 30, DE 29.6.2015 - DOU 30.6.2015, o teor de AEAC a ser adicionado à gasolina deve estar entre 24 a 26%. Esse valor para o teor de AEAC na gasolina tem reflexos direto em parâmetros ambientais tal como a poluição por exemplo, dado o volume do combustível que é consumido pelo mercado interno nacional (ANP, 2013).

Segundo relatório do Ministério de Minas e Energia sobre o balanço energético nacional divulgado em 2015, a gasolina é o segundo combustível mais consumido no Brasil no setor de transportes, ficando atrás apenas do diesel que é usado principalmente na frota de veículos pesados destinados ao transporte de carga. O gráfico da Figura 1 apresenta os dados de consumo de combustíveis pelo setor de transportes no período de 2013 a 2014.

Figura 1 - Consumo de energia pelo setor de transportes, balanço 2013-2014.

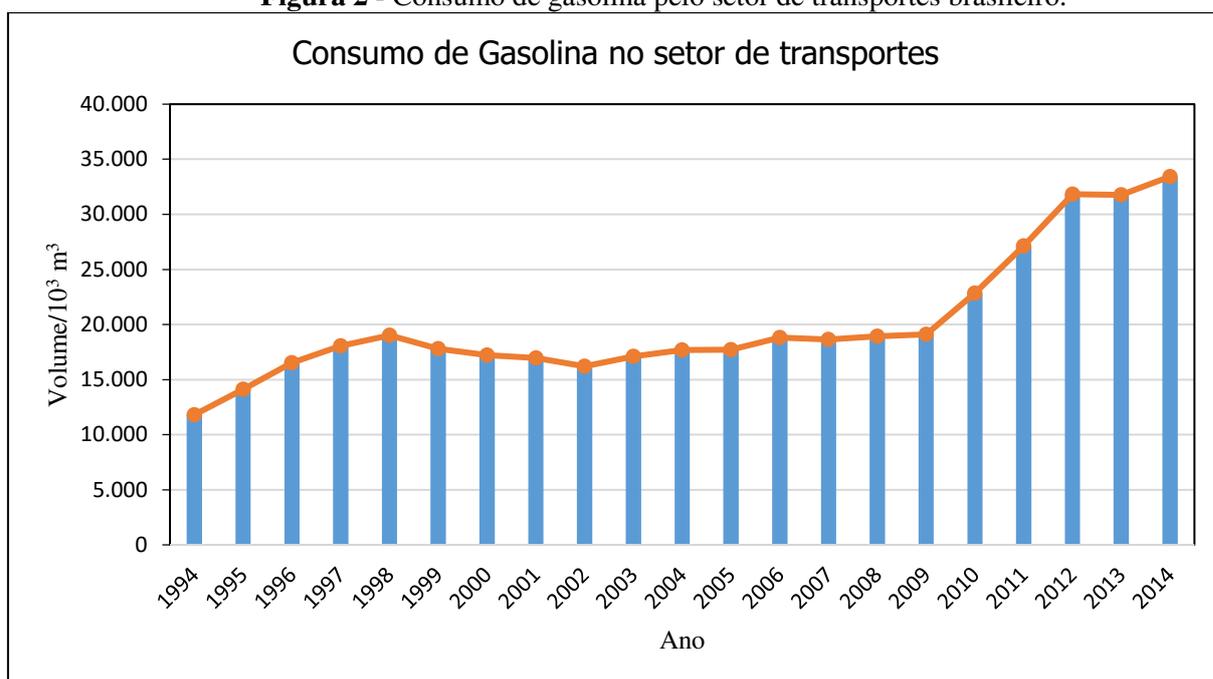


Fonte: <https://ben.epe.gov.br/>.

Ainda segundo dados do Ministério de Minas e Energia, a participação da matriz energética renovável no setor de transportes corresponde a apenas 18% do total de energia consumida por essa classe. Isso significa que a matriz petroquímica ainda continua sendo a principal fonte de energia do setor de transportes no Brasil, sendo que a predominância do diesel como principal combustível usado por esse mercado se deve ao seu uso pela grande frota nacional de veículos pesados. Por outro lado, a gasolina é o combustível mais usado pela frota de veículos leves, dedicados principalmente ao transporte de pessoas (DETRAN, 2016).

Segundo dados divulgados pela ANP em fevereiro de 2015, o consumo de gasolina no Brasil no ano de 2014 foi de 44,364 bilhões de litros, um aumento de 7,09% se comparado com o ano de 2013, que registrou um consumo de 41,428 bilhões de litros do combustível. No mesmo contexto, a frota de automóveis do Brasil cresceu 5,5% no período de dezembro de 2013 a dezembro de 2014 (DETRAN, 2016). O gráfico da Figura 2 apresenta a variação do consumo de gasolina nos últimos 20 anos no setor de transportes.

Figura 2 - Consumo de gasolina pelo setor de transportes brasileiro.



Fonte: BALANÇO ENERGÉTICO NACIONAL, 2015

Como pode ser acompanhado pelo gráfico da Figura 1, o consumo de gasolina no setor de transportes tem sofrido um gradativo aumento, que tem se intensificado nos últimos anos. Esse aumento no consumo de gasolina tem reflexo direto em outros setores que fazem parte da cadeia de produção e consumo da gasolina. Um deles é o setor de controle de qualidade, que também passou por transformações ao longo desse período.

1.1.2. Controle de qualidade e propriedades físico-químicas da gasolina

O controle de qualidade da gasolina é regulamentado pela ANP, que dispõe de uma legislação específica que estabelece limites para uma série de parâmetros físico-químicos da gasolina, os quais devem ser obedecidos pelo setor de distribuição do combustível.

A Tabela 1 apresenta os principais testes de controle de qualidade indicados pela norma ANP nº 40. Essa tabela resume o conjunto de testes físico-químicos aos quais a gasolina brasileira precisa ser submetida para que sua qualidade seja atestada e sua distribuição e revenda autorizadas pela ANP (ANP, 2013).

Tabela 1 - Parâmetros de qualidade da gasolina exigidos pela ANP segundo a RESOLUÇÃO ANP Nº 40, DE 25.10.2013.

CARACTERÍSTICA	UNIDADE	LIMITE				MÉTODO	
		Gasolina Comum		Gasolina Premium		ABNT NBR	ASTM
		Tipo A	Tipo C	Tipo A	Tipo C		
Teor de Etanol Anidro Combustível	% volume	(6)	(7)	(6)	(7)	13992	-D5501 (8)
Massa específica a 20°C	kg/m ³	anotar				7148	D1298
						14065	D4052
Destilação							
10% evaporado, máx.	°C	65				9619	D86
50% evaporado, máx.		120	80	120	80		
90% evaporado, máx.		190					
PFE, máx.		215					
Resíduo, máx.	% volume	2					
Nº de Octano Motor - MON, mín. (9)	-	-	82	-	-	-	D2700
Índice Antidetonante - IAD, mín. (9) (10)	-	-	87	-	91	-	D2699
							D2700

Pressão de Vapor a 37,8°C (11)	kPa	45,0 a 62,0	69,0 (máx.)	45,0 a 62,0	69,0 (máx.)	14149	D4953
						14156	D5191
						-	D5482
							D6378
(6) Proibida a adição. Deve ser medido quando houver dúvida quanto à ocorrência de contaminação. Considera-se o limite máximo de 1 % em volume; (7) O teor de etanol anidro combustível (EAC) a ser misturado à gasolina A para produção da gasolina C deverá estar em conformidade com a legislação vigente.							

Fonte: ANP, 2013

Conforme é possível acompanhar na Tabela 1, existem diversos testes e limites estabelecidos que cumulativamente atestam a qualidade final do produto revendido nos postos de combustíveis. Dentre os parâmetros citados acima, muitos são relativamente de fácil verificação, como é o caso do teor de AEAC, que requer uma instrumentação simples e produz resultados relativamente rápidos, se comparado a outros procedimentos, como é o caso da curva de destilação.

Geralmente, dependendo das condições de infraestrutura laboratorial e de equipamentos, são feitos apenas alguns desses testes que servem de base para a comprovação da qualidade, bem como para a identificação de possíveis fraudes na composição da gasolina. Esses testes incluem a determinação do Teor de AEAC, da Massa específica a 20°C, da Curva de destilação e da Pressão de Vapor à 37,8°C, que são parâmetros ligados à composição do combustível.

Embora a legislação brasileira não especifique um limite máximo ou mínimo para a massa específica da GC, essa propriedade é um parâmetro de qualidade a partir do qual pode-se identificar possíveis adulterações na composição da gasolina (TACKESHITA, 2006). De modo geral, a massa específica da GC fica em torno de 720 a 780 kg/m³, e as adulterações frequentemente provocam uma diminuição nesses valores. Esse efeito é provocado pela adição de produtos menos densos e mais baratos do que aqueles que realmente fazem parte da composição da gasolina.

No presente trabalho a determinação da massa específica à 20°C foi feita por meio do densímetro automático modelo DMA 4500 da Antoon Paar conforme o método ASTM D 4052. Nesse procedimento a amostra é introduzida, com auxílio de uma seringa, em um tubo capilar em forma de U, vibrando com frequência definida. Um banho controlado ajusta a temperatura da amostra a 20°C. A modificação da frequência, provocada pela introdução da amostra, em conjunto com os parâmetros de calibração determina a massa específica da amostra, na

temperatura escolhida (DMA, 2001 apud COCCO, 2008; ABNT, 2001). O tempo de análise para este método é de dois minutos.

A pressão de vapor da gasolina é definida como sendo a pressão total de 1 ml de gasolina com ar saturado à 37,8°C dentro de uma câmara de 5 ml (Valt, 2006). A partir dos resultados de pressão de vapor e da curva de destilação, é possível que se estabeleça um perfil da volatilidade do combustível. Esse perfil de volatilidade está ligado à uma série de características de grande importância para a caracterização da qualidade da gasolina.

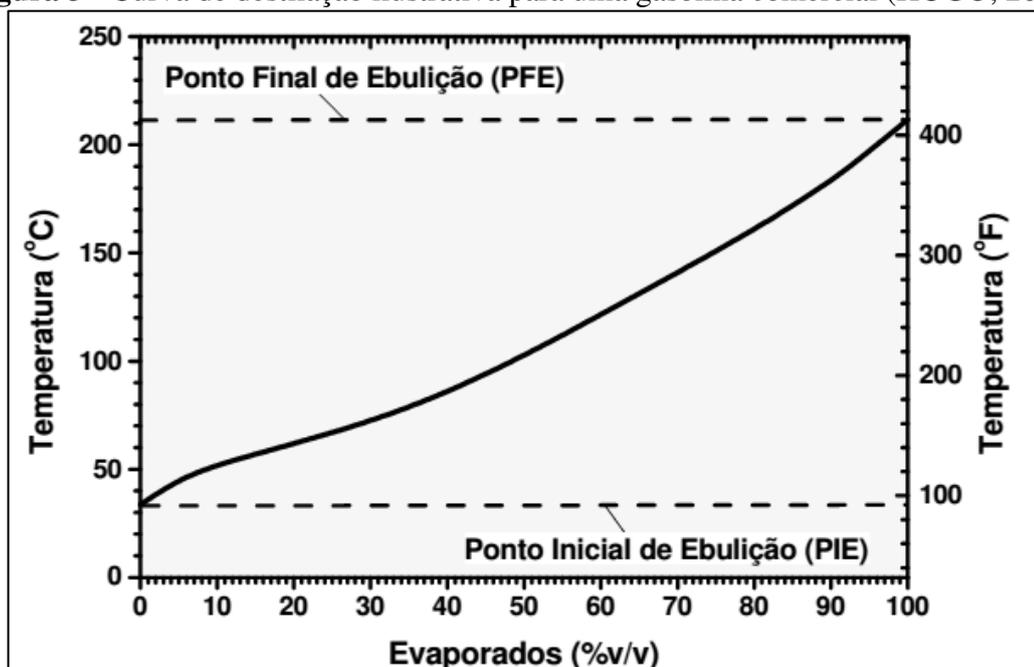
Por meio da pressão de vapor pode-se conhecer qual o comportamento do combustível no motor, principalmente no que se refere à performance e à formação de vapores no interior no motor. Uma gasolina fora das especificações com relação à sua pressão de vapor pode vir a ocasionar problemas nos sistemas internos de distribuição da gasolina. Onde a formação de vapores pode prejudicar o fluxo de gasolina para a câmara de combustão do motor (Cruz, 2003).

Para esse trabalho a determinação da pressão de vapor foi feita no aparelho de pressão de vapor, modelo Minivap VPS da Grabner Instruments segundo a norma ASTM D 5191.

Os resultados de temperatura de ebulição também são importantes na caracterização das amostras de gasolina, pois a partir dos resultados da curva de destilação é possível que se estabeleça uma conexão entre a volatilidade do combustível e sua inteira faixa ebulição. No caso da gasolina, e de combustíveis que são formados por uma mistura de diversos compostos, a destilação ocorre em uma faixa de temperaturas diferentes, onde os compostos mais leves evaporam em temperaturas menores e os mais pesados são destilados em temperaturas maiores. A representação gráfica da temperatura de ebulição em função do volume destilado é chamada de curva de destilação (COCCO, 2008).

A Figura 3 ilustra como fica a representação gráfica da curva de destilação de um destilado típico do petróleo.

Figura 3 - Curva de destilação ilustrativa para uma gasolina comercial (HUGO, 2005).



Fonte: BUARQUE, 2006

Embora a curva de destilação possa ser construída a partir de uma série de resultados de temperaturas e de volumes de combustível destilado, a ANP regulamenta apenas determinados pontos dessa curva. A atual legislação determina limites máximos de temperaturas nas quais 10% (T10), 50% (T50) e 90% (T90) do combustível é destilado. Assim como determina o valor máximo para a temperatura na qual todo combustível é destilado (Ponto Final de Ebulição – PFE) (ANP, 2013).

Esses pontos da curva de destilação da gasolina estão relacionados com o desempenho da gasolina dentro do motor dos veículos. Por isso, a partir dos resultados de temperatura de ebulição é possível que se conheça o comportamento do combustível durante a queima e, conseqüentemente, quais serão os reflexos dessa combustão na dirigibilidade do veículo (HUGO, 2005).

A temperatura na qual 10% do combustível é destilado diz respeito ao comportamento do combustível quando é dada à partida no motor do veículo. Essa temperatura deve ser baixa o suficiente para que o veículo consiga dá a partida, porém alta o bastante para evitar problemas de bloqueio evaporativo causado pela formação de vapores a partir da gasolina.

A T50 está relacionada com as condições ótimas de funcionamento do motor. Essa temperatura deve ser baixa o suficiente para permitir um aquecimento adequado do motor, além de possibilitar uma boa dirigibilidade em climas frios. No entanto, deve ser alta o bastante para

impedir bloqueios evaporativos. Já a T90 e a temperatura do PFE devem ser baixas o suficiente para minimizar a formação de depósitos na câmara de combustão e no cárter do motor, além de facilitar a diluição do óleo do motor (HUGO, 2005 apud RIGA, 2001).

Os resultados dos pontos da curva de destilação da gasolina do banco de dados do presente trabalho foram obtidos a partir de 02 destiladores automáticos Herzog mod. HDA628 (máster) e 02 destiladores automáticos Herzog mod. HDA627 (slave). O procedimento experimental para a determinação das temperaturas de ebulição consiste em destilar 100 ml de gasolina em condições padrão, aferindo, concomitantemente, a temperatura e o volume evaporado nesta. Permitindo que se estabeleça uma relação entre a percentagem de combustível destilado e a temperatura da fração coletada, o que possibilita a construção da curva de destilação do combustível.

1.2. Técnicas de predição de resultados

Com o avanço das técnicas instrumentais nas diversas áreas da química, o volume de dados experimentais à disposição dos pesquisadores cresceu de forma considerável, além disso a interpretação desse conjunto de resultados ficou cada vez mais complexa, uma vez que muitos dos equipamentos utilizados para a coleta de dados o faz de forma indireta, ou seja, a medida experimental baseia-se em um fenômeno físico ou químico que de alguma forma está relacionado com uma determinada propriedade do sistema em estudo. Com isso, para facilitar e melhorar a qualidade e interpretação dos dados experimentais, diversas ferramentas matemáticas foram incorporadas à química, produzindo o que hoje chamamos de quimiometria (BARROS, 2006).

A aplicação de ferramentas matemáticas à química teve um grande impacto no desenvolvimento da instrumentação química, onde hoje se observa que boa parte dos equipamentos usados em laboratório faz uso, em algum nível, de ferramentas matemáticas que possibilitam a obtenção e interpretação dos resultados experimentais.

No ramo da quimiometria, existe um grande número de ferramentas matemáticas que possibilitam a extração de informações de diversos tipos de amostras (COCCO 2008; COCCO, 2005; HUGO, 2005; SOBREIRO, 2009). Dentre as diversas ferramentas quimiométricas usadas no tratamento de dados, as técnicas de PLS, MLR e ANN são alguns dos métodos de predição frequentemente aplicados na obtenção de resultados (ANJOS, 2003; PAULO et al., 2016; TORRES, 2005)

1.2.1. Regressão por mínimos quadrados parciais

Essa técnica de predição baseia-se na relação linear entre as variáveis instrumentais X e as variáveis de interesse Y . Ambas as variáveis são usadas ao mesmo tempo na construção do modelo, durante a etapa de calibração. A matriz de dados instrumentais é decomposta em matrizes de variações das propriedades instrumentais (*loadings* ou pesos) e a posição das amostras (*scores*). Os valores das propriedades originais podem ser considerados combinações lineares das propriedades (pesos) onde os scores representam suas contribuições (Geladi, 1987).

Fazendo a notação em termos de componentes principais, as matrizes X e Y podem ser expressas pelas equações:

$$X = TP^T + E \quad (1)$$

$$Y = UQ^T + F \quad (2)$$

Onde T é a matriz *score* de X , e U é a matriz *score* de Y ; P e Q são as matrizes peso de X e Y respectivamente. Os valores E e F se referem aos resíduos.

A relação linear entre os dois conjuntos de dados pode ser realizada pela correlação entre os escores para cada componente de cada vez. Para isso, faz-se uso da relação linear (MARETTO, 2011):

$$U_h = b_h T_h \quad (3)$$

onde “ h ” refere-se ao número de componentes principais.

Com o intuito de maximizar a covariância de T e U faz-se o ajuste do modelo que fornece os menores valores para as matrizes dos resíduos E e F . Sendo que, ao mesmo tempo, almeja-se atingir uma relação linear ótima entre T e U . Para o PLS faz-se pequenos ajustes nos valores dos escores, o que possibilita que seja encontrada a melhor relação possível. A partir desse ponto as componentes principais passam a ser chamadas de variáveis latentes, pois elas não mais descrevem a máxima variância dos dados. (MARETTO, 2011).

1.2.2. Regressão linear múltipla

Pode-se definir a MLR como sendo a relação linear entre um conjunto de características m , representada por x_j ($j = m - 1$), e uma determinada variável dependente y . Matematicamente, a MLR é expressa como (GELADI, 1987):

$$y = b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m + e \quad (4)$$

$$y = \sum_{j=1}^m b_jx_j + e \quad (5)$$

$$y = X'b + e \quad (6)$$

Na equação (4) x_j são chamadas de variáveis independentes e y é a variável dependente, os termos b_j se referem a sensibilidade e e representa o erro ou resíduo. Na equação (6) y é um escalar, b é um vetor coluna e X' é um vetor linha.

Na resolução de problemas por MLR podem ser encontradas três situações diferentes com relação ao número de variáveis m e o número de amostras n :

- 1) $m > n$. O número de variáveis é maior que o número de amostras. Nesse caso chega-se a conclusão de que existem infinitos valores de b na resolução da equação (4). O que não é viável para a resolução do problema.
- 2) $m = n$. O número de variáveis é igual ao número de amostras. Nesse caso, tem-se apenas uma única solução para o valor de b .
- 3) $m < n$. O número de variáveis é menor que o número de amostras. Nessa situação não é possível encontrar o valor exato de b , no entanto a partir da regressão por mínimos quadrados, através da minimização dos resíduos, é possível que se encontre um valor para b de modo que a variável dependente y possa ser descrita com certa precisão pelas variáveis independentes x_i .

1.2.3. Redes neurais artificiais

Uma ANN é um sistema computacional e matemático cujo processamento baseia-se nas redes neurais dos sistemas biológicos (HAYKIN, 1999). Esse sistema tornou-se atrativo graças a uma série de características que o tornam uma ferramenta de investigação e predição de

resultados altamente sofisticada e flexível. Dentre essas características, a não-linearidade no processamento de dados e a sua alta capacidade de aprender e assimilar padrões de diferentes tipos de amostras, são algumas de suas características mais visadas no estudo de diversos tipos de sistemas (HAYKIN, 1999).

O tipo de rede neural frequentemente usado no tratamento de dados e na predição de resultados é a ANN cuja estrutura é formada por múltiplas camadas de neurônios, do inglês *Multilayer Perceptron* (MLP). Que é formada por uma camada de entrada; uma camada de saída, e entre a camada de entrada e a camada de saída pode haver uma ou mais camadas ocultas (BARRADAS FILHO et al., 2015; COCCO, 2008). A organização das conexões entre essas camadas também é um importante característica durante o processamento de dados.

Na camada de entrada, quando se trata de uma predição, por exemplo, ficam os neurônios artificiais que correspondem as propriedades a partir das quais se deseja fazer a predição de uma outra propriedade. A propriedade a ser predita é representada pelos neurônios da camada de saída da rede. Entre a camada de entrada e a camada de saída podem ser colocadas diferentes quantidades de camadas com diferentes quantidades de neurônios, sendo que quanto maior a quantidade de camadas e de neurônios, maior será a complexidade dos cálculos para a resolução do problema de predição (HAYKIN, 1999).

A ANN usada no presente trabalho, é uma das mais usadas em trabalhos de predição, é uma rede do tipo *full-connected feedforward*. Nesse tipo de rede o processamento do sinal se dá da camada de entrada para a camada de saída, sem que haja conexões entre os neurônios de uma mesma camada (ANJOS, 2003).

1.3. Estado da arte

Na literatura são retratados diversos estudos envolvendo a aplicação de ferramentas quimiométricas no tratamento de resultados referentes a propriedades físico-químicas da gasolina. Boa parte desses trabalhos estão relacionados com a área de controle de qualidade desse combustível, visando principalmente o desenvolvimento de métodos alternativos aos métodos padrões de obtenção de resultados. A seguir são apresentados quatro trabalhos de pesquisa que envolvem a aplicação de ferramentas quimiométricas no tratamento de dados de análises físico-químicas da gasolina com foco na área de controle de qualidade desse combustível.

Em trabalho sobre gasolina C, BUARQUE (2005) fez a predição da densidade relativa a 20°C, do Número de Octano Motor (MON), da Pressão de Vapor Reid (PV REID) e da curva

de destilação da gasolina representada pelas temperaturas T10, T50, T90 e PFE. As ferramentas de predição utilizadas foram as técnicas de MLR e ANN, sendo que os dados de entrada nos dois métodos de predição se referem à composição das amostras de gasolina obtidas por cromatografia gasosa. Nesse trabalho foi demonstrado que as predições obtidas por ANNs apresentaram resultados mais precisos do que os resultados obtidos por MLR. E, apesar de serem observados pequenos desvios nos resultados de predição obtidos por ANNs, o autor considera satisfatório os resultados preditivos obtidos. Uma relevante consideração acerca do desenvolvimento dessa metodologia de predição de dá por conta da utilização de dados de entrada que são obtidos por uma técnica experimental que requer o uso de reagentes de alto custo financeiro, além da utilização de gases que também encarece o processo de análise.

Em seu trabalho sobre gasolina petroquímica (gasolina A) ANJOS (2003) fez a predição de cinco propriedades físico-químicas da gasolina, usando como dados de entrada os resultados de composição obtidos por cromatografia. As propriedades preditas foram a densidade, a pressão de vapor, e as temperaturas de ebulição que representam os pontos T10, T50 e T90. Para isso foram usadas ferramentas quimiométricas de caráter linear e não linear sendo a técnica de ANNs a ferramenta não linear escolhida. Do ponto de vista dos métodos lineares foram construídos polinômios de grau 1 a 3. De modo geral, foram obtidos bons resultados de predição, sendo que, com exceção da densidade, os resultados mais precisos foram obtidos com o uso das ANNs como técnica de predição. Novamente o uso da técnica de cromatografia para a obtenção dos dados de entrada das ferramentas de predição, apresenta-se como um inconveniente nesse método alternativo.

Medeiros (2009) utilizou a técnica de PLS para determinar a concentração de três aditivos diferentes usados em misturas de gasolina C. Os dados de entrada usados no modelo de PLS se referem aos espectros de infravermelho das amostras previamente adulteradas com os três aditivos escolhidos. Ao final do trabalho, constatou-se que é possível determinar a concentração dos aditivos na gasolina tipo C, sendo considerado satisfatório o limite mínimo de concentração de aditivo que o modelo foi capaz de quantificar.

CÔCCO (2008) usou os resultados de infravermelho de 128 amostras de gasolina C para a prever os resultados de densidade, curva de destilação, pressão de vapor e composição desse combustível. As ferramentas quimiométricas usadas no trabalho foram o PLS e a Regressão por Componente Principais (do inglês *Principal Component Regression- PCR*). Sendo que ainda foi proposto a utilização do método de predição por MLR, no entanto, em virtude da extensão da matriz de dados dos espectros da gasolina, não foi possível a implementação desse método. Após as etapas de construção dos modelos de predição, verificou-se que os resultados preditivos

apresentaram pequenos desvios com relação aos resultados experimentais, sendo, dessa forma, considerados resultados satisfatórios. Além disso, verificou-se que os resultados obtidos por PLS foram os que apresentaram as menores taxas de erro de predição.

Na elaboração dos trabalhos descritos anteriormente, a predição das propriedades físico-químicas da gasolina teve como dados de entrada resultados experimentais que estão diretamente relacionados com a composição do combustível. No entanto, a utilização das técnicas instrumentais de cromatografia e infravermelho apresentam inconvenientes relacionados a custo de equipamentos e operação, no caso da cromatografia, e à manipulação da matriz de resultados experimentais no caso do infravermelho. Por isso, torna-se vantajoso o desenvolvimento de métodos de predição baseados em propriedades relacionadas a composição da gasolina que sejam obtidas por técnicas instrumentais mais baratas e cuja manipulação dos resultados se dê de forma mais simples, como é caso das propriedades de massa específica à 20°C e de Pressão de Vapor.

2. OBJETIVOS

2.1. Objetivo geral

Determinar a curva de destilação da gasolina C a partir da pressão de vapor e da massa específica a 20°C desse combustível, aplicando técnicas de predição de resultados no tratamento do banco de dados físico-químicos.

2.2. Objetivos Específicos

- Obter o banco de dados com as propriedades físico-químicas em estudo.
- Otimizar o banco de dados por meio da aplicação do algoritmo KSXY.
- Realizar a calibração de modelos de predição por MLR, PLS e ANN.
- Utilizar os modelos obtidos na etapa de calibração para prever a curva de destilação da gasolina representada pelos pontos nos quais 10, 50 e 90% do combustível é destilado, assim como o ponto no qual todo o combustível é evaporado (PFE).
- Comparar os modelos preditivos obtidos.

3. METODOLOGIA

A metodologia adotada para a elaboração do presente trabalho, se deu em três momentos principais. Primeiramente foram obtidos os dados utilizados na construção dos modelos de predição. Em seguida esses dados foram tratados segundo os métodos de predição estudados, e por fim os resultados obtidos em cada modelo foram comparados para a verificação do modelo mais adequado para a predição dos pontos da curva de destilação da gasolina.

3.1. Obtenção do banco de dados

O banco de dados utilizado foi obtido a partir da literatura, extraído da tese de doutoramento da pesquisadora LÍLIAN CRISTINA CÔCCO, defendida em 2008 sob o título de “PREVISÃO DE PROPRIEDADES FÍSICO-QUÍMICAS E COMPOSIÇÃO QUÍMICA DA GASOLINA A PARTIR DE ESPECTROS NO INFRAVERMELHO” no programa de pós-graduação da Universidade Federal do Paraná - UFPR. Os dados obtidos desse trabalho se referem a 105 amostras de gasolinas que foram coletadas no estado do Paraná e analisadas no Laboratório de Análises de Combustíveis Automotivos da UFPR, através do Programa de Monitoramento de Qualidade de Combustíveis (PMQC) no período de 2005 a 2007.

Nesse trabalho são disponibilizados os dados experimentais de algumas das propriedades físico-químicas cujo resultados são exigidos como parâmetros de qualidade pela ANP. Dentre essas propriedades encontram-se a Massa Específica da gasolina à 20°C, a Pressão de Vapor Reid (PV REID) e a curva de destilação expressa pelas temperaturas nas quais 10 (T10), 50 (T50) e 90% (T90) da gasolina destilada é colocada, bem como especifica o valor do Ponto Final de Ebulição (PFE).

A Tabela 2 resume algumas das características do banco de dados obtido.

Tabela 2 - Resultados exploratórios para o banco de dados utilizado no presente trabalho.

	Valor Médio	Maior resultado	Menor Resultado	Desvio Padrão	Número de amostras
T10 (°C)	56,7	75,1	50,1	4,3	105
T50 (°C)	72,2	76	64,7	1,6	105
T90 (°C)	157,4	190,6	98,1	18,3	105
PFE (°C)	211,0	327,9	135	24,6	105
Massa Específica a 20°C (kg/m ³)	754,5	768	736,9	6,1	105
Pressão de vapor (kPa)	49,4	66,3	18,4	8,2	105

Fonte: Adaptado de CÔCCO, 2008

Essa base de dados inclui amostras de gasolina comum (GC) e gasolina comum aditivada, sendo que algumas dessas amostras encontram-se fora das especificações determinadas pela ANP, o que possibilitou a obtenção de modelos preditivos treinados em uma ampla faixa de variação de resultados das propriedades físico-químicas estudadas.

Os equipamentos utilizados para a determinação dos resultados experimentais das propriedades físico-químicas estudadas apresentaram os desvios de reprodutibilidade expressos na Tabela 3.

Tabela 3 - Resultados de erros experimentais para os equipamentos usados para a obtenção dos resultados experimentais das propriedades físico-químicas estudadas.

Propriedade	Equipamento	Variação (%) 2005	Variação (%) 2006	Variação (%) 2007
Massa específica	Densímetro DMA 4500	0,04	0,04	0,05
PV Reid	Minivap VPS	0,03	0,33	0,14
T10	Destiladores Herzog HDA 628/627	1,37	1,05	1,08
T50	Destiladores Herzog HDA 628/627	0,4	0,36	0,42
T90	Destiladores Herzog HDA 628/627	1,03	1,6	1,04
PFE	Destiladores Herzog HDA 628/627	0,76	4,3	0,85

Fonte: Adaptado de CÔCCO, 2008

Os erros experimentais expressos na Tabela 3 se referem ao desvio de reprodutibilidade de cada equipamento, e são um indicativo da precisão com que cada aparelho determinou os valores experimentais de cada propriedade físico-química estudada. Esses resultados também revelam qual a máxima precisão que poderá ser obtida pela aplicação dos modelos de predição desenvolvidos nesse trabalho, uma vez que os resultados preditivos não podem ser mais precisos do que os resultados experimentais usados na construção do modelo de predição.

3.2. Construção dos modelos de predição

A construção dos modelos de predição seguiu uma sequência a partir do método mais simples (MLR) ao mais complexo (ANN). Partindo-se de uma ferramenta linear para uma não linear, é possível que seja avaliado que tipo de relação que as variáveis de entrada guardam em relação às variáveis de saída.

Para a construção dos modelos de predição, adotou-se o princípio da validação cruzada, na qual a base de dados é dividida em dois subgrupos, e em seguida cada subgrupo é aplicado em um determinado momento da construção do modelo de predição. Esses subgrupos se referem as amostras de calibração e de testes. As amostras da etapa de calibração são usadas para a construção do modelo propriamente dito, já as amostras de testes são usadas para

verificar a robustez e precisão do modelo obtido simulando uma situação real de controle de qualidade.

Com relação a seleção das amostras a serem usadas nas etapas de calibração e testes, utilizou-se o algoritmo intitulado KSXY que se baseia na distância euclidiana para classificar as amostras de acordo com cada etapa do processo de construção dos modelos de predição (GALVAO, 2005). No presente trabalho essa metodologia consistiu em organizar o banco de dados na forma de tabela cujas linhas, enumeradas de 1 a 105, representaram as amostras de gasolina, e as colunas da matriz, enumeradas de 1 a 6, representaram as propriedades físico-químicas em análise. Sendo que nas colunas 1 e 2 foram colocados os dados de massa específica e de pressão de vapor, e nas colunas 3, 4, 5 e 6 às temperaturas T10, T50, T90 e PFE, respectivamente, conforme pode ser acompanhado pela Tabela 15 especificada no Anexo A. Para o processamento do algoritmo optou-se por destinar 70% da base de dados para a etapa de calibração (74 amostras), e 30% para etapa de teste dos modelos preditivos (31 amostras).

A construção dos modelos de predição foi realizada nos softwares MATLAB® R2015a da MathWorks® e no STATISTICA (data analysis software system), versão 10.0 da StatSoft®. Ambos os programas disponibilizam pacotes específicos para o tratamento de dados por PLS, MLR e ANN. No entanto, por questões de praticidade, os modelos de predição por MLR e PLS foram obtidos no programa da StatSoft® e as ANNs foram construídas no MATLAB®.

3.3. Configurações e parâmetros da técnica de predição por MLR

Na construção dos modelos de predição por MLR a equação (4) assumiu a seguinte forma geral:

$$C = a + b_1M + b_2PV \quad (7)$$

Onde **C** é a curva de destilação da gasolina; **a** é o coeficiente linear da equação; **b₁** e **b₂** são os fatores de conversão das variáveis independentes; **M** é a massa específica a 20°C; e **PV** é a pressão de vapor. Dessa forma, a determinação das temperaturas de ebulição da gasolina por MLR consistiu em determinar os coeficientes **a**, **b₁**, e **b₂** da equação (7).

3.4. Configurações e parâmetros da técnica de predição por PLS

A obtenção dos modelos preditivos por PLS consistiu em determinar os parâmetros de predição apresentados nas equações (1), (2) e (3). Sendo que no software STATÍSTICA 10.0 a determinação desses fatores levou em consideração os parâmetros de configuração expressos na Tabela 4.

Tabela 4 - Parâmetros de configuração para os modelos de PLS obtidos para a predição da curva de destilação da gasolina.

Parâmetros	Especificações
Algoritmo de convergência:	NIPALS
Número máximo de iterações:	50
Critério de convergência:	0,0001
Número mínimo de casos válidos por variável:	80%
Número mínimos de variáveis válidas:	80%

Fonte: Autoria própria

3.5. Configurações e parâmetros da técnica de predição por ANN

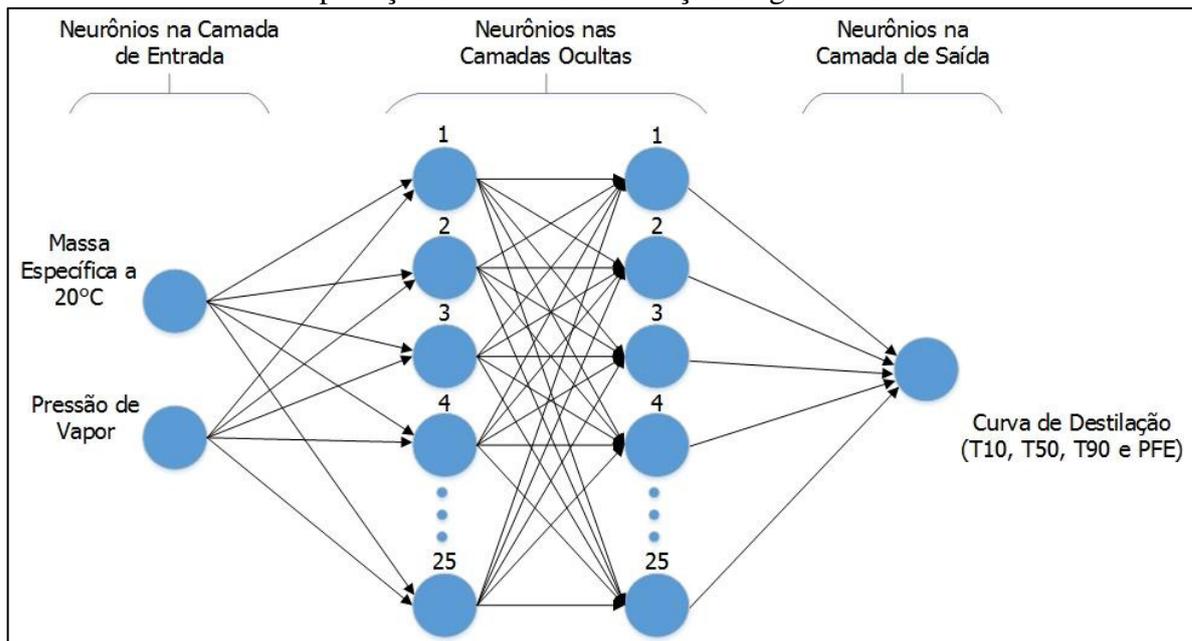
A construção das ANNs é feita com base em um conjunto de parâmetros que devem ser ajustados para a obtenção dos melhores resultados de predição a partir de uma determinada estrutura de rede. O conjunto desses parâmetros e suas especificações estão resumidos na Tabela 5.

Tabela 5 - Parâmetros de ajuste na construção das ANNs e suas especificações.

Parâmetro		Especificação
Função de ativação na camada oculta	Camada 1	Tangente Hiperbólica (Tansig), Logística (Logsig).
	Camada 2	Tangente Hiperbólica (Tansig), Logística (Logsig).
Função de ativação na camada de saída		Linear
Método de convergência		Levenberg-Marquardt (LM)
Número de neurônios nas camadas ocultas		3 a 25
K – particionamento		10
Função de erro		Erro quadrado médio (MSE)
Condições de interrupção do treinamento		Número máximo de épocas = 10000 Meta de desempenho (MSE) = 0 Número máximo de falhas = 20

Fonte: Autoria própria

De modo geral, as ANNs obtidas com base nos parâmetros especificados na Tabela 5 apresentaram a taxonomia ilustrada na Figura 4.

Figura 4 - Diagrama de representação da arquitetura das ANNs treinadas para a predição da curva de destilação da gasolina.

Fonte: Autoria própria

Como ilustrado na Figura 4, os neurônios da camada de entrada correspondem às propriedades usadas para prever a curva de destilação da gasolina. Na camada oculta estão os neurônios usados no processamento dos dados das variáveis de entrada para a obtenção dos resultados da variável de saída. E os neurônios da camada de saída correspondem aos pontos da curva de destilação.

3.6. Avaliação dos modelos de predição

A eficiência dos modelos de predição estudados foi avaliada pelos valores da raiz do erro médio quadrático (do inglês, *Root Mean Squared Error* - RMSE), pelo erro percentual absoluto médio (do inglês *Mean Absolute Percentage Error* - MAPE) e pelo coeficiente de correlação (r) além da análise dos gráficos de erro percentual relativo que auxiliaram na visualização do perfil dispersivo das predições realizadas. Esses parâmetros estatísticos são calculados pelas seguintes equações:

A raiz do erro quadrático médio (RMSE) é calculada pela raiz quadrada do erro médio quadrático dado pela equação (8):

$$RMSE = \sqrt{\sum_i \frac{(y_i - \hat{y}_i)^2}{n}} \quad (8)$$

O erro percentual absoluto medido (MAPE) é encontrado depois de se ter calculado o erro percentual (PE), subtraindo o valor medido pelo estimado e então o dividindo pelo medido conforme a equação (9):

$$PE = \frac{|Valor Estimado - Valor Medido|}{Valor Medido} \quad (9)$$

Em seguida, tirou-se a média dos erros percentuais para que o MAPE fosse calculado, dado pela equação (10):

$$MAPE = \frac{\sum |PE|}{n} \quad (10)$$

O coeficiente de correlação entre as respostas observadas e as previstas foi calculado pela equação (11):

$$r = \frac{n \sum(x_i \cdot y_i) - (\sum x_i)(\sum y_i)}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (11)$$

4. RESULTADOS E DISCUSSÕES

A etapa de otimização do banco de dados, feita através do algoritmo KSXY, aconteceu no software MATLAB® R2015a da MathWorks®. As amostras de teste, selecionadas na etapa de otimização do banco de dados, foram inseridas manualmente no script usado para construir as ANNs e aplicadas nas equações características de PLS e MLR obtidas para cada ponto da curva de destilação. As amostras do grupo de calibração foram utilizadas na construção de cada modelo preditivo de acordo com os parâmetros especificados anteriormente, constituindo uma etapa com características próprias de cada método de predição.

4.1. Otimização do banco de dados

As melhores amostras selecionadas para a etapa de testes dos modelos de predição para cada ponto da curva de destilação estão expressas na Tabela 6.

Tabela 6 - Otimização do banco de dados pelo algoritmo KSXY para a aplicação nos modelos de predição em cada ponto da curva de destilação.

Ponto da curva	Amostras selecionadas para a etapa de teste
T10	3, 6, 10, 12, 14, 15, 16, 17, 18, 19, 22, 23, 25, 32, 33, 34, 35, 55, 56, 60, 61, 67, 69, 76, 79, 80, 87, 88, 91, 98, 101
T50	5, 6, 8, 10, 12, 14, 15, 16, 17, 22, 23, 24, 25, 29, 32, 33, 36, 37, 46, 47, 52, 56, 61, 67, 71, 76, 80, 87, 88, 103, 105
T90	4, 6, 7, 8, 12, 15, 17, 19, 21, 24, 28, 29, 32, 33, 35, 36, 37, 38, 41, 42, 47, 54, 61, 63, 67, 80, 87, 91, 93, 101, 105
PFE	4, 6, 7, 8, 10, 12, 15, 16, 20, 21, 24, 29, 33, 36, 37, 38, 41, 43, 46, 54, 56, 61, 63, 66, 68, 87, 88, 91, 98, 101, 103

Fonte: Autoria própria

O conjunto de testes foi utilizado para verificar a eficiência dos modelos de predição simulando uma situação real, sendo de grande importância na verificação da eficiência dos modelos preditivos obtidos. Por conta disso, para cada ponto da curva de destilação foram

utilizadas as mesmas amostras no grupo de testes de cada método de predição, o que possibilitou a comparação da performance de cada modelo preditivo.

4.2. Predição do ponto T10 da curva de destilação

Na determinação das temperaturas do ponto T10 da curva destilação da gasolina foram obtidos os resultados para os parâmetros de caracterização dos modelos de predição expressos na Tabela 7.

Tabela 7 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto T10 da curva de destilação.

Parâmetros	Métodos de Predição					
	MLR		PLS		ANN	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
RMSE	1,6°C	0,8°C	1,6°C	0,8°C	0,8°C	0,9°C
MAPE	1,82%	1,31%	1,82%	1,31%	0,76%	1,24%
r	0,94154	0,97128	0,94154	0,97128	0,98513	0,94338

Fonte: Autoria própria

Os resultados do modelo de predição por MLR deram origem a equação:

$$T_{10} = 39,03 + 0,0556 \times M - 0,488 \times PV \quad (12)$$

Com base nos parâmetros apresentados na Tabela 7, constata-se que os resultados de predição para a T10 obtidos por MLR ficaram próximos dos resultados experimentais de temperatura tanto na etapa de calibração quanto na fase de testes do modelo. Isso pode ser verificado pelos baixos valores de RMSE e MAPE e pelo bom resultado de coeficiente de correlação. Além disso, os desvios percentuais da Tabela 7 estão próximos dos valores de desvios para os equipamentos de destilação apresentados na Tabela 3.

O processamento dos resultados de massa específica e de pressão de vapor para a determinação dos resultados de temperatura do ponto T10 feito através da técnica de PLS deu origem a seguinte equação preditiva.

$$T_{10} = 39,03 + 0,0556 \times M - 0,488 \times PV \quad (13)$$

A análise dessa equação revela a semelhança dos resultados obtidos nos métodos de MLR e PLS, onde praticamente não houve diferença significativa entre os resultados de predição obtidos em ambas técnicas de predição. Pela análise da Tabela 6, verifica-se que o método de predição por PLS também produziu bons resultados. Apresentando baixos valores de erro e um alto valor de coeficiente de correlação.

O modelo de predição por ANN apresentou resultados de RMSE, MAPE e de coeficiente de correlação melhores do que aqueles obtidos nos modelos de MLR e de PLS. A estrutura da melhor ANN obtida para o ponto T10 é especificada na Tabela 8.

Tabela 8 - Características estruturais para a melhor ANN treinada para a T10.

Arquitetura da ANN			
Estrutura	Algoritmo de treinamento	Função de ativação na camada 1	Função de ativação na camada 2
MLP-2-11-13-1	LM	Tansig	Logsig

Fonte: Autoria própria

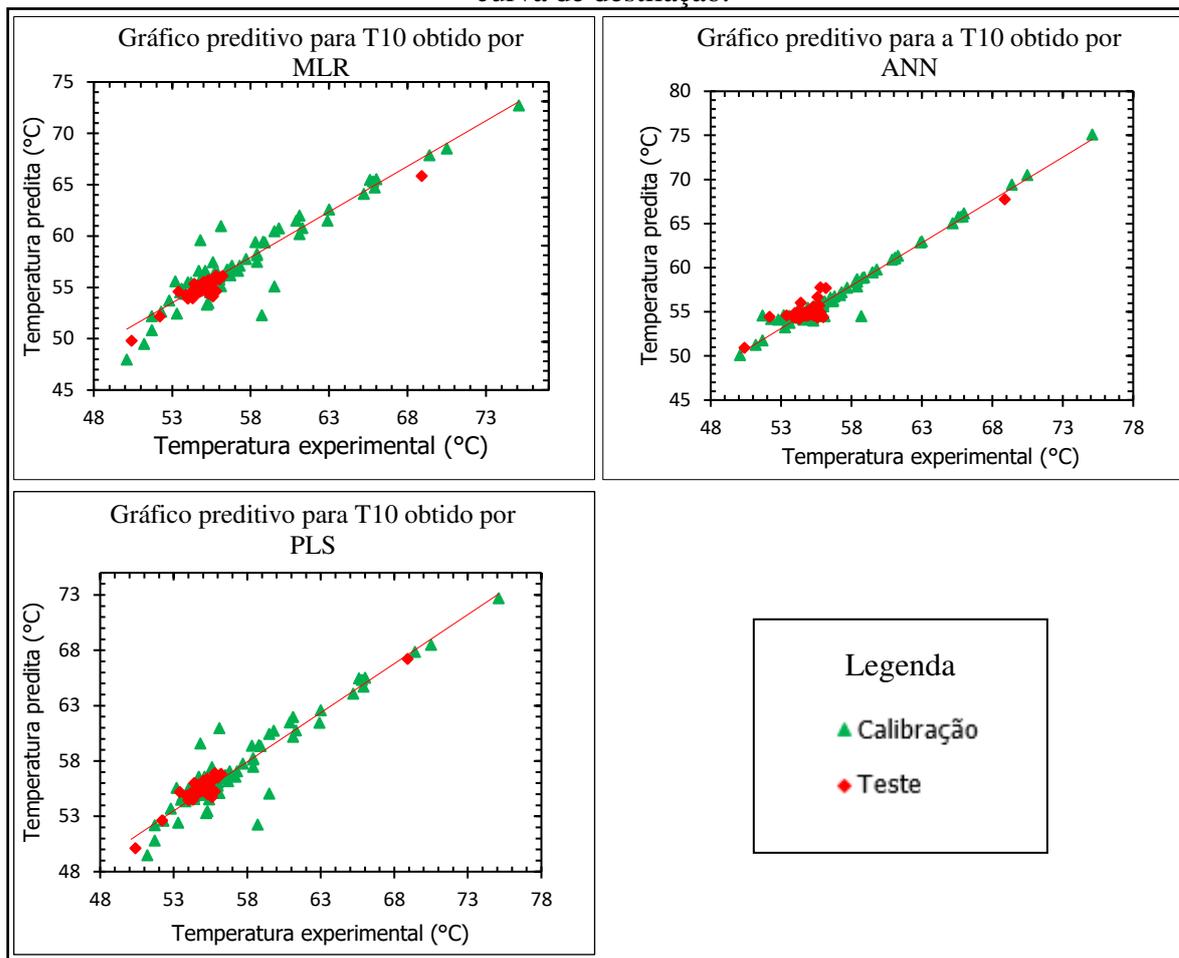
A partir da tabela acima, pode-se inferir que se trata de uma estrutura do tipo MLP, que possui 2 neurônios na camada de entrada (variáveis de entrada); 11 na camada 1; 13 na camada 2 e 1 na camada de saída (variável de saída). O algoritmo de treinamento é o Levenberg-Marquardt (LM), tendo a função Tangente hiperbólica (Tansig) como função de ativação da camada 1, e a função Logística (Logsig) na camada 2.

Diferentemente dos métodos de MLR e PLS, na técnica de ANN não é possível estabelecer de forma direta a relação entre as variáveis de entrada e as variáveis de saída, principalmente por que a processamento dos dados de entrada na direção dos dados de saída é feito com base em funções não lineares, e por esse motivo essa técnica é comparada a uma “caixa preta”, na qual as informações sobre o conteúdo interno não podem ser acessadas facilmente.

A partir da Tabela 7, pode-se inferir ainda que a aplicação dos modelos preditivos no tratamento das amostras de teste, não acarretou em perda de desempenho, o que pode ser verificado através dos resultados de RMSE, MAPE e coeficiente de correlação calculados para essa etapa do processo de construção dos modelos de predição.

Os gráficos de dispersão ilustrados na Figura 5 evidenciam o grau de ajuste dos resultados preditivos com relação aos resultados experimentais de temperatura de ebulição.

Figura 5 - Gráficos preditivos obtidos para os modelos de predição no ponto T10 da curva de destilação.



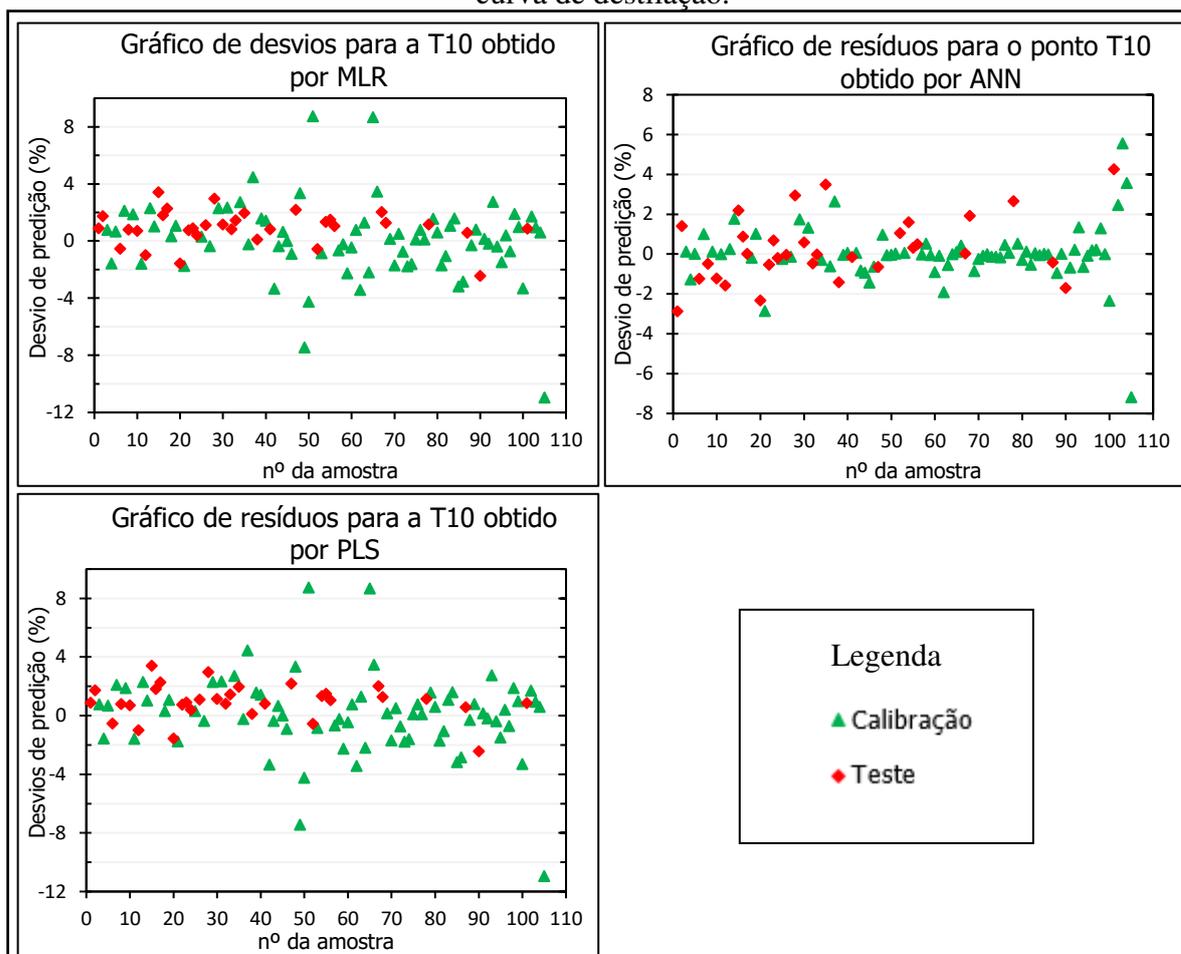
Fonte: Autoria própria

Os gráficos preditivos ilustrados na Figura 5 revelam as semelhanças nos resultados preditivos obtidos pelos métodos de PLS e MLR, onde é possível verificar a semelhança no comportamento dispersivo dos resultados obtidos por ambos os modelos de predição. Verifica-se também o alto grau de ajuste observado no modelo de predição obtido por ANN, onde os resultados preditivos ficaram pouco dispersos com relação os resultados experimentais de temperatura.

Através dos gráficos da Figura 5 é possível notar também o efeito da otimização KSXY, onde as amostras selecionadas para a fase de calibração dos modelos apresentaram maior faixa de variação, enquanto que as amostras do grupo de testes ficaram próximas do resultado médio de temperatura. Essa forma de distribuição do banco de dados possibilitou que modelos de predição tivessem um aprendizado em uma ampla faixa de variação de temperatura, tornando-os mais robustos e ao mesmo tempo sensíveis o suficiente para determinarem a T10 mesmo diante de pequenas variações nos resultados de massa específica e pressão de vapor.

O comportamento dispersivo dos resultados de predição também pode ser avaliado através dos gráficos de resíduos obtidos para cada método de predição. Nesse tipo de gráfico fica mais claro o caráter dispersivo de cada modelo preditivo, uma vez que são mostrados os desvios de predição para cada amostra da base de dados.

Figura 6 - Gráficos de resíduos obtidos para os modelos de predição no ponto T10 da curva de destilação.



Fonte: Autoria própria

Pela análise dos gráficos de resíduos obtidos a partir das predições realizadas por cada modelo de predição fica ainda mais clara a semelhança nos resultados preditivos obtidos por MLR e PLS. Dessa forma, a utilização de ambas as metodologias de predição na determinação do ponto T10 implica na obtenção dos mesmos resultados preditivos, sendo que o método de MLR tem a vantagem de utilizar os dados experimentais sem a necessidade de transformações lineares que por vezes dificultam a interpretação do fenômeno físico-químico.

No caso do modelo preditivo por ANN, observa-se um maior grau de dispersão dos erros de predição na etapa de teste em relação a etapa de calibração do modelo. Esse comportamento

tem efeito mais pronunciado sobre o resultado de MAPE, que consiste de um parâmetro sensível a pequenas variações nos resultados de predição. Por conta disso observa-se um resultado de MAPE ligeiramente maior para a etapa de teste das ANNs, que, no entanto, não chega a comprometer a qualidade do modelo de predição, que nesse caso apresentou os melhores resultados de predição para o ponto T10 da curva de destilação da gasolina.

4.3. Predição do ponto T50 da curva de destilação

A qualidade dos modelos de predição obtidos no tratamento da T50 pode ser avaliada com base na Tabela 9.

Tabela 9 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto T50 da curva de destilação.

Parâmetros	Métodos de Predição					
	MLR		PLS		ANN	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
RMSE	1,1°C	0,4°C	1,1°C	0,4°C	0,3°C	0,3°C
MAPE	0,97%	0,39%	0,97%	0,38%	0,26%	0,33%
r	0,80780	0,76687	0,80780	0,76660	0,98630	0,80446

Fonte: Autoria própria

Os resultados expressos na Tabela 9 indicam a boa qualidade dos modelos de predição obtidos para a predição do ponto T50. Onde são observados pequenos valores de erro, e um coeficiente de correlação que pode ser considerado bom, principalmente quando se leva em consideração os pequenos desvios com relação aos resultados experimentais.

Quando considerado os resultados de MAPE da Tabela 9 em relação aos resultados de desvios dos equipamentos usados na obtenção dos resultados experimentais apresentados na Tabela 3, fica ainda mais claro o grau de ajuste dos modelos de predição obtidos para esse ponto da curva de destilação.

Para o modelo de predição por MLR foi obtido a seguinte equação preditiva:

$$T_{50} = -88,8232 + 0,212585 \times M + 0,012956 \times PV \quad (14)$$

De modo geral, o desempenho do modelo de predição por MLR foi melhor na etapa de testes, apresentando resultados de RMSE e de MAPE menores do que os resultados observados

para a etapa de calibração do modelo. Esse comportamento se deve, principalmente, às características do conjunto de dados usado na etapa de calibração, onde estão presentes as amostras com maior grau de variação com relação à média, o que influencia no processo de mapeamento dos padrões para o estabelecimento da equação de predição. No entanto, esse fator não afetou de forma significativa o desempenho geral do modelo, que ainda pode ser considerado satisfatório para a determinação do ponto T50 da curva de destilação da gasolina.

Para o método de predição por PLS, foi obtida a seguinte equação preditiva.

$$T_{50} = -88,6674 + 0,212360 \times M + 0,013406 \times PV \quad (15)$$

Essa equação é semelhante a equação obtida para a determinação do ponto T50 por MLR. Na verdade, esse comportamento é esperado tendo em vista que a pequena quantidade de variáveis de entrada usadas para predizer a curva de destilação torna o efeito da combinação linear do PLS pouco expressivo para a determinação da variável de saída, tornando os resultados desse método muito semelhantes aos resultados obtidos por MLR.

E tal como foi observado para o método de MLR, no PLS os resultados de erro foram melhores na etapa de testes se comparado com a fase de calibração do modelo. E da mesma forma que no modelo de predição por MLR, essa característica não afetou o desempenho geral do modelo, que obteve resultados preditivos satisfatórios.

No modelo preditivo por ANN nota-se uma pequena melhora nos parâmetros de caracterização em relação aos resultados obtidos por MLR e PLS. Essa melhora pode estar relacionada às características da técnica de ANN, que, por trabalhar também com a parte não linear da relação entre as variáveis de entrada e saída, consegue mapear de forma mais completa os padrões inerentes ao sistema físico-químico e, conseqüentemente, pode obter resultados de predição mais precisos do que aqueles obtidos por MLR e PLS.

A melhor ANN obtida no tratamento desse ponto da curva de destilação é apresentada na Tabela 10.

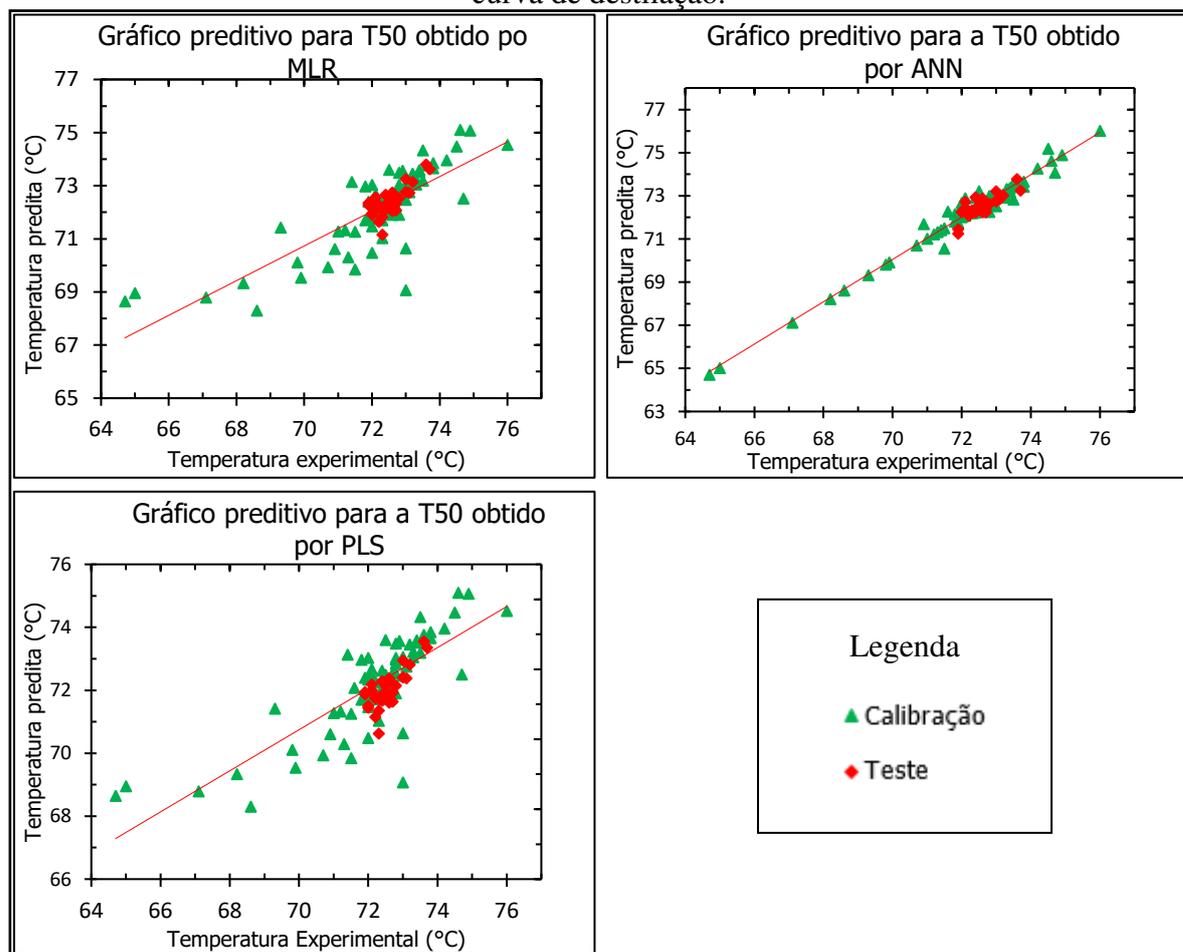
Tabela 10 - Características estruturais para a melhor ANN treinada para a T50.

Arquitetura da ANN			
Estrutura	Algoritmo de treinamento	Função de ativação na camada 1	Função de ativação na camada 2
MLP-2-19-21-1	LM	Tansig	Logsig

Fonte: Autoria própria

A partir da Tabela 10 apresenta uma rede do tipo MLP, que possui 2 neurônios na camada de entrada (variáveis de entrada); 19 na camada 1; 21 na camada 2 e 1 na camada de saída (variável de saída). O algoritmo de treinamento é o Levenberg-Marquardt (LM), tendo a função Tangente hiperbólica (Tansig) como função de ativação da camada 1, e a função Logística (Logsig) na camada 2.

O comportamento dispersivo dos resultados obtidos por cada modelo de predição pode ser verificado através dos gráficos de dispersão ilustrados na Figura.

Figura 7 - Gráficos de predição obtidos para os modelos de predição no ponto T50 da curva de destilação.

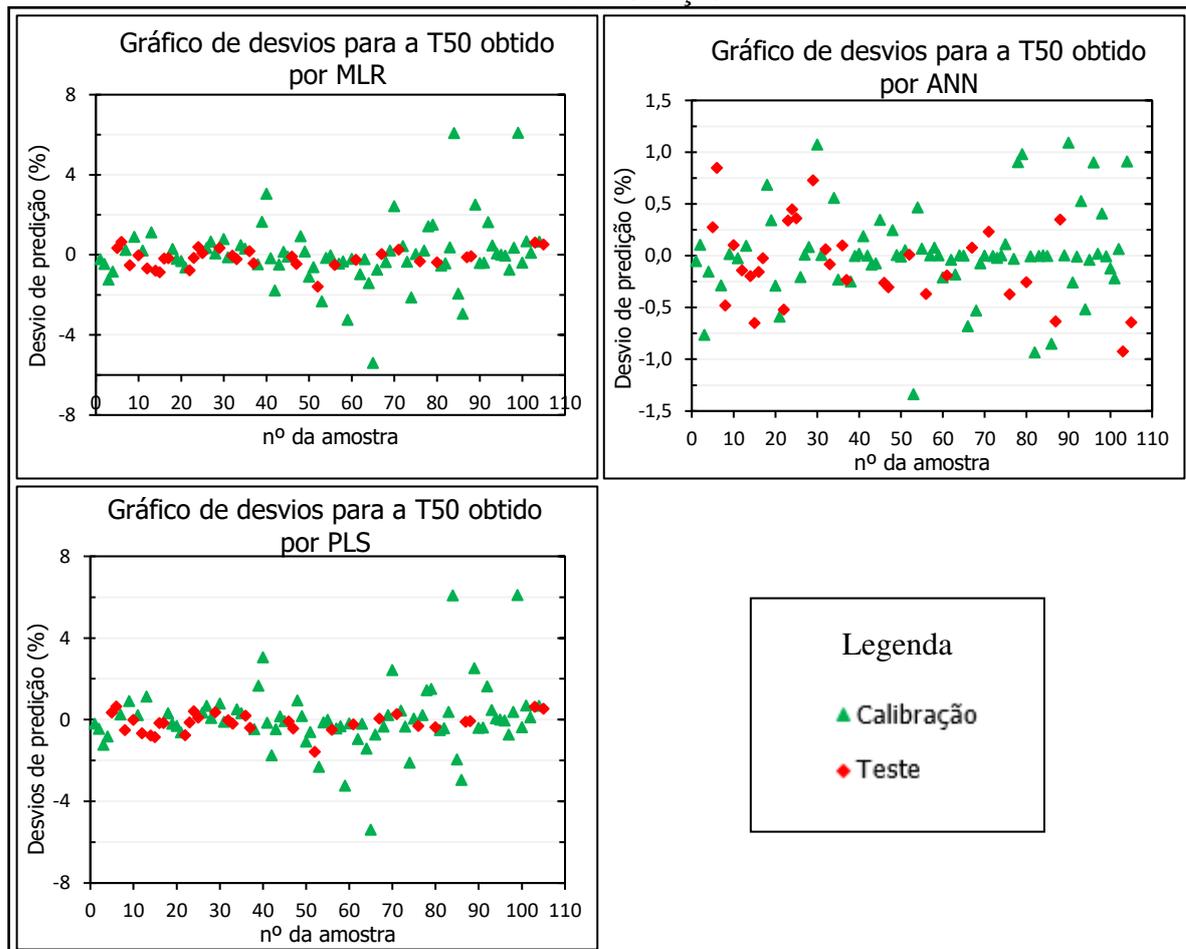
Fonte: Autoria própria

O caráter mais dispersivo das predições realizadas para esse ponto da curva provocou a queda de desempenho com relação ao coeficiente de correlação, embora os valores de erro de predição tenham permanecido na mesma faixa de valores que foram obtidos para o ponto T10. Esse comportamento revela a importância da utilização de mais de um parâmetro de avaliação de desempenho para a caracterização do modelo de predição.

Novamente fica evidente o efeito do processo de otimização KSBY, onde as amostras mais representativas do banco de dados foram designadas para a etapa de calibração dos modelos preditivos e as amostras de teste novamente ficaram em torno do valor médio de temperatura.

Para uma análise mais precisa com relação ao comportamento dispersivo das predições realizadas por cada modelo de predição, são mostrados na Figura 8 os gráficos de desvios para cada método preditivo.

Figura 8 - Gráficos de resíduos obtidos para os modelos de predição no ponto T50 da curva de destilação



Fonte: Autoria própria

Pela Figura 8 nota-se as semelhanças entre as predições realizadas pelas técnicas de MLR e PLS. Porém, verifica-se que em alguns pontos o modelo de predição por PLS obteve um desvio maior do que os valores observados nos resultados obtidos por MLR.

Verifica-se ainda que em ambos os modelos houveram alguns resultados com alto grau de dispersão com relação ao valor experimental de temperatura. Provavelmente esses resultados mais dispersivos foram responsáveis pela queda de desempenho com relação ao coeficiente de correlação desses modelos preditivos.

Por outro lado, os resultados preditivos obtidos por ANN ficaram bem próximos dos resultados experimentais, o que pode ser percebido a partir dos pequenos valores de desvio de predição observados no gráfico de desvios obtido por ANN ilustrado na Figura 8.

De modo geral, considerando os resultados para os parâmetros de caracterização dos modelos de predição apresentados na Tabela 9, pode-se considerar satisfatório o desempenho dos modelos preditivos para o tratamento desse ponto da curva de destilação.

4.4. Predição do ponto T90 da curva de destilação

Os resultados dos parâmetros de caracterização dos modelos de predição obtidos para o tratamento desse ponto da curva de destilação são apresentados na Tabela 11.

Tabela 11 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto T90 da curva de destilação.

Parâmetros	Métodos de Predição					
	MLR		PLS		ANN	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
RMSE	14,6°C	4,6°C	14,6°C	4,6°C	7,7°C	2,8°C
MAPE	6,76%	2,37%	6,76%	2,37%	3,14%	1,34%
r	0,71870	0,26499	0,71870	0,26450	0,93023	0,60205

Fonte: Autoria própria

Para esse ponto da curva de destilação houve uma queda significativa no desempenho dos modelos preditivos, havendo um aumento nos valores de erro e uma grande diminuição nos resultados de coeficiente de correlação, principalmente para a etapa de testes dos modelos de predição. Verifica-se ainda, de modo geral, o aumento considerável dos resultados de MAPE,

indicando o grau de desvio dos resultados de predição com relação aos resultados reais de temperatura para esse ponto da curva de destilação.

Para a técnica de predição por MLR, foi obtida a seguinte equação preditiva.

$$T_{90} = -646,282 + 0,976052 \times M + 1,344869 \times PV \quad (16)$$

A aplicação dessa equação para a predição da T90 produziu resultados muito dispersos com relação aos resultados reais de temperatura. Isso pode ser verificado pelos altos valores de erro e pelo pequeno resultado de coeficiente de correlação, tanto para a etapa de calibração quanto para a fase de testes.

O modelo de predição por PLS também produziu resultados com baixa precisão, o que fica evidente quando são considerados os resultados para os parâmetros de caracterização calculados para esse modelo preditivo. Nesse caso o modelo de PLS produziu a seguinte equação preditiva.

$$T_{90} = -646,282 + 0,976052 \times M + 1,344869 \times PV \quad (17)$$

Essa equação apresenta a mesma forma da equação obtida para o modelo de predição por MLR, o que reforça a ideia de que a pequena quantidade de variáveis de entrada torna o método de PLS muito semelhante à técnica de MLR, uma vez que a combinação linear das duas variáveis de entrada não afeta de forma significativa a informação contida nos resultados dessas propriedades.

O modelo de predição por ANN também sofreu uma queda de desempenho se comparado com os resultados do modelo obtidos nos pontos T10 e T50. Entretanto, a queda no desempenho desse modelo foi menor do que para os métodos de PLS e MLR. A Tabela 12 apresenta a estrutura da melhor ANN construída para esse ponto da curva de destilação.

Tabela 12 - Características estruturais para a melhor ANN treinada para a T90.

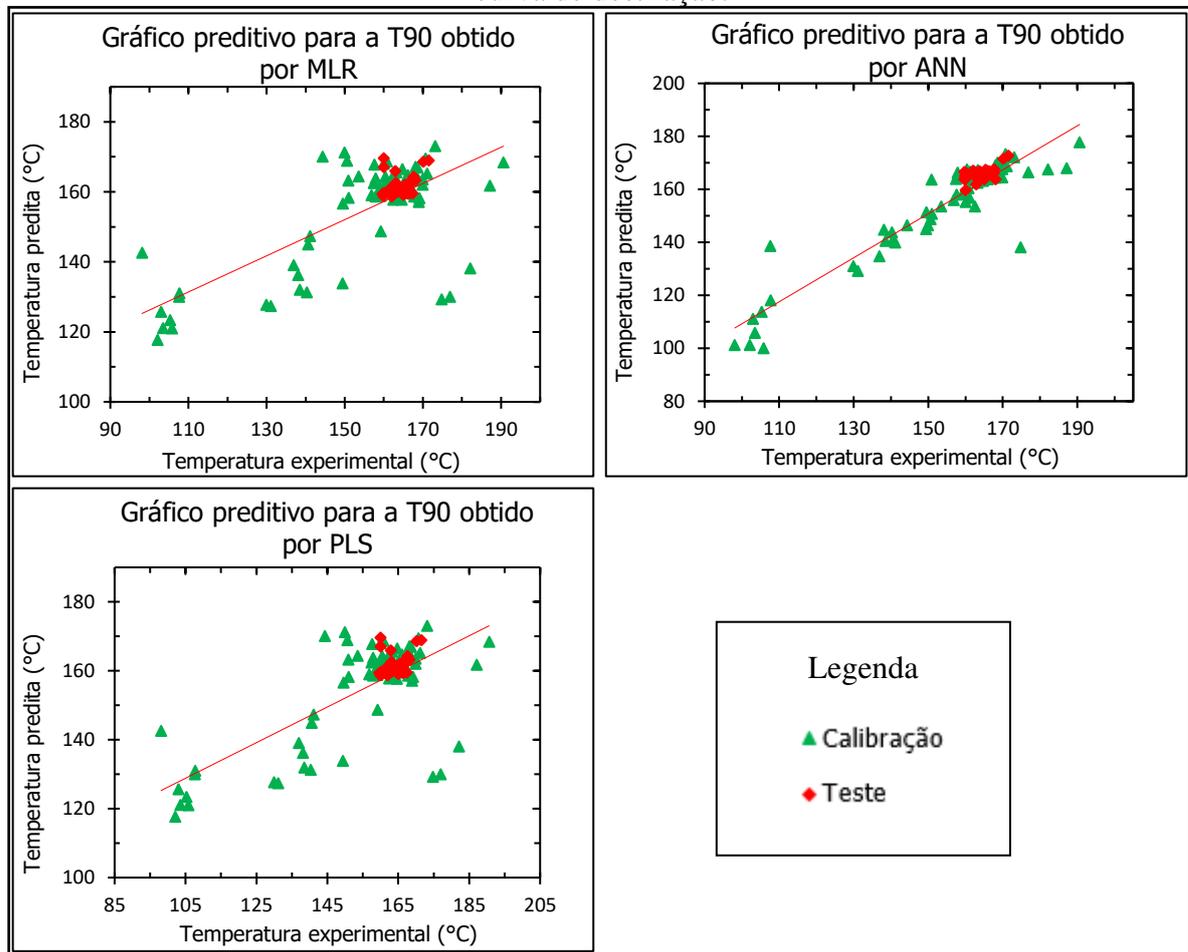
Arquitetura da ANN			
Estrutura	Algoritmo de treinamento	Função de ativação na camada 1	Função de ativação na camada 2
MLP-2-17-11-1	LM	Tansig	Logsig

Fonte: Autoria própria

A rede apresentada na Tabela 12 é do tipo MLP; contém 2 neurônios na camada de entrada (variáveis de entrada), 17 na camada 1; 11 na camada 2 e 1 na camada de saída (variável de saída). O algoritmo de treinamento é o Levenberg-Marquardt (LM), tendo a função Tangente hiperbólica (Tansig) como função de ativação da camada 1, e a função Logística (Logsig) na camada 2.

O comportamento dispersivo dos resultados de predição calculados para esse ponto da curva de destilação pode ser verificado com base nos gráficos de predição ilustrados na Figura 9.

Figura 9 - Gráficos de predição obtidos para os modelos de predição no ponto T90 da curva de destilação.



Fonte: Autoria própria

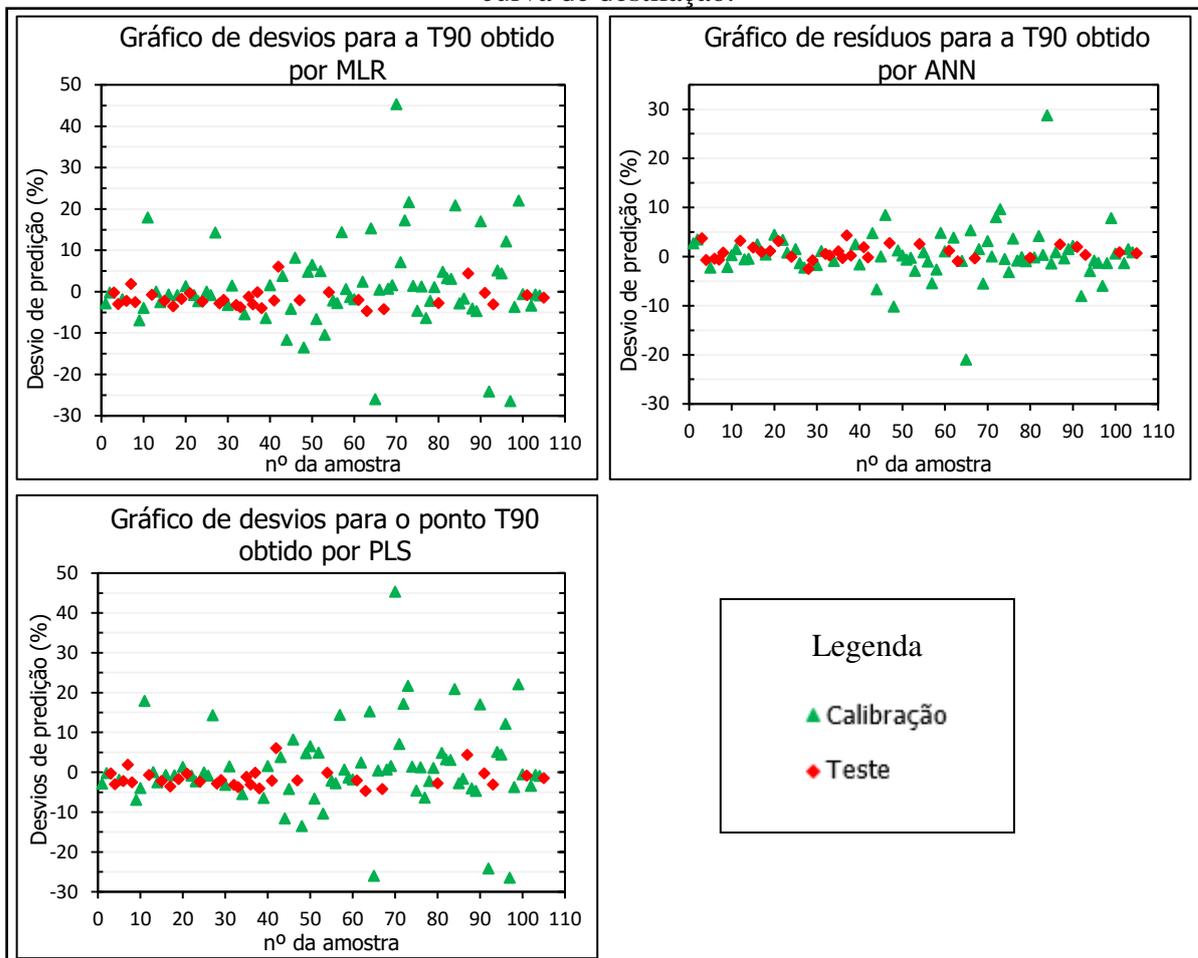
Para os modelos de predição MLR e PLS os gráficos preditivos ilustrados na Figura 9 permitem justificar os resultados para os parâmetros de caracterização desses modelos observados na Tabela 11. E é possível notar também que os resultados de predição da etapa de

testes de fato ficaram mais próximos dos valores experimentais de temperatura, o que justifica os resultados de erro menores se comparado com os resultados da etapa de calibração.

Na comparação do gráfico de predição obtido por ANN com aqueles obtidos por MLR e PLS nota-se o grau de superioridade da técnica de ANN, onde o comportamento dispersivo dos resultados calculados por ANN é bem menor do que aquele observado nos outros dois modelos de predição.

A Figura 10 ilustra os gráficos de desvios obtidos para cada modelo preditivo, e a partir desses resultados pode-se avaliar o comportamento de cada modelo considerando todo o conjunto de dados.

Figura 10 - Gráficos de desvios obtidos para os modelos de predição no ponto T90 da curva de destilação.



Fonte: Autoria própria

Com base nos gráficos ilustrados na Figura 10 verifica-se que os resultados de predição calculados na etapa de testes para cada modelo de predição foram bem mais precisos do que aqueles obtidos na etapa de calibração. Esse comportamento é devido, principalmente, às

características das amostras usadas em cada etapa do processo de obtenção dos modelos, onde verifica-se que as amostras do grupo de calibração apresentam uma ampla faixa de variação enquanto que as amostras do grupo de testes apresentam pouca variação.

Pela análise dos gráficos de predição ilustrados na Figura 9 e considerando os resultados de desvios ilustrados nos gráficos da Figura 10, pode-se inferir que os modelos de predição tendem a produzir resultados com altas taxas de desvios a medida que as amostras se distanciam do valor médio de temperatura. Esse comportamento indica que os modelos preditivos não são muito precisos na determinação da temperatura de ebulição de amostras fora de conformidade com relação as propriedades físico-químicas em estudo.

No caso da técnica de ANN, apesar da significativa melhora com relação aos dois outros métodos, o gráfico de desvios para esse modelo ilustrado na Figura 10 apresenta alguns pontos com valores consideráveis de desvios, com desvios acima de 20% em alguns pontos.

Essa queda de desempenho a partir desse ponto da curva de destilação já era esperada, dado que a partir desse ponto de ebulição, as temperaturas as quais a gasolina é exposta começa a promover a quebra de moléculas do combustível. Essa quebra das moléculas libera fragmentos de moléculas a diferentes temperaturas, provocando variações na temperatura de ebulição do combustível.

4.5. Predição do ponto PFE da curva de destilação

Os resultados dos parâmetros de caracterização dos modelos de predição obtidos para esse ponto da curva de destilação estão expressos na Tabela 13.

Tabela 13 - Resultados dos parâmetros de caracterização para os modelos de predição obtidos para o ponto final da curva de destilação.

Parâmetros	Métodos de Predição					
	MLR		PLS		ANN	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
RMSE	24,2°C	10,0°C	24,2°C	10,0°C	6,9 C	7,8°C
MAPE	6,96%	4,45%	6,98%	4,45%	2,34%	3,11%
r	0,54198	0,20893	0,54198	0,20893	0,96288	0,38355

Fonte: Autoria própria

O mesmo comportamento observado para o ponto T90 também foi verificado para esse ponto da curva de destilação, havendo o aumento ainda mais significativo dos valores de RMSE e MAPE e a diminuição dos valores de coeficiente de correlação. Essa queda de desempenho também já era esperada, haja vista que a temperatura de ebulição desse ponto da curva é maior que a temperatura observada para o ponto T90, o que leva a quebra de moléculas que compõem o combustível, provocando alterações na temperatura de ebulição ainda mais acentuadas.

O tratamento preditivo por MLR deu origem a seguinte equação de predição.

$$PFE = -554,203 + 0,921997 \times M + 1,445600 \times PV \quad (18)$$

As predições realizadas com base nessa equação produziram resultados de temperatura com RMSE de 10°C, sendo que o desvio percentual dado pelo valor do MAPE foi de 4,45%. O que indica pouca precisão nas predições realizadas por esse modelo preditivo.

O modelo de predição por PLS produziu resultados similares aos observados para o modelo de MLR. A equação preditiva obtida para esse modelo é apresentada a seguir.

$$PFE = -554,203 + 0,921997 \times M + 1,445600 \times PV \quad (19)$$

Os resultados dos parâmetros de caracterização para o modelo de PLS foram similares aos calculados para o método de MLR. Dessa forma, com base nos resultados de RMSE, MAPE e coeficiente de correlação expressos na Tabela 13 constata-se que as predições realizadas com base na equação (19) também foram pouco precisas.

Para o modelo de predição por ANN também houve uma piora nos parâmetros de caracterização do modelo que, mesmo assim, exibiu resultados um pouco melhores do que aqueles obtidos pelos métodos de MLR e PLS.

As características da melhor ANN obtida para esse ponto da curva de destilação estão expressas na Tabela 14.

Tabela 14 - Características estruturais para a melhor RNA treinada para o PFE.

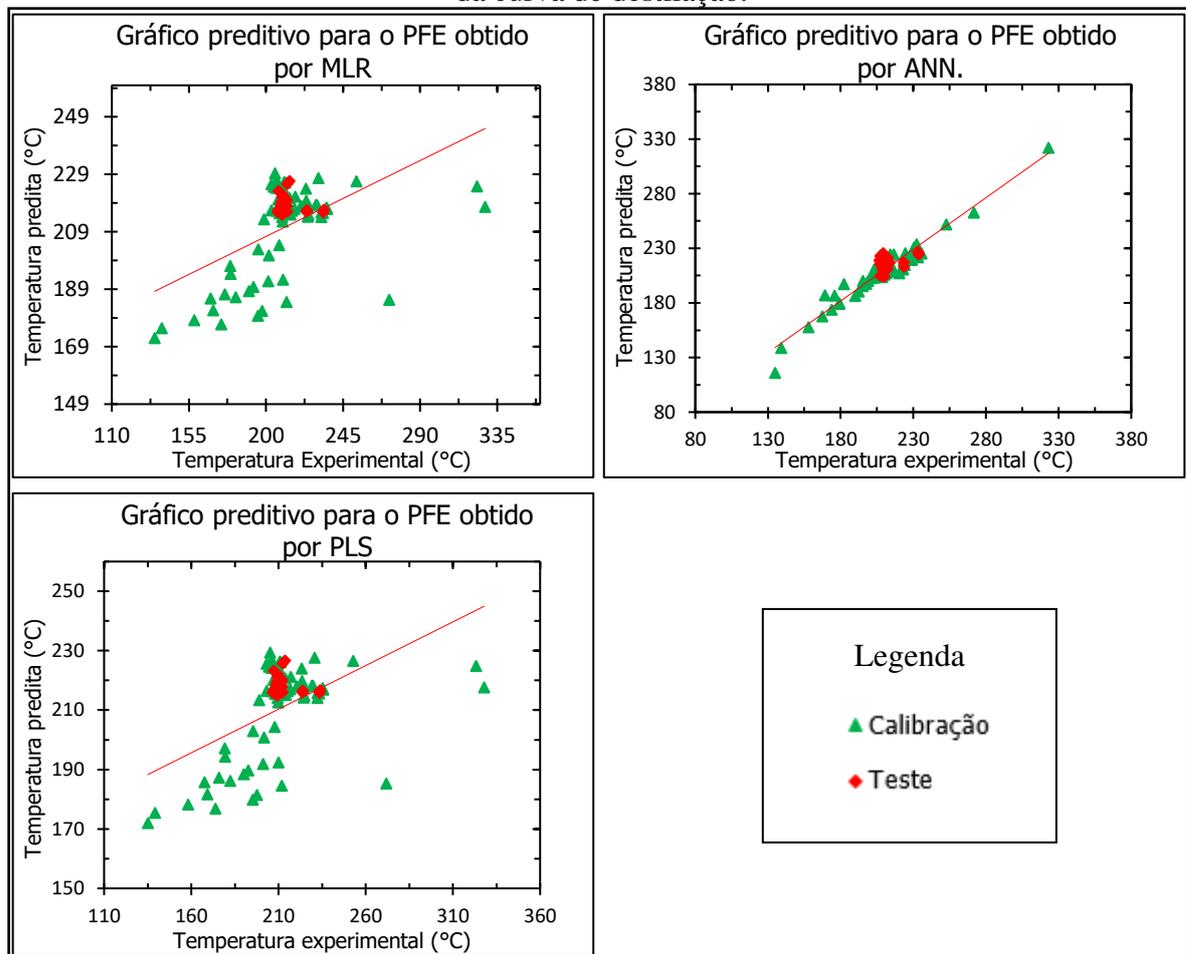
Arquitetura da ANN			
Estrutura	Algoritmo de treinamento	Função de ativação na camada 1	Função de ativação na camada 2
MLP-2-19-24-1	LM	Tansig	Logsig

Fonte: Autoria própria

A ANN descrita na Tabela 14 apresenta uma estrutura de rede do tipo MLP, que possui 2 neurônios na camada de entrada (variáveis de entrada), 19 na camada 1; 24 na camada 2 e 1 na camada de saída (variável de saída). O algoritmo de treinamento é o Levenberg-Marquardt (LM), tendo a função Tangente hiperbólica (Tansig) como função de ativação da camada 1, e a função Logística (Logsig) como função de ativação da camada 2.

Apesar do grau de complexidade da ANN obtida para esse ponto da curva de destilação, os resultados preditivos obtidos por esse modelo não foram tão bons quanto nos pontos anteriores. Com base nos gráficos ilustrados na Figura 11 nota-se a falta de ajuste dos modelos obtidos para esse ponto da curva de destilação.

Figura 11 - Gráficos de predição obtidos para os modelos de predição no ponto final da curva de destilação.



Fonte: Autoria própria

Os resultados de predição obtidos para os modelos de predição por MLR e PLS apresentaram alto grau de dispersão, havendo pouca correlação entre as temperaturas preditas

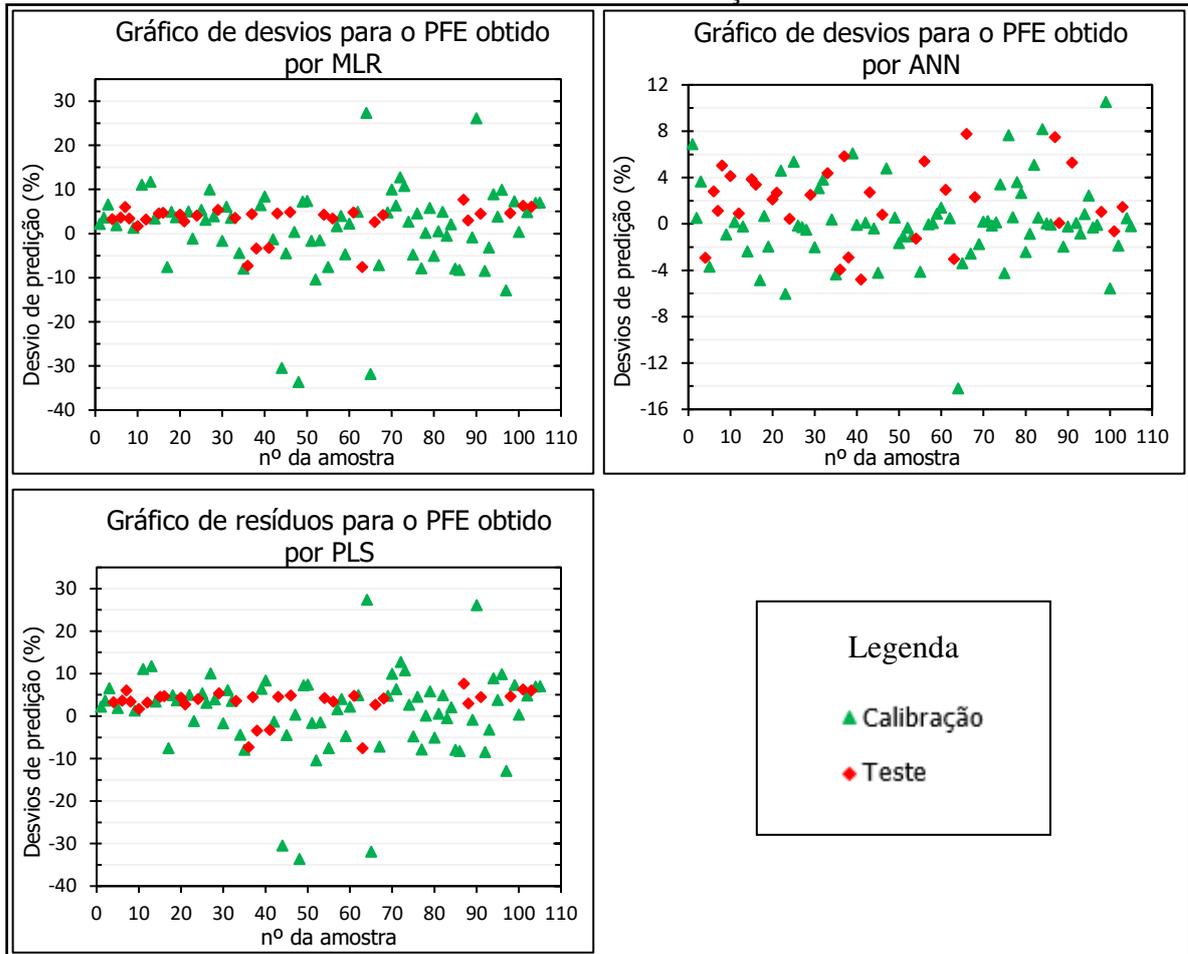
e as temperaturas experimentais, conforme é possível acompanhar pelo distanciamento dos pontos com relação a reta de correlação dos gráficos obtidos para os dois modelos preditivos.

Tal como foi observado para os modelos de predição do ponto T90, as técnicas preditivas também apresentaram desempenho inferior na etapa de calibração se comparado com a fase de testes dos modelos. O que pode ser justificado com base nas características dos dados pertencentes ao grupo de calibração, pois trata-se do grupo com maior caráter dispersivo com relação a faixa de temperatura de ebulição.

Os resultados de predição obtidos para o modelo de ANN foram visivelmente melhores do que os resultados calculados por MLR e PLS. Entretanto, mesmo para esse modelo preditivo os dados de predição apresentaram um grau de dispersão superior aos resultados encontrados para os pontos anteriores.

A Figura 12 ilustra os gráficos de desvios para cada modelo de predição e a partir dos quais é possível que seja verificado de forma detalhada os desvios de predição para cada amostra de gasolina.

Figura 12 - Gráficos de resíduos obtidos para os modelos de predição para o ponto final da curva de destilação.



Fonte: Autoria própria

Para o modelo preditivo por ANN os desvios, embora menores do que os desvios observados nos modelos de MLR e PLS, chegaram a atingir 14% em alguns pontos, entretanto, pelos resultados dos parâmetros de caracterização expressos na Tabela 13, verifica-se que o valor médio de desvio ficou em torno de 2,34% para a etapa de calibração e 3,11% para a etapa de testes.

4.6. Comparativo entre os métodos de predição utilizados

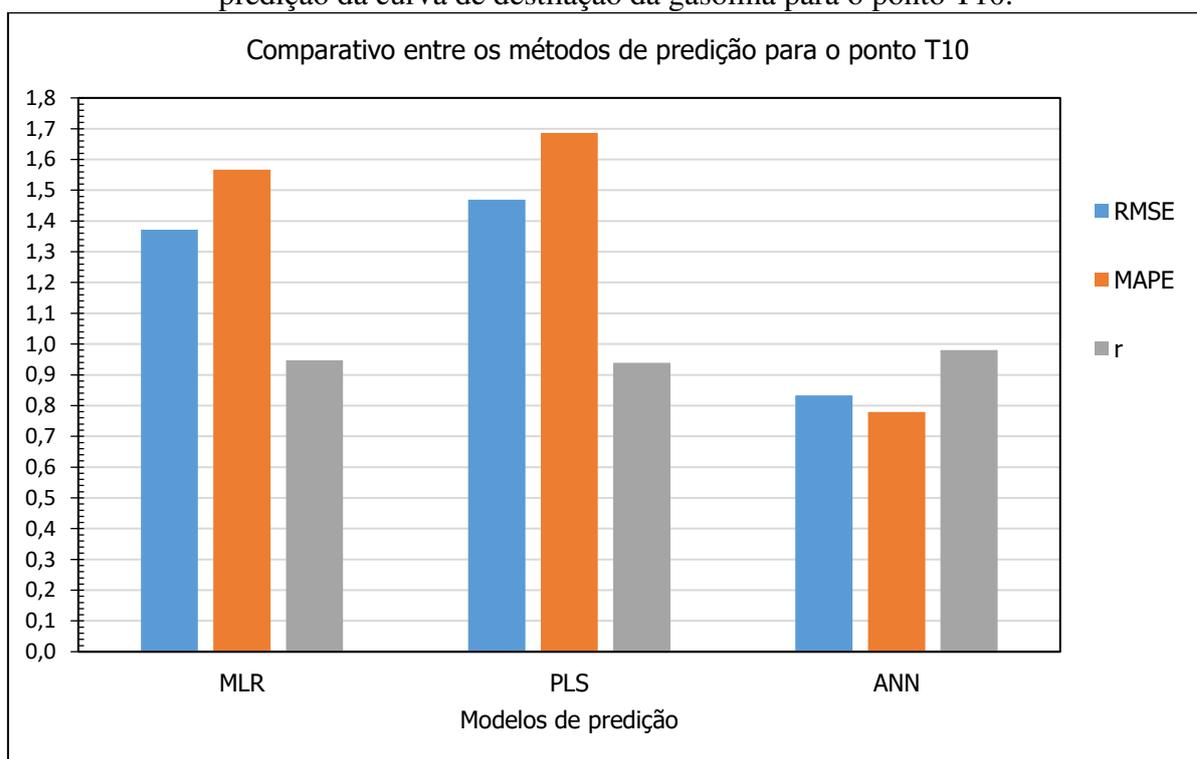
Como foi apresentado até o momento, cada modelo de predição produziu resultados preditivos com diferentes graus de precisão. Dessa forma, para que se possa conhecer aquele com a melhor performance preditiva, é preciso que se estabeleça uma comparação entre os resultados estatísticos de caracterização de cada modelo de predição.

Para a comparação entre os métodos de predição utilizados ao longo do trabalho, serão considerados os resultados de RMSE, MAPE e coeficiente de correlação calculados considerando tanto a etapa de calibração quanto a etapa de testes. Isso quer dizer que será exibido apenas um resultado de RMSE, MAPE e coeficiente de correlação para cada modelo preditivo em cada ponto da curva de destilação, pois dessa forma será possível verificar o desempenho geral de cada modelo de predição.

4.6.1. Modelos preditivos obtidos para o ponto T10

O gráfico da Figura 13 ilustra os resultados para os parâmetros de caracterização dos modelos de predição obtidos para a determinação do ponto T10 da curva de destilação da gasolina.

Figura 13 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para o ponto T10.



Fonte: Autoria própria

Com base no gráfico da Figura 13 fica evidente que os modelos MLR e PLS apresentaram resultados semelhantes para esse ponto da curva de destilação. Entretanto, verifica-se que o modelo de predição por PLS apresentou resultados de MAPE e de RMSE maiores do que no caso da MLR.

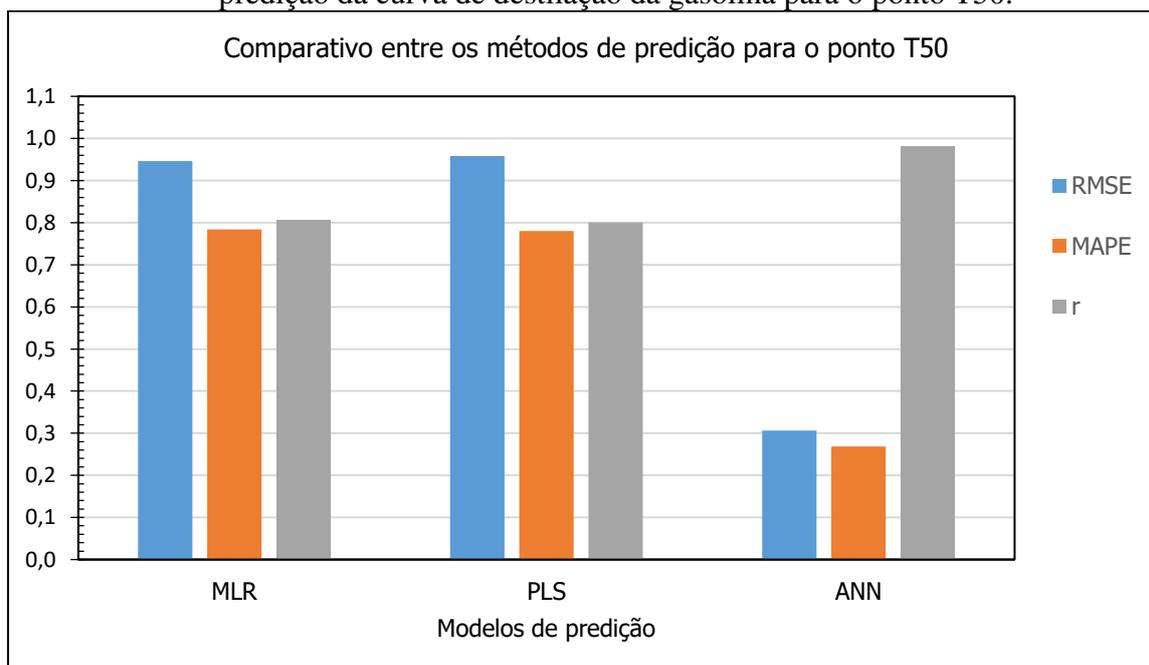
Na comparação dos três modelos preditivos verifica-se que a técnica de ANN apresentou os melhores resultados para os parâmetros de caracterização dos modelos de predição, com baixos resultados de RMSE e MAPE, e com um valor de coeficiente de correlação (r) maior do que o calculado para os demais modelos. Dessa forma, dentre os modelos estudados, o modelo de ANN é o que teve melhor desempenho na predição do ponto T10.

A superioridade da técnica de ANN pode estar relacionada com a forma de processamento dos dados realizados ao longo do processo preditivo. Uma vez que o processo de predição por ANN envolve o uso de funções não lineares e algoritmos que tem a finalidade de otimizar o processo de aprendizagem dos padrões presentes na base de dados usada para treinar as redes, o que aumenta a possibilidade de mapeamento dos padrões presentes na base de dados e que não são verificados facilmente pelo uso de técnicas de predição de caráter linear.

4.6.2. Modelos preditivos obtidos para o ponto T50

A diferença de desempenho entre os modelos de predição para esse ponto da curva de destilação é ainda mais acentuada que no caso anterior. A Figura 14 ilustra o gráfico de comparação entre os parâmetros de caracterização modelos preditivos estudados.

Figura 14 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para o ponto T50.



Fonte: Autoria própria

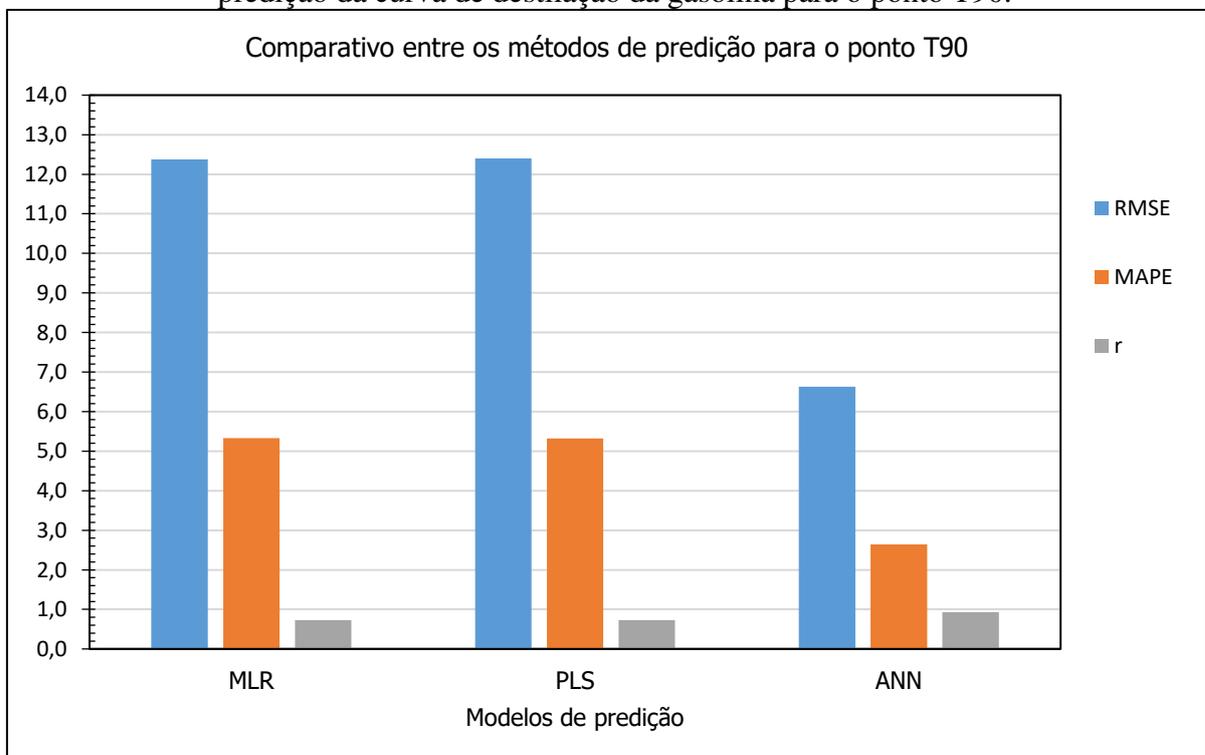
Para o ponto T50 a diferença entre o modelo de ANN e os demais modelos é ainda mais acentuada do que aquela observada na comparação entre os modelos obtidos para o ponto T10. Com base no gráfico da Figura 14 verifica-se que os resultados obtidos para os modelos MLR e PLS foram similares, não sendo observadas diferenças significativas no desempenho desses modelos de predição.

Baseado no resultado do gráfico da Figura 14 constata-se que o modelo de predição por ANN é o mais adequado para a predição do ponto T50 da curva de destilação da gasolina, exibindo pequenos valores de RMSE e de MAPE e um bom resultado de coeficiente de correlação.

4.6.3. Modelos preditivos obtidos para o ponto T90

Para esse ponto da curva de destilação também foi verificado uma diferença acentuada entre os resultados de predição obtidos por MLR, PLS e ANN. A Figura 15 ilustra o gráfico com os resultados de caracterização desses modelos preditivos.

Figura 15 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para o ponto T90.



Fonte: Autoria própria

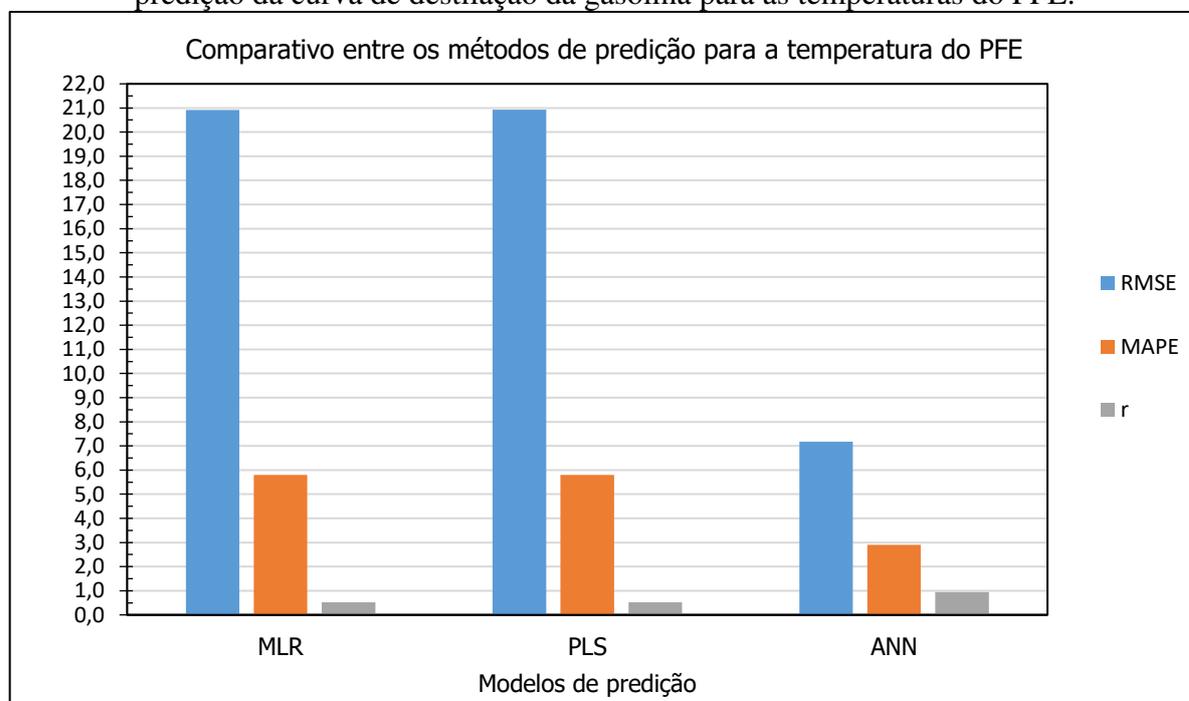
O resultado comparativo ilustrado na Figura 14, evidencia a superioridade da técnica de ANN, que apresentou um desempenho consideravelmente melhor que os demais modelos de predição. Fica clara novamente a semelhança entre os modelos MLR e PLS que não apresentaram diferenças significativas de desempenho.

Com base na Figura 14, constata-se novamente que o modelo preditivo por ANN é o mais adequado para a predição do ponto T90 da curva de destilação. Entretanto, mesmo para essa técnica de predição, os resultados dos parâmetros de caracterização para esse modelo preditivo indicam que é preciso cautela ao se considerar os resultados de predição obtidos por esse modelo, uma vez que foram verificados desvios de predição consideráveis.

4.6.4. Modelos preditivos obtidos para a temperatura do PFE

Para o PFE o resultado da comparação de desempenho entre os modelos de predição foi semelhante ao verificado nos casos anteriores. Esse fato pode ser verificado com base no gráfico ilustrado na Figura 16.

Figura 16 - Comparação entre os resultados de caracterização para os métodos de predição da curva de destilação da gasolina para as temperaturas do PFE.



Fonte: Autoria própria

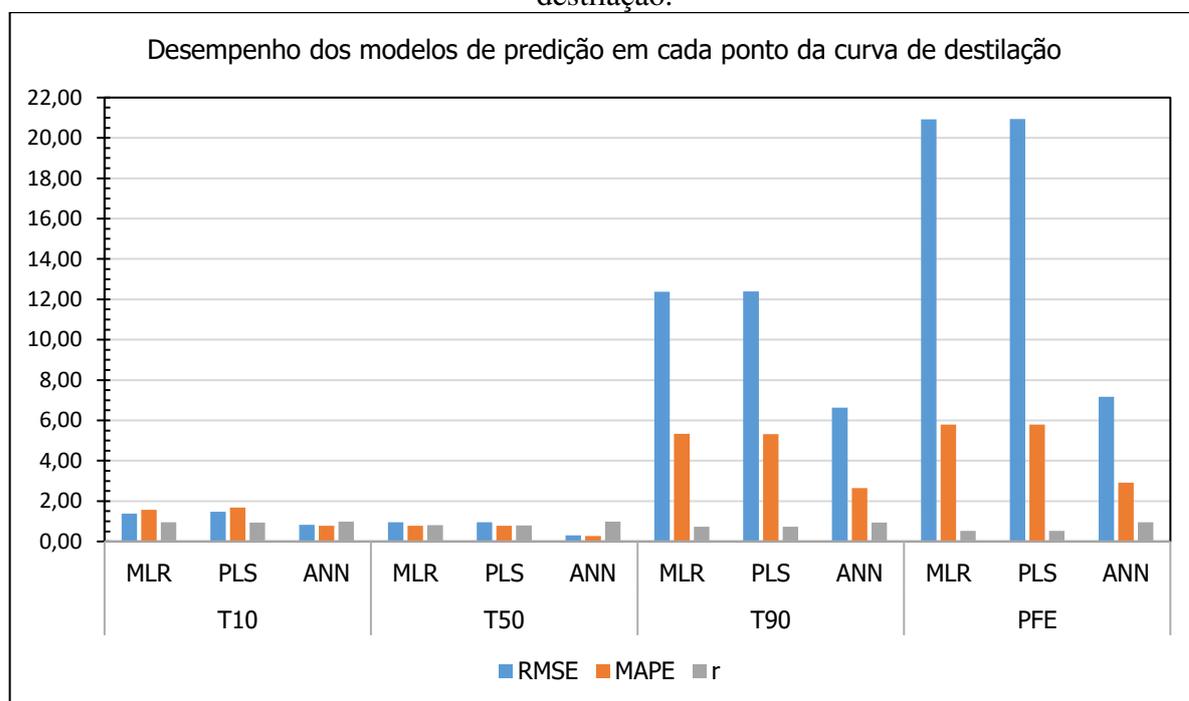
Tal como foi verificado para os demais pontos da curva de destilação, o desempenho do modelo de predição por ANN foi superior ao dos demais métodos preditivos. Entretanto,

novamente é preciso cautela na aplicação do modelo de predição por ANN, uma vez que mesmo para essa técnica foram obtidos resultados com resultados consideráveis de RMSE e de MAPE.

4.6.5. Desempenho geral dos modelos de predição

O gráfico da Figura 17 ilustra o desempenho dos modelos preditivos considerando os quatro pontos da curva de destilação estudados.

Figura 17 - Comportamento preditivo dos modelos de predição em cada ponto da curva de destilação.



Fonte: Autoria própria

Com base na Figura 17 é possível conferir que à medida que os modelos foram sendo construídos partindo do ponto T10 indo até o PFE, a diferença na qualidade dos modelos obtidos por técnicas lineares (MLR e PLS) e o método não linear (ANN) tornou-se cada vez maior.

Podemos correlacionar esse comportamento ao fato de que a medida que a temperatura de ebulição aumenta ao longo da curva de destilação, os sistemas físicos e químicos tornaram-se ainda mais difíceis de serem modelados uma vez que as altas temperaturas verificadas nos pontos T90 e PFE promovem a quebra de moléculas tornando o sistema físico-químico mais complexo e difícil de ser modelado por técnicas lineares. Dessa forma, em função da capacidade limitada, os modelos MLR e PLS apresentaram resultados razoáveis apenas para os dois

primeiros pontos da curva de destilação. Apresentando uma queda de desempenho considerável para o tratamento dos pontos T90 e PFE.

O modelo preditivo obtido por ANN também seguiu a tendência de deterioração da qualidade das predições à medida que a temperatura de ebulição da curva de destilação aumentou. No entanto, essa deterioração foi mais suave do que aquela observada para os demais modelos preditivos. Esse comportamento pode estar relacionado com o caráter não linear da técnica de predição por ANN, que possibilita o aprendizado de padrões que normalmente não são detectados por técnicas lineares.

5. CONCLUSÃO

Com base na literatura foi possível obter a base de dados com os resultados experimentais de curva de destilação, massa específica a 20°C e pressão de vapor da gasolina, possibilitando a aplicação dos modelos preditivos no tratamento das variáveis de massa específica e pressão de vapor para a obtenção da curva de destilação da gasolina.

A aplicação do método de otimização do banco de dados feita por meio do algoritmo KSXY possibilitou obtenção de resultados preditivos otimizados, possibilitando que os modelos fossem calibrados a partir do conjunto de dados com maior faixa de variação, possibilitando a construção de modelos preditivos robustos e precisos.

Com relação à performance dos modelos de predição, verificou-se que para o tratamento do ponto T10, os resultados obtidos pelos métodos preditivos foram semelhantes. Verificando-se apenas uma ligeira superioridade do modelo ANN com relação aos modelos de MLR e PLS. Entretanto, a aplicação de ambos os modelos preditivos na determinação da temperatura de ebulição desse ponto da curva de destilação produz resultados com uma pequena margem de desvios, principalmente quando se considera o próprio desvio experimental para os equipamentos de destilação.

Os resultados de predição obtidos para o ponto T50 representaram uma diferença significativa de performance entre os métodos de predição lineares (MLR e PLS) e a técnica não linear (ANN), onde os resultados obtidos por ANN foram mais precisos do que aqueles obtidos por MLR e PLS. Dessa forma, o modelo de predição por ANN pode ser usado na predição da temperatura de ebulição T50, produzindo resultados com boa margem de precisão tendo em vista a pequena taxa de desvios observada na construção do modelo, e considerando os resultados de desvios experimentais para os próprios equipamentos de destilação.

Com relação aos pontos T90 e PFE da curva de destilação, o modelo preditivo por ANN foi visivelmente superior aos modelos preditivos por MLR e PLS. Entretanto os resultados de predição obtidos por ANN apresentaram taxas de desvios consideráveis, principalmente quando se considera os resultados de desvios experimentais para os equipamentos de destilação. Dessa forma, ainda é necessário um maior aperfeiçoamento desse modelo preditivo para que ele possa ser usado na determinação desses pontos da curva de destilação.

6. ETAPAS FUTURAS

Considerando que a técnica de ANN produziu os modelos de predição com os melhores resultados preditivos, pretende-se fazer ajustes nos modelos de predição obtidos considerando os seguintes aspectos:

- a) Otimizar dos modelos de predição com amostras com diferentes características de não conformidade, possibilitando que os modelos preditivos possam ter um bom desempenho na determinação da temperatura de ebulição de amostras fora das especificações da ANP.
- b) Otimizar das ANNs obtidas no tratamento dos pontos T90 e PFE, considerando os resultados de desvios de predição observados para esses pontos da curva de destilação.

REFERÊNCIAS

- ABNT NBR 7148. **Determinação da massa específica, densidade relativa e API - Método do densímetro.** 2001.
- ABNT NBR 9619. **Produtos de petróleo - Destilação à pressão atmosférica.** 2009.
- AMERICAN SOCIETY FOR TESTING AND MATERIALS – ASTM. ASTM D 86: **Distillation of petroleum products.** 2008.
- ANDRADE, E. T; Carvalho, S. R.G; Souza, L. F. Programa do Proálcool e o etanol no Brasil. **Engvista**, v. 11, n. 2, P. 127-136. 2009.
- ANJOS, L. C. A. **Desenvolvimento e aplicação de modelos matemáticos baseados em redes neurais artificiais e modelos estatísticos na predição de propriedades físico-químicas da gasolina petroquímica.** 2003. 116f. Dissertação (Mestrado em Engenharia Química). Universidade Federal de Pernambuco, Recife-PE. 2003.
- ANP. **Consumo de combustíveis no Brasil cresceu 5,28% na comparação entre 2013 e 2014.** Disponível em < <http://www.anp.gov.br/?pg=74189&m=consum>> Acesso em, 15 mar. 2016. 22:23:45.
- ANP, **Mapa de ações de fiscalização da anp.** Disponível em: <<http://www.anp.gov.br/?id=2906>> Acesso em, 19 jun. 2016. 19:22.
- EPE, **BALANÇO ENERGÉTICO NACIONAL.** Disponível em < <https://ben.epe.gov.br/>>. Acesso em: 20 mar. 2016. 00:12:35.
- BARRADAS FILHO, A. O. et al. Application of artificial neural networks to predict viscosity, iodine value and induction period of biodiesel focused on the study of oxidative stability. **Fuel**, v. 145, p. 127-135, 2015.
- BARROS NETO, Benício; SCARMINIO, Ieda S.; BRUNS, Roy E. 25 anos de quimiometria no Brasil. **Química Nova**, v. 29, n. 6, p. 1401, 2006.
- BASTOS, V. D. Etanol, alcoolquímica e biorrefinarias. **BNDES Setorial**, Rio de Janeiro, n. 25, p. 5-38, 2007.
- BORSATO, D. et al. Aplicação de redes neurais artificiais na identificação de gasolinas adulteradas comercializadas na região de Londrina-Paraná. **Química. Nova**, Vol. 32, n. 9, p. 2328-2332, 2009.
- BRAGA, J. W. B.; POPPI, R. J. Validação de modelos de calibração multivariada: uma aplicação na determinação de pureza polimórfica de carbamazepina por espectroscopia no infravermelho próximo. **Quim. Nova**, v. 27, p. 1004-1011, 2004.
- BRASIL, Portaria MAPA nº 105, de 28 de fevereiro de 2013, ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis; Estabelece as especificações para a adição de álcool

etílico anidro nas gasolinas automotiva no Brasil, **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 01 mar. 2013. Seção 1, p. 3.

BRASIL, Resolução nº 40, de 25 de outubro de 2013, ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Estabelece as especificações para a comercialização de gasolina automotivas no Brasil. **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 28 outub. 2013. Seção 1, p. 84.

BRASIL, Resolução nº 30, de 29 de junho de 2015, ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis; Estabelece as especificações para a comercialização de gasolina automotivas no Brasil. **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 30 jun. 2015. Seção 1, p. 59.

BRUICE, P. Y. **Química orgânica**. 4. ed. São Paulo: Pearson Prentice Hall, 2006. v.1,

BUARQUE, H.L.B. **Predição de propriedades de gasolinas a partir das suas composições**. 2006. 225f. Tese (Doutorado em Física). Universidade Federal do Ceará, Fortaleza-CE. 2006.

BUIATTE, J. E. et al. Qualitative and Quantitative Monitoring of Methyl Cotton Biodiesel Content in Biodiesel/Diesel Blends Using MIR Spectroscopy and Chemometrics Tools. **Journal of the Brazilian Chemical Society**, v. 27, n. 1, p. 84-90, 2016.

CATALUÑA, R.; SILVA, R. Desenvolvimento de um equipamento para avaliação do efeito do etanol na pressão de vapor e entalpia de vaporização em gasolinas automotivas. **Química Nova**, v. 29, n. 3, p. 580-585, 2006.

CÔCCO, L. C. **Aplicação de redes neuronais artificiais para previsão de propriedades da gasolina a partir de sua composição química**. 2003. 169f. Dissertação (Mestrado em Processos Químicos). Universidade Federal do Paraná, Curitiba. 2003.

CÔCCO, L. C. **Previsão de propriedades físico-químicas e composição química da gasolina a partir de espectros no infravermelho**. 2008. 873f. Tese (Doutorado em Engenharia Química) – Universidade Federal do Paraná, Curitiba-PR. 2008.

CRUZ, M. J. F. **Determinação experimental e predição da pressão de vapor de gasolinas com aditivos oxigenados**. 2003. 103f. Dissertação (Mestrado em Engenharia Química) - Universidade Federal de Pernambuco, Recife –PE. 2003.

DETRAN. **Departamento Nacional de Trânsito**. Disponível em: <<http://www.denatran.gov.br/frota.htm>>. Acesso em 15 mar. 2016. 23:18:23.

EPE, **Demanda de Energia 2050**. Disponível em: <<http://www.epe.gov.br/Estudos/Documents/DEA%2013-14%20Demanda%20de%20Energia%202050.pdf>> Acesso em 27 jun. 2016. 01:12:16.

FERREIRA, F. O. **Determinação da composição e da octanagem de gasolinas C por espectroscopia de RMN 'INTPOT. 1 H' e métodos quimiométricos**. 2006. 153f.

Dissertação (Mestrado em Química). Instituto de Química - Universidade Estadual Paulista, Araraquara-SP. 2006.

GALVAO, R. K. H. et al. A method for calibration and validation subset partitioning. **Talanta**, v. 67, n. 4, p. 736-740, 2005.

GELADI, Paul; KOWALSKI, Bruce R. Partial least-squares regression: a tutorial. **Analytica chimica acta**, v. 185, p. 1-17, 1986.

GUEDES, C. L. B. et al. Avaliação de biocombustível derivado do bio-óleo obtido por pirólise rápida de biomassa lignocelulósica como aditivo para gasolina. **Química Nova**, v. 33, n. 4, p. 781-786, 2010.

GUARNIERI, L. C.; JANNUZZI, G. M. Proálcool: impactos ambientais. **Revista Brasileira de Energia**, v. 2, n. 2, p. 147-161, 1992.

HAYKIN, S. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001. 900 p.

LEME, R. M. Álcool combustível derivado da cana-de-açúcar e o desenvolvimento sustentável. **5º Encontro de Energia no Meio Rural**, 2004.

MAHAN, B M.; MYERS, R. J. **Química, um curso universitário**. 4-ed. Traduzido por: Koiti Araki; Denise de Oliveira Silva; Flávio Massao Matsumoto. São Paulo: Edgard Blücher, 1995. 582p

MALFATTI, L. **Análise qualitativa do ciclo real e tempo de combustão em um motor padrão ASTM-CFR operando com mistura de gasolina e etanol hidratado**. 2009. 166f. Dissertação. (Mestrado em Engenharia Mecânica). Universidade Federal do Rio Grande do Sul, Porto Alegre-RS. 2009.

MARETTO, D. A. **Aplicação de máquinas de vetores de suporte para desenvolvimento de modelos de classificação e calibração multivariada em espectroscopia no infravermelho**. 2011. 133f. Tese (Doutorado em Química Analítica). Universidade Estadual de Campinas, Campinas – SP. 2011.

MEDEIROS, A. R. B. **Uso de ATR/FTIR e FTNIR associado a técnicas quimiométricas para quantificação de aditivos em gasolina automotiva**. 2009. Dissertação (Mestrado em Química). Universidade de Brasília, Brasília – DF. 2009.

MORAES, J. E. B. Módulo Experimental para determinação da octanagem de combustíveis. **Braz. Soc. of Mechanical Sciences and Engineering. Proceedings of ENCIT 2006 – ABCM**. Curitiba, PR, dezembro. 5-8, 2006.

PANDORFI, H. et al. Artificial neural networks employment in the prediction of evapotranspiration of greenhouse-grown sweet pepper. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 20, n. 6, p. 507-512, 2016.

PAULO, J. M.; BARROS, J. E. M; BARBEIRA, P. J. S. A PLS regression model using flame spectroscopy emission for determination of octane numbers in gasoline. **Fuel**, v. 176, p. 216-221, 2016.

PETROBRAS, **FATOS E DADOS**. Disponível em: <<http://www.petrobras.com.br/fatos-e-dados/entenda-dez-questoes-sobre-a-nossa-gasolina.htm>> Acesso em 15 mar. 2016. 00:14:08.

PETROBRAS, **Como são obtidas as gasolinas?** Disponível em: <http://www.br.com.br/wps/portal/portalconteudo/produtos/automotivos/gasolina!/ut/p/c4/04_SB8K8xLLM9MSSzPy8xBz9CP0os3hLf0N_P293QwP3YE9nAyNTD5egIEcnQ4MgQ_2CbEdFAGTIInk!/?PC_7_9O1ONKG10GSIC025HRRAB10F4000000_WCM_CONTEXT=/wps/wcm/connect/portal+de+conteudo/produtos/automotivos/gasolina/duvidas+frequentes+gasolinas+petrobras#1> Acesso em 20 jun. 2016. 23:14:35.

SCHIFER, T S.; JUNIOR, S. B.; MONTANO, M. A. E. Aspectos toxicológicos do chumbo. **Infarma**, v. 17, n. 5/6, p. 67-72. 2005.

SCHALE, S. P.; TRANG M. L.; PIERCE, K. M. Predicting feedstock and percent composition for blends of biodiesel with conventional diesel using chemometrics and gas chromatography–mass spectrometry. **Talanta**, v. 94, p. 320-327, 2012.

SILVA, F. L. N. et al. Determination of benzene, toluene, ethylbenzene and xylenes in commercial gasoline from Piauí state. **Química Nova**, v. 32, n. 1, p. 56-60, 2009.

SOBREIRO, V. A.; ARAÚJO, P. H. S. L.; NAGANO, M. S. Precificação do etanol utilizando técnicas de redes neurais artificiais. **Revista de Administração da Universidade de São Paulo**, v. 44, n. 1, p. 46-58. 2009.

TAKESHITA, E. V. **Adulteração de gasolina por adição de solventes: Análise dos Parâmetros Físico-Químicos**. 2006. 102f.. Tese de Doutorado. Dissertação (Mestrado em Engenharia Química) -Universidade Federal de Santa Catarina, Florianópolis. 2006.

TORRES, A. A. L. C. **Caracterização, análise multivariada e predição de propriedades do álcool etílico combustível**. 2005. 100f. Dissertação (Mestrado em Engenharia Química) - Universidade Federal de Pernambuco, Recife – PE. 2005.

VALT, E. **Estimação de parâmetros por algoritmo evolucionário para otimização de um modelo termodinâmico preditivo de propriedades físico-químicas da gasolina**. 2006. 112f. Dissertação (Mestrado em engenharia de Processos). Universidade Federal do Paraná, Curitiba – PR. 2006.

ANEXO A

**Propriedades Físico-Químicas das amostras de gasolina utilizadas no estudo de
predição**

Tabela 15 - Propriedades físico-químicas das amostras de gasolina usadas na construção dos modelos de predição.

Amostra	T10/(°C)	T50/(°C)	T90/(°C)	PFE/(°C)	Massa Espec./(kg.m^{-3})	PV Reid/(KPa)
1	56	72,4	162,4	209,4	754,6	50,2
2	54,2	72,1	159,1	207,9	752,2	52,7
3	54,6	72,8	160,2	203,1	752,8	53
4	55,4	72,3	165	210	751,8	53,9
5	55,4	73	168,1	217	759,2	52,2
6	55,3	72,1	167,1	212,1	755,9	53,4
7	54,9	73,4	162,9	209,4	760,8	51,8
8	54,9	72,7	165,4	210,5	754,8	52,6
9	58,3	72,9	168,9	209,8	761,1	45
10	55,3	72,2	166,1	212,4	754,2	51,8
11	58,4	74,9	144,3	203,1	767,9	49,7
12	55,8	72,3	159,9	208,7	752,4	52,5
13	53,6	73,5	173,1	205,2	764,1	54,7
14	54,2	72,2	163,2	208,7	751,5	53,4
15	53,4	72,7	163,8	207,3	753,6	52,7
16	53,9	72,4	163,3	208,9	754,5	53,5
17	54,6	72,3	165	233,2	754,2	51,5
18	54,4	72,4	166,1	211	756,1	54,3
19	53,8	72	164,1	210,2	752,5	54,3
20	55,6	72,5	160,4	210,1	754,5	53,8
21	56,1	72,7	162	212,1	754,4	53
22	54,9	72,6	161,2	206,1	753,5	52,5
23	54	72	165	220,5	752,7	54,1
24	54,3	72,1	167,7	211,8	755	54,3
25	55,1	72,6	163,2	208,4	756,4	52,9
26	55,7	73,2	166	213,9	760,2	51,2
27	57,3	74,6	149,9	206,2	768	50,5
28	54,4	73	168,1	211,2	758,3	51,6
29	53,6	72,4	167,6	209,5	756,3	53,8
30	54,6	72,1	168,7	223,5	756,4	53
31	54,2	73,3	163,2	209,1	759	52,8
32	54,8	72,2	166	209,7	754,1	52,7
33	54,6	72,3	166,4	209,2	754	52,4
34	55,1	72,1	167,8	224,7	755,6	50,1
35	55,8	72,8	162,8	235,6	758,3	49,8
36	56,6	72,6	165,5	233,7	756,9	50,5
37	53,2	72,4	159,7	206,7	753,8	52
38	55,4	72,4	166,2	223,6	753,6	52,2

Amostra	T10/(°C)	T50/(°C)	T90/(°C)	PFE/(°C)	Massa Espec./(kg.m ⁻³)	PV Reid/(KPa)
39	59,8	68,2	140,3	175,9	741,5	40
40	61,1	69,3	136,9	179,3	751,4	38,6
41	54,7	72,1	163,4	223,8	753,2	52,8
42	51,2	72,3	160	230,6	748	63,8
43	55,4	73,1	157,9	210,7	756,8	53,1
44	54,6	73,4	190,6	323,2	760,3	54
45	55,7	71,9	164,6	224,3	752,6	51,6
46	56,7	73,2	150,9	209,2	758,7	51,3
47	55	72,6	162	214,2	754,7	50,8
48	55,6	72,8	187,1	327,9	760,5	48,9
49	59,5	71,2	149,5	198,9	750,1	52,6
50	50,1	70,7	157,6	210,7	742,7	66,3
51	54,8	72,5	159,2	207,7	754,1	43,8
52	50,4	72,3	160,8	252,8	748,7	62,6
53	61,3	71,5	149,4	192,6	743,9	40,2
54	54	71,8	159,7	207,4	751,8	53,5
55	54,9	72,5	164,6	235,2	755,6	51,9
56	55,3	73,1	166,6	211	756,8	51,7
57	63	68,6	105,8	173,9	736,9	35,7
58	58,4	74,2	161,1	209,3	762,8	47,5
59	62,9	73	138,1	201,2	747,7	39,2
60	58,4	73,8	163,9	211,8	761,4	47,6
61	54,4	72,4	165,4	208,7	754,3	53,6
62	55,2	72,6	160,4	210,9	752,6	56,5
63	55,7	72,7	167,4	233,4	756	50,5
64	69,4	71,3	102,1	135	746,9	26
65	56,1	73	174,8	271,8	740,3	39,4
66	54,7	73	157,9	209,3	755,6	50,1
67	55,3	72,5	166,5	232,3	755,9	50,5
68	55,6	73,3	161,4	209,8	758,3	51
69	56,7	73,6	162,5	210,8	761,7	50,4
70	65,2	71,4	98,1	179,2	759,7	35,2
71	56,8	73,6	153,5	207	761,9	49,8
72	65,8	69,8	105,3	158,1	745,7	31,1
73	65,9	71,5	107,7	167,6	751	32,9
74	53,3	72	156,8	210,8	745,8	57,5
75	56	72,8	169,9	229,1	757,3	51,4
76	55,9	73	159	207,6	757	50,8
77	57,7	72,8	169,1	232,3	758,1	48
78	56,2	72	164,7	216,7	758,3	49,9
79	55	72,5	164,7	210,4	760,8	52,2
80	55,2	72,8	166,4	229,8	755,8	52,3
81	51,7	69,9	151	215,1	741,2	60,3
82	57,2	73,5	157,4	208,3	759	50,5
83	58,8	71	140,6	201,7	750,4	43,7

Amostra	T10/(°C)	T50/(°C)	T90/(°C)	PFE/(°C)	Massa Espec./(kg.m ⁻³)	PV Reid/(KPa)
84	59,5	65	107,6	182,3	739,7	40,4
85	75,1	76	131,1	195,3	767,3	18,4
86	70,5	74,7	129,9	197,7	757,3	25,9
87	55,4	73,7	160	207,3	760,9	52,5
88	55,1	72	167,1	210,6	753	53,2
89	58,9	67,1	138,5	190	738,8	42,5
90	68,9	70,9	103,5	139,1	748,3	27,5
91	55,1	72,8	163,1	209,7	755,7	53
92	65,6	71,8	182,1	210	759,1	32,3
93	54	72	166	224,5	754,9	52,3
94	56	73,8	160	205,9	762	52,5
95	61,1	72,1	141,1	195,4	754,4	42,6
96	56,5	74,5	150,6	204,4	765	50,9
97	66	72	176,9	211,9	752,1	31,4
98	53,5	72,1	169,9	210,5	754,9	54,3
99	60,9	64,7	103	169,2	738,4	38,1
100	55,3	72,8	168,1	223,2	755,5	56,5
101	52,2	71,9	170,1	212,3	754,8	58,1
102	52,8	72,2	171,1	211,8	754,4	55,9
103	51,7	71,9	170,7	213,7	754,5	59
104	52,3	71,6	168,6	209,6	753,3	58
105	58,7	71,9	171,5	211,3	754,2	58,8

Fonte: Adaptado de CÔCCO, 2008