



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE QUÍMICA

TALITA MAGALI DE FARIA

ANÁLISE EXPLORATÓRIA DA QUALIDADE DOS COMBUSTÍVEIS
NA REGIÃO DO DISTRITO FEDERAL E ENTORNO POR MEIO DA
ANÁLISE DE COMPONENTES PRINCIPAIS

BRASÍLIA-DF

1º/2019



UNIVERSIDADE DE BRASÍLIA

INSTITUTO DE QUÍMICA

TALITA MAGALI DE FARIA

ANÁLISE EXPLORATÓRIA DA QUALIDADE DOS COMBUSTÍVEIS
NA REGIÃO DO DISTRITO FEDERAL E ENTORNO POR MEIO DA
ANÁLISE DE COMPONENTES PRINCIPAIS

Trabalho de Conclusão de Curso em
Química apresentado ao Instituto de
Química da Universidade de Brasília, como
requisito parcial para a obtenção do título
de bacharelado em Química.

Orientador: Prof. Dr. Jez Willian Batista Braga

BRASÍLIA-DF

1º/2019

AGRADECIMENTOS

Agradeço a todos que me ajudaram durante esta jornada:

Primeiramente a Deus, que me deu sabedoria, saúde e paciência para a conclusão do curso.

A minha família (Huga, Sineobelim e Otaviano), que é quem mais acredita em mim e sempre me deu suporte em todos os momentos difíceis.

Aos meus amigos Amanda Lima, Daniel Reis, Marcos Felipe, Mateus Portela e todos aqueles que presenciaram minha luta durante a graduação.

A Professora Dra. Fernanda Vasconcelos pela orientação durante o TCC 1 ao meu orientador, Professor Dr. Jez Braga, que teve paciência e disposição para assumir a minha orientação no TCC 2 mesmo com tantas adversidades.

A ANP pela oportunidade de realização do estágio no período de junho de 2016 a dezembro de 2017, no qual pude obter experiência profissional na área de análise de combustíveis e realizar parte significativa das análises incluídas nesse trabalho.

Aos especialistas, técnicos e amigos da ANP pelo incentivo e esclarecer minhas dúvidas quanto ao projeto.

Aos professores e amigos da Universidade de Brasília pela amizade, ensinamentos, companhia e colaboração. Aqui deixo o meu muito obrigada pelos 5 anos de convivência, alegrias compartilhadas e desafios superados.

Ao Instituto de Química e à UnB, pela infraestrutura e auxílio necessário.

RESUMO

Os combustíveis são utilizados na geração de energia em indústrias e meios de transporte. O órgão que regula as atividades dos combustíveis automotivos no Brasil é a Agência Nacional de Petróleo, gás natural e biocombustíveis (ANP). Com o intuito de se estabelecer o panorama da qualidade de combustíveis no País, foi criado o Programa de Monitoramento de Qualidade de Combustíveis (PMQC). Esse programa consiste na coleta, transporte e análise dos combustíveis etanol, gasolina e óleo diesel. Diversos ensaios físico-químicos são realizados regularmente pelos laboratórios credenciados pela ANP. Um de seus laboratórios, o Centro de Análises e Pesquisas Tecnológicas (CPT) está localizado em Brasília e analisa combustíveis do Distrito Federal e Tocantins. Neste trabalho foi realizada a análise exploratória dos resultados das análises realizadas no programa de monitoramento, disponibilizados pela ANP em seu portal online, utilizando Análise de Componentes Principais (PCA). As variáveis originais escolhidas foram ensaios (testes) como: cor, teor de etanol ou biodiesel, pontos específicos da curva de destilação e ponto de fulgor. O modelo PCA é uma das ferramentas da quimiometria que elimina a correlação entre as variáveis (ensaios realizados) e facilita a observação de padrões entre as amostras. Foram analisadas amostras de gasolina e óleo diesel, com base nos resultados disponíveis nos boletins do PMQC. Os cálculos foram realizados no software Matlab®, onde buscou-se informações de acordo com a UF onde as amostras foram coletadas, bem como distribuidora e tipo de combustível. Os modelos PCA criados para os combustíveis foram construídos com 4 PCs e capturaram variâncias acima de 80% dos dados. Com PCA foram gerados gráficos de escores e pesos que permitiram observar que as amostras de gasolina tiveram influência do local onde foram coletadas, provavelmente por terem vindo de refinarias diferentes. Além disso, a gasolina comum, aditivada e Premium apresentaram agrupamentos distintos principalmente devido à cor e teor de etanol. Já o diesel apresentou escores que dividiram as amostras entre S500 e S10. As variáveis cor e massa específica estão diretamente ligadas e basicamente capturaram a mesma informação. Foi possível observar como PCA possibilitou a análise exploratória de uma grande quantidade de dados, destacando a presença de amostras com comportamento anômalo. Os resultados mostram que a quimiometria é uma área imprescindível para obter informações sobre as análises realizadas em laboratórios e pode melhorar o estudo sobre a qualidade de combustíveis no Brasil.

Palavras chaves: Combustíveis, Quimiometria, Análise multivariada, Petróleo, Qualidade

ABSTRACT

Fuels are used for energy generation in industries and means of transport. The official agency that regulates the activities of the automotive fuels in Brazil is the National Agency of Petroleum, natural gas and biofuels (ANP), in order to draw up a simplified overview of the quality of fuels in the country was created the Program of Monitoring of Quality of Fuels (PMQC). This program consists of the collection, transporting and analyzing of ethanol, gasoline and diesel fuels. Several physical-chemical tests have been running regularly by accredited laboratories by the ANP. One of them is the Center for Analysis and Technological Research (CPT) located in Brasília, which analyzes fuels from Federal District and Tocantins. This work has come out with the exploratory study of the results of the experiments of monitoring program, that are available on the ANP website and it was performed using Principal Component Analysis (PCA). The original variables chosen were tests called: color, ethanol or biodiesel content, specific points of the distillation curve and flash point. The PCA is one of the tools of chemometrics that eliminates the correlation between the variables (performed tests) and facilitates the observation of patterns between the samples. Gasoline and diesel oil samples were analyzed based on the results available in PMQC bulletins. The calculations were performed in the Matlab® software and patterns among the samples according to the location where the samples were collected, as well as distributor and fuel type were sought. The PCA models created for fuels was made with 4 PCs and captured variances above 80% of the data. The PCA scores and loadings plots allowed observing that the gas samples had influence of the place where they were collected, probably because they come from different refineries. In addition, gasoline with additives, regular or Premium presented distinct clusters, mainly due to the color and ethanol content. The diesel presented scores that divided the samples between S500 and S10. The variables color and specific mass were linked and it basically captured the same information. It was possible to observe how PCA was useful to enable an exploratory analysis of a large amount of data, highlighting the presence of few samples with anomalous behavior. The results show that chemometrics is an essential area for obtaining information about analyzes carried out in laboratories and can improve the study on the quality of fuels in Brazil.

Keywords: Fuels, Chemometrics, Multivariate analysis, Petroleum, Quality.

SUMÁRIO

1. INTRODUÇÃO	7
2. OBJETIVOS	10
3. REVISÃO BIBLIOGRÁFICA	11
3.1 Combustíveis	11
3.2 Legislação da Agência Nacional do Petróleo (ANP)	12
3.2.1 Programa de monitoramento de qualidade de combustíveis na ANP	13
3.2.2 Ensaios realizados na gasolina, etanol e óleo diesel no Programa de Monitoramento de Qualidade de Combustíveis (PMQC)	15
3.3 Quimiometria	18
3.3.1 Análise de Componentes Principais (PCA)	19
4. MATERIAIS E MÉTODOS	22
4.1 Conjunto de dados	22
4.2 Estrutura e análise de dados	22
4.3 Variáveis de estudo	23
4.3.1 Variáveis para óleo diesel	23
4.3.2 Variáveis para gasolina	24
5. RESULTADOS E DISCUSSÃO	26
5.1 Modelagem por PCA para amostras de óleo diesel	27
5.2 Modelagem por PCA para amostras de gasolina	31
6. CONSIDERAÇÕES FINAIS.....	37
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	37

1. INTRODUÇÃO

Os combustíveis são de grande interesse econômico, pois são utilizados na geração de energia (PESQUISA; EPE, 2018). A descoberta do petróleo no Brasil se deu por volta de 1930 (PAULA; SOARES, 2012), o seu refino transforma o óleo cru em derivados de elevado valor comercial, que são utilizados em meios de transportes e indústrias como combustíveis, tais como: óleo diesel, gasolina, GLP, querosene e outros (PORTAL DA ANP, 2019). Contudo, por ser um recurso esgotável, está sendo substituído por fontes renováveis, como é o caso do álcool etílico ou etanol combustível, um biocombustível considerado limpo para o meio ambiente e que é obtido majoritariamente através da fermentação de açúcares da cana de açúcar, uma fonte renovável (LEITE et al., 2009).

O setor de transportes está entre os setores que mais consomem energia no país, representando 32,7%, do total, ficando apenas atrás das indústrias com 32,9% (PESQUISA; EPE, 2018). De acordo com o relatório síntese do Balanço Energético Nacional (BEN) de 2018, é possível observar o consumo final de produtos energéticos no Brasil no ano de 2017 (Figura 1), sendo que dentre os combustíveis destinados ao transporte, o óleo diesel ocupa o primeiro lugar em termos de utilização, sendo que a gasolina e o etanol ficam, respectivamente, em segundo e em terceiro lugar. Tendo em vista a maior contribuição desses combustíveis, o presente trabalho de conclusão de curso será restrito a uma análise exploratória dos dados de qualidade de 2 combustíveis: óleo diesel e gasolina.

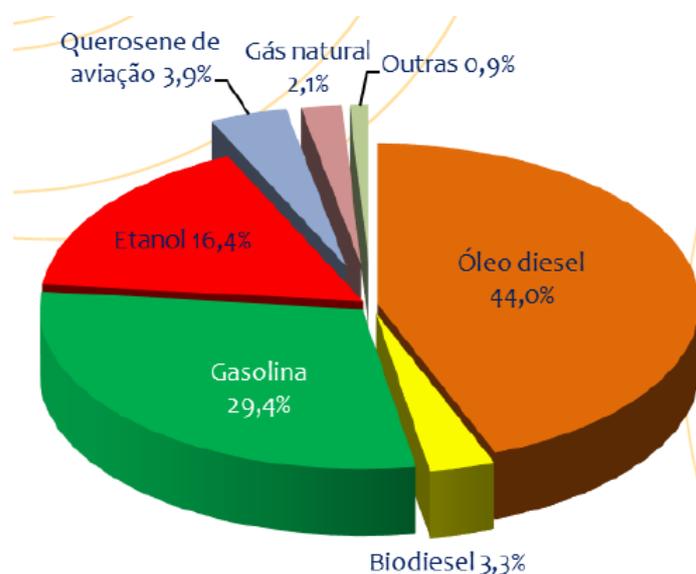


Figura 1 - Consumo final de energia nos transportes do Brasil em 2017. Fonte: (PESQUISA; EPE, 2018).

A gasolina e o óleo diesel são produtos derivados do petróleo, que basicamente, consistem em uma mistura complexa de hidrocarbonetos obtidos a partir do seu refino. A separação em frações é feita utilizando métodos físicos e químicos como decantação, destilação e craqueamento. A composição final dos combustíveis depende do seu processo de refinaria tanto quanto da composição do petróleo. A Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) é o órgão brasileiro responsável por estabelecer normas a fim de regular e monitorar as atividades da indústria que envolvem desde a extração até o comércio dos combustíveis automotivos.

A ANP publica resoluções que mantêm o controle de qualidade dos produtos comercializados no país, ou seja, especificam as características físico-químicas de cada combustível, de modo que tenham a qualidade e o desempenho esperados para a segurança do veículo e do consumidor minimizando impactos ambientais. Neste sentido, a Agência possui o Programa de Monitoramento da Qualidade dos Combustíveis (PMQC) que é realizado com amostras coletadas em todo o território nacional. Este monitoramento é feito ininterruptamente, com o intuito de monitorar a conformidade de cada combustível e os resultados das análises são disponibilizados no próprio site da Agência (PORTAL DA ANP, 2019). Algumas dessas análises são os ensaios denominados: cor e aspecto, destilação, massa específica a 20 ° C, ponto de fulgor, teor de biodiesel e teor de etanol anidro combustível.

Diversos laboratórios são contratados pela ANP para a realização das análises. O principal objetivo do PMQC é levantar indicadores gerais da qualidade, a fim de homogeneizar os produtos de norte a sul do País e verificar que todo combustível cumpra as especificações estabelecidas. O programa identifica a origem de amostras conformes ou não e assim aperfeiçoa a fiscalização da Agência (PORTAL DA ANP, 2019).

Os resultados das amostras coletadas pelos laboratórios que participam do PMQC são disponibilizados mensalmente e incluem dados desde o local da coleta até os resultados de cada ensaio (teste) realizado para cada combustível. Ao fazer uma análise exploratória da qualidade dos produtos no Brasil, pode-se comparar esse conjunto de dados e observar, por exemplo, quais são as semelhanças e diferenças entre as amostras de combustíveis de acordo com local, data ou ensaio. Para facilitar a visualização dessas informações são utilizados métodos multivariados que fazem parte da área de pesquisa de quimiometria.

A definição que será utilizada neste trabalho resume a quimiometria como “A área da química que usa métodos matemáticos e estatísticos para: planejar ou otimizar

procedimentos experimentais; e extrair o máximo da informação química relevante, através da análise de dados” (Kowalski, B. R, 1978). Como pode ser ilustrado pela Figura 2, essa ferramenta reúne a interdisciplinaridade entre estatística, química e matemática. A ênfase no controle, análise de sinais e comportamento experimental coloca a quimiometria como uma vertente da química analítica, porém com natureza multidisciplinar (Ferreira, M. M. C., 2015).

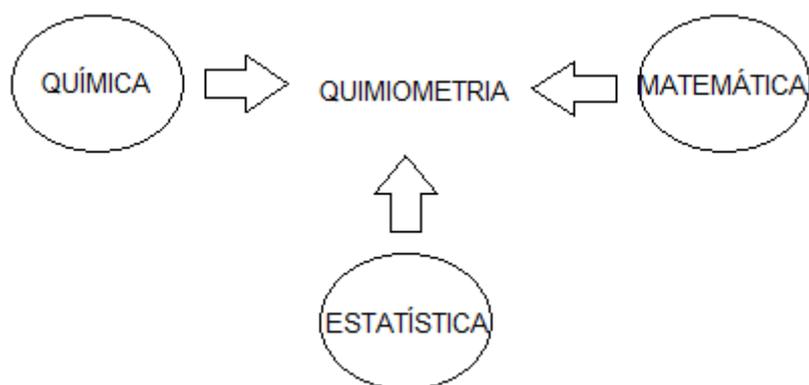


Figura 2 - Interdisciplinaridade da quimiometria. *Fonte: Adaptado de Ferreira, M. M. C.: Quimiometria: conceitos, métodos e aplicações. Editora Unicamp, 2015.*

A Análise de Componentes Principais, do inglês “*Principal Component Analysis*” (PCA) é um método utilizado na quimiometria para reconhecer padrões, ou seja, similaridades a partir de uma tabela de dados. A PCA projeta os dados multivariados em um espaço de dimensão reduzido (Ferreira, M. M. C., 2015). Na área de estudo em questão, este tipo de análise estatística permite detectar tendência nos dados e agrupá-los por semelhança, além de mostrar aspectos subjacentes das correlações entre as variáveis originais.

Tudo isso é feito por meio computacional, graças ao desenvolvimento da informática. O conhecimento obtido por meio da quimiometria pode gerar informações muito úteis para um laboratório. Em suma, a PCA pode facilitar a interpretação de dados, detectar falhas em processos industriais bem como aprimorar o sistema de qualidade em um laboratório (Ferreira, M. M. C, 2015).

2. OBJETIVOS

Tendo em vista a variabilidade existente nos combustíveis comercializados no País, o banco de dados produzido pelas análises de rotina da ANP e a capacidade da ferramenta PCA em evidenciar as possíveis relações entre as amostras em um conjunto de dados, o objetivo do presente trabalho de conclusão de curso é avaliar a existência de algum tipo de padrão entre as amostras, seja ele o local que foi coletada a amostra, a data ou os ensaios realizados. Esse tipo de informação ajuda a melhorar a gestão de um laboratório e neste caso pode reconhecer os ensaios físico-químicos de maior influência para o relatório da qualidade dos combustíveis.

Em síntese, a PCA será utilizada para melhor explorar as relações entre as amostras e os resultados de suas análises, a fim de avaliar a variabilidade entre os combustíveis e as principais informações que podem ser acessadas com base nos dados e resultados disponibilizados pela ANP. Cabe destacar ainda que o tema escolhido para esse TCC possibilitará que sejam obtidos conhecimentos sobre uma das principais ferramentas quimiométricas empregadas na análise de dados químicos, que é uma área de pouco ou nenhum acesso para nas disciplinas regulares dos currículos dos cursos de graduação. Portanto, um dos objetivos também é a oportunidade de introduzir o estudo da quimiometria e saber como aplicá-la ao conhecimento acadêmico, desde a formatação, pré-processamento até a análise exploratória dos dados.

3. REVISÃO BIBLIOGRÁFICA

3.1 Combustíveis

Do ponto de vista da química, autores como Dionysio e Meirelles, descrevem combustíveis como substâncias que, quando reagem com o oxigênio, liberam energia na forma de calor. Atualmente são classificados como fósseis ou renováveis. Os fósseis são aqueles obtidos de fontes não renováveis, que se originaram há milhares de anos por meio da fossilização lenta e complexa de animais e plantas. Exemplos: petróleo, gás natural e carvão mineral. Já os renováveis são obtidos a partir de fontes naturais que estão constantemente se renovando no meio ambiente. Exemplos: etanol, biodiesel e madeira (DIONYSIO; MEIRELLES, 2016).

O carvão mineral teve grande importância na primeira revolução industrial, ainda no século XVIII. Logo depois, o petróleo se destacou como a maior fonte energética mundial. Porém, seu uso tem sido substituído no mercado devido ao fato de ser um recurso esgotável e, quando queimado, liberar gases que são nocivos à atmosfera. Por razões econômicas, sociais e ambientais, hoje vivemos a era do incentivo ao uso de biocombustíveis, como o etanol e o biodiesel (OLIVEIRA, 2008).

Os combustíveis automotivos são aqueles utilizados para abastecimento em veículos: carros, motos e caminhões. A gasolina e o álcool são os combustíveis mais utilizados em veículos leves, já o diesel é o mais comercializado no mercado brasileiro e utilizado em embarcações, indústrias e veículos pesados. Quando os carros são abastecidos somente a álcool, eles utilizam o álcool etílico hidratado. Já o que é adicionado à gasolina é o álcool etílico anidro (PORTAL DA PETROBRÁS, 2019).

A gasolina comercializada no Brasil pode ser de 2 tipos: tipo A ou C. As gasolinas tipo A são obtidas nas refinarias, sem etanol e as gasolinas tipo C, conhecidas como comum, possuem o etanol anidro adicionado e são as vendidas em postos de revenda de combustíveis. A composição química das gasolinas é basicamente dada por cadeias normais e ramificadas de hidrocarbonetos, compostos oxigenados e contaminantes tóxicos, dentre eles o benzeno e enxofre (CARVALHO; FILHO, 2014). Aditivos como detergentes dispersantes também são adicionados ao produto. Para que a gasolina atenda às especificações da ANP, ela deve ser formulada a partir de frações (correntes) do refino do petróleo de forma a atender as especificações da ANP (PORTAL DA ANP, 2019).

Já o álcool biocombustível é composto pela molécula de etanol, C_2H_5OH , e pequenas quantidades de água e outros contaminantes, podendo ser utilizado em veículos do tipo *Flex Fuel*, ou seja, em motores que funcionam com a mistura de 2 combustíveis em qualquer proporção (GONÇALVES et al., 2013). Os veículos abastecidos apenas com etanol tiveram grande importância nos anos 70, quando foi lançado o Programa Brasileiro do Alcool (PróAlcool) e o Brasil configurava-se como um grande produtor de cana-de-açúcar. Na época, o petróleo passava por crises e o preço do barril obrigava os brasileiros a procurarem alternativas para abastecer seus veículos (LEITE et al., 2009). O etanol é comercializado exclusivamente na forma de etanol hidratado ou misturado à gasolina. (PORTAL DA ANP, 2019).

O diesel é um óleo de origem petrolífera e é uma mistura de hidrocarbonetos, além de conter compostos de enxofre, nitrogênio e oxigênio em menores quantidades. Esses compostos são removidos (não completamente) nas etapas de craqueamento catalítico, hidrocrackeamento e/ou hidrotreatamento do petróleo, dependendo do esquema de refino. Assim como a gasolina, existem 2 tipos: A e B. O tipo A é proveniente das refinarias e não possui o biodiesel adicionado. Já o B é comercializado em postos de revenda de combustíveis com a adição obrigatória de biodiesel. O óleo diesel possui diferentes concentrações de enxofre, as quais dão a classificação ao combustível. No Brasil, após a resolução da ANP nº 50 de 2013, os veículos só podem ser abastecidos com diesel S10, que possui o teor máximo de enxofre de 10mg/kg ou S500, que possui o teor máximo de 500mg/kg (GARCIA, 2013).

3.2 Legislação da Agência Nacional do Petróleo

A ANP foi criada em 06 de agosto de 1997, sob a Lei Nº 9.478. O artigo 177 da Constituição Federal Brasileira estabelece que o petróleo, o gás natural e seus derivados são monopólio da União. Desta forma, cabe à União fiscalizar e regulamentar as atividades relacionadas a exploração, produção e comercialização desses produtos e garantir a qualidade deles.

Vinculada ao Ministério de Minas e Energia (MME), a ANP é responsável por estabelecer as resoluções, normas, portarias e decretos para cada combustível comercializado nos postos. Os laboratórios contratados pela ANP, escolhidos por licitação, realizam análises exigidas no escopo do controle de qualidade dos combustíveis, a fim de garantir a segurança do consumidor. Comparações interlaboratoriais verificam regularmente o desempenho desses laboratórios contratados em todo o País. As análises

feitas no DF seguem os requisitos exigidos pela Norma NBR ISO/IEC 17025, que preza pela qualidade dos ensaios (PORTAL DA ANP, 2019). Os analistas podem ser especialistas em regulamentação, técnicos ou estagiários em química.

Além disso, a Agência mantém compromisso com a saúde humana e do meio ambiente, pois seus regulamentos e ações de fiscalização garantem o atendimento às novas exigências ambientais e para o bom desempenho dos motores. Assim, os produtos regulados necessitam ter as características de qualidade revisadas e atualizadas constantemente (PORTAL DA ANP, 2019).

3.2.1 Programa de Monitoramento de Qualidade de Combustíveis

Em 14 de janeiro de 1998 foi criado o Programa de Monitoramento de Qualidade de Combustíveis (PMQC) um programa nacional com o intuito de reduzir os índices de não conformidade dos produtos regulados. O programa fornece um panorama da qualidade dos combustíveis e consiste na coleta, transporte e análises físico-químicas dos combustíveis líquidos automotivos. As análises realizadas são consolidadas em boletins mensais disponibilizados em seu portal online, os quais permitem o planejamento de ações fiscais mais eficientes, tanto pela ANP quanto por outros órgãos que firmam convênios com a Agência, como Ministérios Públicos, Procons e Secretarias de Fazenda (PORTAL DA ANP, 2019).

No DF, o laboratório da ANP é o Centro de Pesquisas e Análises Tecnológicas (CPT) e monitora constantemente a conformidade da gasolina tipo C, do etanol hidratado e do óleo diesel tipo B comercializados no Distrito Federal e, eventualmente, em estados vizinhos. Atualmente, o PMQC é regulado pela resolução N° 8, de 9.2.2011 e estabelece que a coleta e o transporte das amostras são responsabilidade dos laboratórios contratados pela ANP, os pontos de coleta são escolhidos por sorteio e as datas têm caráter sigiloso. A cada visita feita a um revendedor varejista de combustível, é coletado, no mínimo, uma amostra de cada produto. Cabe mencionar que o PMQC não tem caráter fiscalizatório e não gera autuação em caso de não conformidade (PORTAL DA ANP, 2019).

É importante salientar que o laboratório localizado em Brasília é acreditado pela Coordenação Geral de Acreditação do Instituto Nacional de Metrologia, Qualidade e Tecnologia (INMETRO), ou seja, é atestado que o CPT atende à Norma Brasileira NBR ISO/IEC 17025. A acreditação reconhece o laboratório como competente desde o “recebimento de amostras até a emissão de resultados, passando pela capacitação dos

analistas e a calibração de equipamentos analíticos” (PORTAL DA ANP, 2019). Atualmente, o centro de pesquisas é responsável por monitorar 2 unidades federativas: Distrito Federal (DF) e Tocantins (TO).

Os boletins mensais de monitoramento de qualidade de combustíveis trazem dados nacionais por região e por estado. São eles: quantidade de amostras coletadas, ensaios realizados e o tipo de não conformidade. Dados brutos de cada amostra também são disponibilizados. Segundo o site da ANP, desde 1998 a quantidade de amostras conformes à especificação tem aumentado, comprovando a eficácia do Programa. Em fevereiro de 2019, por exemplo, as amostras de gasolina apresentaram índices de conformidade a partir de 98%, as de etanol hidratado, de 96% e as de óleo diesel, de 92%. Esses valores indicam um elevado padrão de conformidade dos combustíveis analisados (PORTAL DA ANP, 2019). A amostra é dada como não conforme quando ela apresenta, no mínimo, um resultado de ensaio fora da especificação estabelecida pela ANP. Na Figura 3, encontra-se o modelo de funcionamento do PMQC, o primeiro passo é a coleta das amostras combustíveis, seguido pelo envio das amostras para análise em laboratório próprio (CPT) ou em contratados, que podem ser instituições de ensino ou pesquisa que venceram licitações. Os resultados das análises seguem para a avaliação da ANP, para serem divulgados mensalmente em boletins (SÉRIES TEMÁTICAS ANP N°5, 2012).

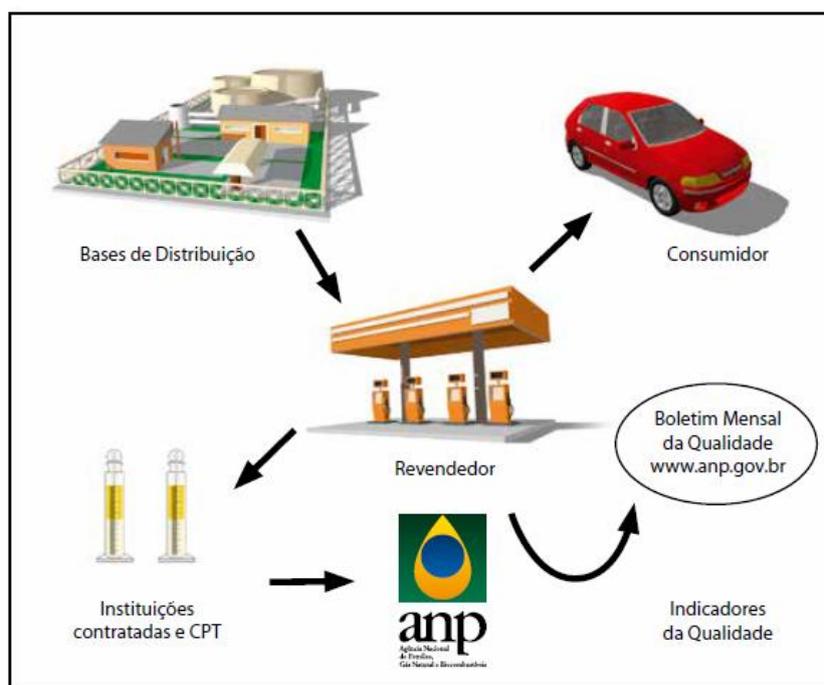


Figura 3 - Modelo de atuação do PMQC. Fonte: Séries temáticas ANP N°5, 2012.

3.2.2 Ensaio realizado na gasolina, etanol e óleo diesel no PMQC

As características avaliadas no PMQC dão nome aos ensaios realizados na gasolina tipo C, no etanol hidratado combustível e óleo diesel tipo B. Todos os testes feitos são pautados nas versões mais atualizadas dos métodos oficiais contidos nas Normas Brasileiras de Regulamentação publicadas pela Agência Brasileira de Normas Técnicas (ABNT), além de órgãos internacionais como: *American Society of the International Association for Testing and Materials* (ASTM), *International Organization for Standardization* (ISO) ou *Comité Européen de Normalisation* (CEN) (CARVALHO; FILHO, 2014). Nos Quadros 1, 2 e 3 podem ser observadas as características, o método utilizado para a realização do ensaio, descrição ou importância de cada uma. Todas as características físico-químicas são avaliadas como conformes ou não, de acordo com um parâmetro, ou seja, uma especificação legal que estabelece se o resultado está dentro ou fora dos limites estabelecidos para cada teste feito.

Quadro 1 - Ensaios realizados na gasolina. *Fonte: Séries temáticas ANP N°5, 2012.*

CARACTERÍSTICA	MÉTODO		DESCRIÇÃO/IMPORTÂNCIA
	ABNT	ASTM	
Aspecto, visual	-	-	Indicação visual da qualidade e de possíveis contaminações.
Cor, visual	-	-	Indicação da tonalidade da cor característica da gasolina.
Massa específica a 20 °C	NBR 7148 NBR 14065	D4052 D1298	Relação entre a massa de combustível e o seu volume. Quanto maior a massa específica, maior o potencial energético, uma vez que maior será a massa de combustível para um mesmo volume considerado.
Teor de etanol anidro combustível	NBR 13992	-	Verificação do atendimento ao dispositivo legal que delimita teores mínimo e máximo de etanol anidro na gasolina. A gasolina com valores inferiores ou superiores de etanol anidro combustível aos especificados pode comprometer o bom funcionamento dos veículos.
Destilação	NBR 9619	D86	Avaliação da volatilidade da gasolina. Os diversos pontos de destilação especificados visam atender as diferentes condições de operação dos motores.
Infravermelho [MON, RON, IAD, teor de benzeno (ASTM D6277) e composição (olefinas, saturados e aromáticos)]	Espectrometria por infravermelho		Indicação da concentração de compostos aromáticos, hidrocarbonetos com dupla ligação carbono-carbono (olefinas) e saturados (parafinas e naftênicos). Importante para o controle de emissões de compostos tóxicos, poluentes e que causam danos ao sistema de combustível e câmara de combustão.
Marcador de solvente	Conforme metodologia de identificação do fornecedor do marcador		Avaliação que verifica se o combustível foi adulterado por adição ilícita de solvente.

Quadro 2 - Ensaio realizados no óleo diesel. *Fonte: Adaptado de Séries temáticas ANP N°5, 2012.*

CARACTERÍSTICA	MÉTODO		DESCRIÇÃO/IMPORTÂNCIA
	ABNT	ASTM	
Aspecto, visual	-	-	Indicação visual da qualidade e de possíveis contaminações.
Cor, visual	-	-	Indicação da tonalidade da cor característica do óleo diesel.
Cor ASTM	NBR 14483	D1500	Avaliação da cor característica do óleo diesel. Cor ASTM fora de especificação pode ser um indicativo de degradação ou contaminação.
Massa específica a 20 °C	NBR 7148 NBR 14065	D1298 D4052	Verificação da relação entre a massa de combustível e o seu volume. Valores de massa específica acima do especificado podem aumentar a produção de poluentes, ao passo que valores abaixo do especificado podem favorecer a perda de potência do motor e o aumento do consumo de combustível.
Ponto de fulgor	NBR 7974 NBR 14598 -	D56 D93 D3828	Verificação de temperatura na qual um combustível, no caso o óleo diesel, gera uma quantidade suficiente de vapor para formar uma mistura inflamável com o ar atmosférico, apresentando uma queima abrupta de tais vapores ao entrar em contato com uma fonte externa de calor, sem que, contudo, haja a combustão do combustível. Está relacionado à inflamabilidade do combustível. Pode indicar a contaminação por produtos de menor ponto de ebulição. Importante também como indicativo dos cuidados a serem tomados durante o manuseio, transporte, armazenamento e utilização do combustível.
Enxofre Total	NBR 14875 - NBR 14533 -	D1552 D2622 D4294 D5453	Compostos de enxofre são altamente corrosivos e poluentes, devido à formação de gases tóxicos.
Destilação	NBR 9619	D86	Avaliação da volatilidade para verificar a contaminação por derivados mais leves, como a gasolina, ou mais pesados, como o óleo lubrificante. Um combustível com valores dentro da faixa de destilação em conformidade com a especificação contribui para o bom desempenho do motor, para minimizar a formação de depósitos no sistema de combustão e no controle das emissões de gases tóxicos e poluentes.
Índice de Cetano	NBR 14759	D4737	Avaliação da qualidade de ignição do combustível, influenciando diretamente na partida do motor. O índice de cetano correlaciona o desempenho do óleo diesel com o desempenho do hexadecano – hidrocarboneto parafínico comercializado como cetano. Um índice de cetano alto permite um aquecimento mais rápido no motor, aumentando a eficiência
Teor de biodiesel	NBR 15568	D7861	Determinação do teor de biodiesel em óleo diesel por espectroscopia na região do infravermelho médio. A adição de biodiesel é obrigatória em determinadas proporções

Quadro 3 - Ensaios realizados no etanol hidratado combustível. *Fonte: Séries temáticas ANP N°5, 2012.*

CARACTERÍSTICA	MÉTODO		DESCRIÇÃO/IMPORTÂNCIA
	ABNT	ASTM	
Aspecto, visual	-	-	Indicação visual da qualidade e de possíveis contaminações.
Cor, visual	-	-	Indicação da tonalidade da cor característica do etanol, podendo indicar possíveis alterações ocasionadas por degradação ou contaminação.
Massa específica a 20 °C	NBR 5992	D4052	Verificação da relação entre a massa de combustível e o seu volume. Variações de valores de massa específica para baixo ou para cima do especificado impossibilita, quando injetada no motor, uma mistura de ar/combustível balanceada.
Teor Alcoólico °INPM	NBR 5992	-	Verificação da quantidade em grama de álcool absoluto contido em cem gramas de etanol combustível. O limite inferior garante um poder calorífico mínimo para seu uso como combustível. Etanol com teor alcoólico inferior ao especificado, por adição de água, pode comprometer os componentes do veículo, dificultar a ignição e levar à diminuição do desempenho do motor.
Condutividade Elétrica	NBR 10547	D1125	Verificação quantitativa do caráter elétrico do combustível. A limitação máxima de condutividade elétrica permite controlar a concentração de contaminantes condutores de corrente elétrica (eletrólitos) no etanol. Combustível com uma condutividade elétrica superior ao permitido pode comprometer a vida útil das partes mecânicas do veículo.
Potencial Hidrogeniônico (pH)	NBR 10891	-	Verificação da acidez, neutralidade e basicidade. Um pH com valores para baixo e para cima do especificado para o etanol combustível pode comprometer partes do veículo por onde haja o fluxo do combustível.
Teor de hidrocarbonetos	NBR 13993	-	É o parâmetro indicativo de presença de contaminantes como gasolina.

3.3 Quimiometria

A quimiometria surgiu com a necessidade de se obter informações químicas a partir de dados multivariados que começaram a ser gerados nos equipamentos modernos disponíveis em um laboratório químico. Em 1971 a expressão “quimiometria” foi utilizada pela primeira vez e a Sociedade de Quimiometria foi criada em 1974, nos Estados Unidos

da América (Ferreira, M. M. C., 2015). Desde então essa nova área da química se concretizou no meio científico, com vários jornais e publicações dedicadas à essa importante disciplina.

O avanço ocorrido nos sistemas computacionais permite que softwares consigam analisar dados por meio de análise exploratória, construir modelos qualitativos, de classificação, ou quantitativos, por meio de análise multivariada (Ferreira, M. M. C., 2015). Contudo, o uso correto desses métodos requer um conhecimento mínimo do formalismo matemático envolvido e do conhecimento químico para interpretar os resultados desses modelos.

Para garantir a confiabilidade e correta interpretação das análises, é necessário escolher as melhores condições para processar, ou seja, analisar os dados. No escopo deste trabalho será utilizada a PCA, que possibilita a análise exploratória dos dados.

3.3.1 Análise de componentes principais (PCA)

A PCA realiza a decomposição de dados e os organiza na forma de uma matriz, utilizando os princípios da álgebra linear. É classificada com um método multivariado, pois analisa mais de uma variável simultaneamente. Nas matrizes, as amostras geralmente ficam dispostas em linhas e as variáveis em colunas. Por exemplo, ao realizar um experimento em um espectrofotômetro que opera na região do infravermelho, o resultado obtido é um espectro de sinal do detector *versus* número de onda em cm^{-1} . Neste estudo, o número de onda é a variável. A técnica é utilizada para facilitar a identificação das variáveis mais importantes para a identificação de uma substância ou observar determinado padrão de distribuição das amostras (Ferreira, M. M. C., 2015).

Geometricamente, a técnica realiza uma projeção dos dados em um novo espaço, pois os dados multivariados obtidos são projetados em um espaço de dimensão menor, definido pelas componentes principais, na qual se espera que esteja contida as informações relevantes para a análise que se pretende realizar (Ferreira, M. M. C., 2015). O método proporciona uma maior facilidade para o reconhecimento de padrões, pois facilita a visualização da tendência presente nos dados em poucas dimensões. A matriz que contém os dados originais em várias dimensões geralmente é representada como \mathbf{X} . Na decomposição em PCA da matriz \mathbf{X} são determinadas novas variáveis que mais expliquem a variância (informação) de um dado, elas são chamadas de componentes principais, ou PC's do inglês *Principal Components* e formam um novo sistema de eixos. O computador

faz isso utilizando cálculos de combinação lineares das variáveis originais (CARVALHO; FILHO, 2014).

As componentes principais são determinadas de forma a serem perpendiculares. Elas são definidas e colocadas em um espaço seguindo a direção que descreva a máxima variância dos dados originais. Uma PC é definida por um par de vetores pesos e vetores escores. Esses vetores dão origem a pontos em gráficos de escores (**T**) e pesos (**P**) (Ferreira, M. M. C., 2015). A técnica computacional desenha um novo sistema de eixos (variáveis), que em algumas situações pode se restringir a um plano bidimensional. As dimensões dependerão do número de PCs escolhidas.

Para compactar os dados, a matriz de dados **X** é decomposta em três termos, matriz de escores **T**, matriz de pesos **P^T** e de resíduos **E**. A matriz **T** contém os escores, que são as coordenadas (valores) das amostras no sistema de eixos formados pelas PCs, mostra a relação entre as amostras. A matriz **P^T** contém os pesos, os quais são os ângulos entre vetor definido pela PC e as variáveis originais, os quais mostram as relações entre as variáveis medidas e as PCs (TRYGG, 2003). A matriz **E** é a matriz dos resíduos, nela é possível observar resultados que não foram explicados no modelo PCA criado.

Neste caso, a estatística realizada com os parâmetros **Q** e **T²** pode ser utilizada para evidenciar esses resultados considerados atípicos. Em resumo, esses dois parâmetros utilizam um limite de confiança para comparar quais amostras diferem do modelo de PCA elaborado e os projetam em um gráfico de resíduos. As amostras que apresentam valores fora do limite para os parâmetros estatísticos **Q** e **T²** de *Hottelling* são consideradas amostras anômalas. O parâmetro **Q** é a soma quadrática dos resíduos **E** do modelo, que caracteriza as amostras que não foram bem ajustadas pelo modelo de PCA descrito. Já o parâmetro **T²** de Hotelling mostra a variação de cada amostra dentro do modelo PCA (FERREIRA et al., 2013). Um esquema da decomposição em PCA pode ser observado na Figura 7.

$$\boxed{\mathbf{X}} = \boxed{\mathbf{T}} \boxed{\mathbf{P}^T} + \boxed{\mathbf{E}}$$

Figura 4 - Estrutura do modelo PCA. *Fonte: adaptado de (BRO; Rasmus, 2014).*

Os dados originais da matriz **X** podem passar por diversos pré-processamentos antes da análise por PCA. O autoescalamento é um recurso que faz com que as variáveis

tenham o mesmo peso, sendo geralmente empregado quando as variáveis têm unidades ou características diferentes. Nesse pré-processamento o software subtrai de cada coluna dos dados sua própria média e os divide pelo seu desvio padrão (CARVALHO; FILHO, 2014). Em análise exploratória, espera-se que o gráfico de escores de PCA revele agrupamentos por semelhança, grupos comuns. O gráfico de pesos permite entender quais variáveis mais contribuem para os agrupamentos observados. O método chamado de não supervisionado, pois a separação dos agrupamentos acontece sem a necessidade de conhecimento, *a priori*, sobre como as amostras se relacionam entre si (DE SOUSA, 2012).

4. MATERIAIS E MÉTODOS

4.1 Conjunto de dados

Todas as amostras que fizeram parte desse trabalho foram analisadas seguindo métodos oficiais para a realização dos ensaios, descritos nos Quadros 1, 2 e 3. Elas foram coletadas durante a rotina do PMQC. As amostras no DF foram coletadas em diferentes bairros, todas no município de Brasília. Já no TO diferentes bairros de diferentes municípios foram selecionados, sempre através dos sorteios. Ao final das análises das amostras, os resultados dos ensaios foram lançados em um sistema eletrônico, o LIMS (*Laboratory Information Management System*) pela ANP. Esses dados são avaliados e consolidados como os dados da qualidade de combustíveis no País. Após esse tratamento realizado pela própria Agência, esses dados foram divulgados no Boletim de Monitoramento (<http://www.anp.gov.br/dados-abertos-anp>) em planilhas que contêm data, local e bandeira da amostra por Unidade Federativa (UF) junto com os resultados de seus respectivos ensaios.

As amostras que tiveram seus dados analisados por PCA foram coletadas pelo CPT no período de junho de 2016 até setembro de 2017. Os dados obtidos por UF nas planilhas de Excel disponibilizadas pela ANP foram reorganizados por local, data, número da amostra e seus ensaios por combustível. Com o auxílio do Software Matlab, foram analisadas 1322 amostras de óleo diesel e 1069 amostras de gasolina, utilizando a técnica de PCA. O etanol combustível não foi objeto de estudo, pois poucos dos ensaios realizados (Quadro 3) estavam disponíveis para análise por PCA.

4.2 Estrutura e análise de dados

Todos os dados que foram utilizados para a realização desse trabalho foram obtidos nas planilhas de Excel disponibilizadas no site da ANP, no endereço <http://www.anp.gov.br/dados-abertos-anp>, e reorganizados a fim de selecionar apenas os ensaios e combustíveis de interesse nos dois estados: DF e TO. Os cálculos foram executados utilizando o software PLS Toolbox 8.5 para uso em ambiente Matlab® 7.5, utilizando a técnica quimiométrica PCA. Cabe destacar que a autora deste trabalho de conclusão de curso foi estagiária na ANP no período de junho de 2016 a dezembro de 2017 e realizou muitos dos ensaios descritos nos Quadros 1, 2 e 3, ao integrar a equipe de análise da agência, cujas responsabilidades incluem a análise das amostras de monitoramento e fiscalização dos combustíveis diesel, gasolina e etanol. Portanto, o

presente trabalho não consistiu apenas em uma análise de dados de uma base pública de resultados de análises.

Os resultados das análises foram organizados no formato de matrizes tanto para óleo diesel como para a gasolina. As variáveis químicas selecionadas para estudo foram aquelas com a maior quantidade de dados disponíveis durante o período de junho de 2016 a setembro de 2017, ou seja, foram a maioria dos ensaios físico-químicos que estavam no escopo das análises da ANP nos meses de referência. As matrizes de dados foram autoescaladas e analisadas por PCA.

4.3 Variáveis de estudo

As variáveis de estudo expostas nos gráficos deste trabalho foram escolhidas de acordo com a sua disponibilidade nos dados. Para cada combustível foram escolhidos os ensaios que mais continham resultados, ou seja, valores numéricos que possibilitassem a análise por PCA. Caso estivessem faltando resultados de análises de alguma variável para um número significativo de amostras, essa variável era excluída da matriz de dados **X**. Em alguns casos cujos resultados dos ensaios eram faltantes na matriz original de dados, ou seja, que não foram disponibilizados nas tabelas de Excel que continham os dados da ANP, os campos vazios foram substituídos por NaN (acrônimo em inglês para *Not a Number*). NaN é um valor especial que no Matlab corresponde a “não é um número”, sendo que para o cálculo das componentes principais o algoritmo substitui os dados faltantes por um ajuste matemático baseado nos valores das outras amostras.

4.3.1 Variáveis para óleo diesel

As variáveis químicas analisadas para este combustível foram os seguintes ensaios: cor, as temperaturas quando 10 e 50% do volume de diesel foi evaporado durante a destilação das amostras, massa específica a 20 ° C, ponto de fulgor e teor de biodiesel. A Tabela 1 contém um resumo das especificações e as variáveis químicas que foram escolhidas, elas correspondem aos ensaios avaliados segundo os métodos adotados pela ANP (Quadro 2). No quesito cor, foi preciso atribuir valores numéricos para cada uma, já que o Matlab só realiza os cálculos com informações numéricas e as cores estavam escritas por extenso. As cores descritas na Tabela 2 foram aquelas observadas para os 2 combustíveis: diesel e gasolina.

Tabela 1 – Ensaios utilizados como variáveis químicas originais para o óleo diesel e suas especificações.

Ensaio	Especificações*
Cor	Vermelho para diesel S500
Destilação 10% (Destil-10%)	Mínimo 180 ° (S10)
Destilação 50% (Destil-50%)	245 ° C a 295 ° C (S10) 245° C a 310 ° C (S500)
Massa específica (Massa)	815 a 853 kg/m ³ (S10) 815 a 865 (S500)
Ponto de fulgor (PF)	Mínimo 38 ° C
Teor de Biodiesel (Teor Bio)	6,5 a 8,0% volume

(*) As especificações mencionadas para cada ensaio eram as vigentes na época em que as amostras foram analisadas, que compreende o período de junho de 2016 a setembro de 2017. *Fonte: ANP, 2013.*

Tabela 2 – Escala numérica de cores das amostras e correspondentes numéricos estabelecidos nesse trabalho para análise no Matlab.

Correspondente numérico	Cor observada no ensaio
1	Castanho
2	Amarelo
3	Laranja
4	Verde
5	Azul
6	Vermelho

4.3.2 Variáveis para gasolina

As variáveis químicas analisadas para a gasolina foram os seguintes ensaios: cor, ponto final de ebulição da destilação, resíduo da destilação, as temperaturas quando 10, 50% e 90% do volume de gasolina foi evaporado durante a destilação das amostras, massa específica a 20 ° C e teor de etanol anidro combustível em gasolina. Todos os ensaios

foram avaliados seguindo as especificações e métodos adotados pela ANP (Quadro 1). Para cor também foi feita a conversão numérica.

Tabela 3 – Ensaios utilizados como variáveis químicas originais para a gasolina e suas especificações.

Ensaio	Especificações*
Cor	Exceto azul
Ponto final de ebulição (Destil-PFE)	Máximo 215 °C
Resíduo da destilação (Destil-Res)	Máximo 2,0 mL
Destilação 10% (Destil-10%)	Máximo 65 °C
Destilação 50% (Destil-50%)	Máximo 80 °C
Destilação 90% (Destil-90%)	Máximo 190 °C
Massa específica (Massa)	Anotar o resultado
Teor de etanol anidro combustível (Teor)	26 a 28 % volume 24 a 26 % volume

(*) As especificações mencionadas para cada ensaio eram as vigentes na época em que as amostras foram analisadas, que compreende o período de junho de 2016 a setembro de 2017. *Fonte: ANP, 2013.*

5. RESULTADOS E DISCUSSÃO

Primeiramente, o pré-processamento dos dados da matriz \mathbf{X} foi realizado para melhorar os resultados da análise multivariada. O autoescalamento das matrizes foi realizado já no Software Matlab®. Como resultado, os índices numéricos nos eixos dos gráficos seguintes são adimensionais, pois a média de cada coluna foi subtraída dos dados e então dividiu-se pelo seu desvio padrão. Essa transformação foi realizada no conjunto original dos dados a fim de que cada variável apresentasse média zero e variância igual a um, para que as influências das variáveis não fossem afetadas pelas unidades dos ensaios e tivessem o mesmo peso.

Após o pré-processamento dos dados, a análise exploratória foi realizada com as amostras de óleo diesel e gasolina. Os dados multivariados obtidos em planilhas foram projetados em um espaço de dimensão menor sem perder as informações relevantes. A matriz original \mathbf{X} foi decomposta em duas menores: a de escores (\mathbf{T}) e a de pesos (\mathbf{P}^t). Amostras que não foram bem ajustadas com os modelos de PCA e que apresentaram comportamento diferente de suas similares compõe a matriz de resíduos \mathbf{E} . Para destacar as amostras com comportamento atípico, foram projetados os parâmetros Q e T^2 com limites de 95% de confiança.

As amostras tiveram suas coordenadas projetadas em gráficos bivariados (*biplots*), que é um sistema de eixos formado pelas componentes principais (novas variáveis) traçadas ortogonalmente. As PCs escolhidas são as que mais explicam a variância dos dados. Dessa forma, foi possível observar tendências e padrões de acordo com os valores próximos ou distantes de escores. Para explicar quais são as variáveis originais químicas que mais influenciaram no comportamento das amostras, foram plotados os gráficos de pesos, neles são observadas as correspondências com as mesmas componentes principais selecionadas no gráfico de escores. Com o autoescalamento utilizado no pré-processamento, as variáveis originais tiveram a mesma influência no cálculo das novas variáveis, as PCs, isso foi necessário já que os resultados continham diferentes unidades, por exemplo ° C e g/cm³, além da grande variação nos dados devido à diferente composição química dos tipos combustíveis.

No gráfico de pesos, as PCs ocupam os mesmos eixos dos gráficos de escores, porém, as coordenadas projetadas são as variáveis originais. Os pesos apresentam índices entre -1 e 1, sendo aqueles com valores próximos de ± 1 as de maior peso e

consequentemente maior contribuição. Aquelas variáveis que tiveram seus valores próximos de 0 tiveram pouca ou nenhuma influência na PC. O gráfico de pesos é diretamente proporcional ao de escores, ou seja, valores positivos de pesos para uma determinada componente principal indica que o teor da variável em questão aumenta com o valor de escore no mesmo fator. Isso também ocorre para os pesos negativos.

5.1 Modelagem por PCA para as amostras de óleo diesel

As amostras de diesel foram modeladas por PCA com 4 componentes capturando 91,42% da variância total dos dados (Tabela 4). A variância explicada de uma PC é a fração da variância global que ela explica no modelo de PCA. A variância total explicada é a soma das variâncias explicadas individualmente por cada PC, ela mostra como a informação química dispersa nas variáveis originais pode ser representada em um número menor de variáveis, as componentes principais (PCs).

Tabela 4 – Percentual de variância explicada por cada componente principal para o modelo PCA construído para amostras de diesel

Componente Principal	% Variância explicada	% Variância total explicada
1	46,80	46,80
2	19,00	65,79
3	19,94	80,73
4	10,69	91,42

Figura 5 – Gráfico dos parâmetros T^2 e Q para as amostras de diesel em função do tipo e do estado onde a amostra foi coletada.

As amostras com altos valores de Q e T^2 tiveram variações não esperadas nas variáveis originais (Tabela 1). Em sua maioria, foram amostras com baixas temperaturas no volume equivalente a 10% da destilação. Apesar do comportamento atípico, comprovado em vista dos altos índices de Q e T^2 , esses resultados (Figura 5) não tiveram relação direta com não conformidade das amostras. Foi observado que a grande maioria dessas amostras eram provenientes do estado de Tocantins. Uma hipótese para essas

amostras é a grande variação na temperatura das amostras se comparada às outras amostras de óleo diesel S500.

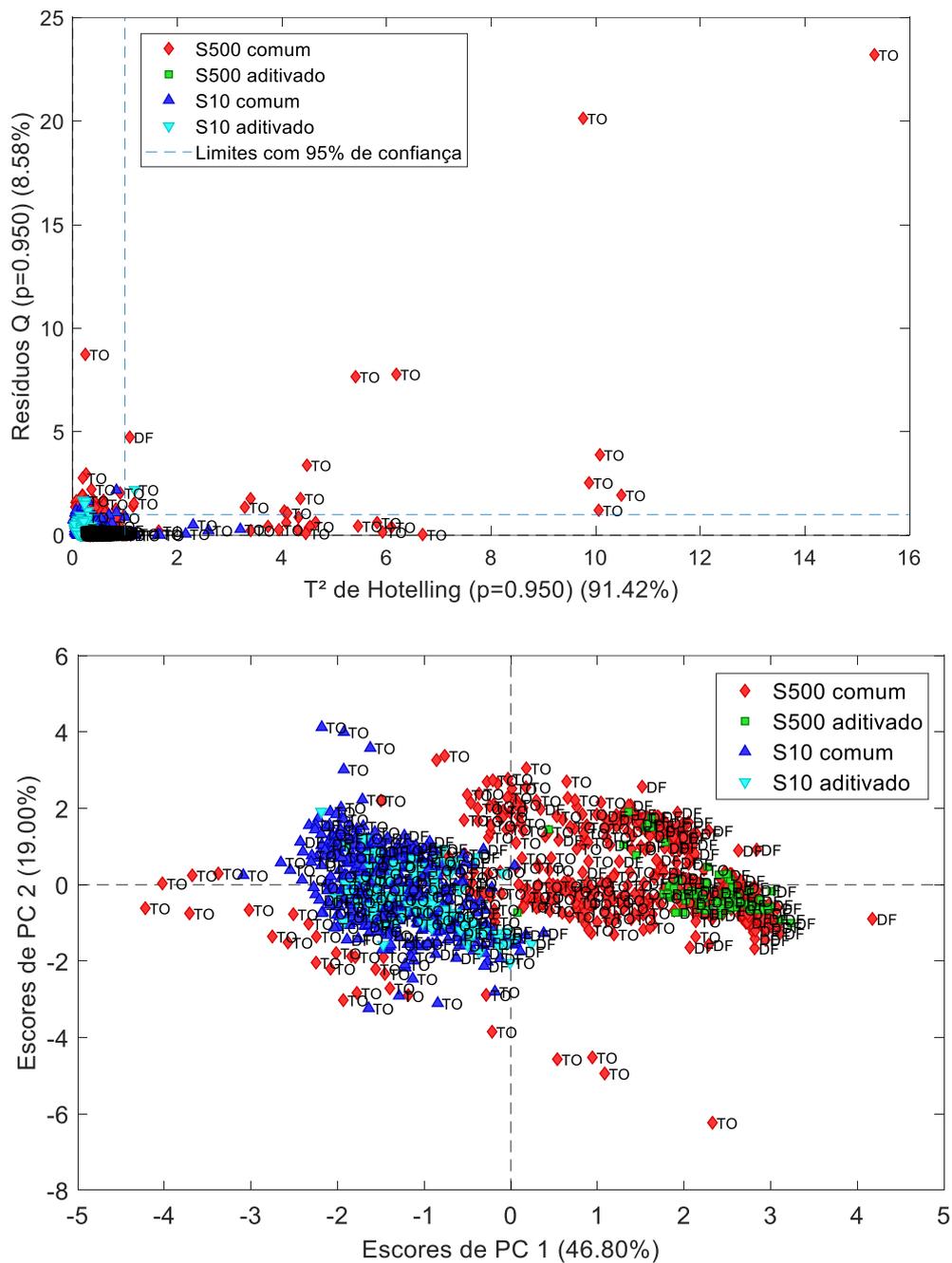


Figura 6 - Distribuição dos escores de PC1 x PC2 para amostras de diesel em função do tipo e do estado onde a amostra foi coletada.

Ao analisar o gráfico de escores de PC1 x PC2 (Figura 6) observou-se que o diesel S500 e S10 se separam relativamente bem, com exceção de algumas amostras de S500 de Tocantins. A legenda separa as amostras de diesel por S500 comum ou aditivado e S10 comum ou aditivado. As sobreposições no gráfico das amostras comuns e aditivadas de um

mesmo tipo de diesel (S10 ou S500) confirmam a similaridade na composição química e que os resultados das análises não sofriam variação em função da presença dos aditivos.

Para explicar quais são as variáveis originais químicas que mais influenciaram no comportamento do gráfico de escores, foi plotado o gráfico de pesos (Figura 7). Nele é observada a correspondência com as mesmas componentes principais selecionadas no gráfico de escores. Com a análise conjunta dos gráficos de escores, o uso das componentes principais permite avaliar a influência das 6 variáveis originais na organização das amostras em cada gráfico. Logo, o conjunto de dados foi reduzido de 6 (variáveis originais) para 4 (número de componentes selecionados no modelo). Com o gráfico PC1 x PC2 já foi possível extrair padrões relevantes.

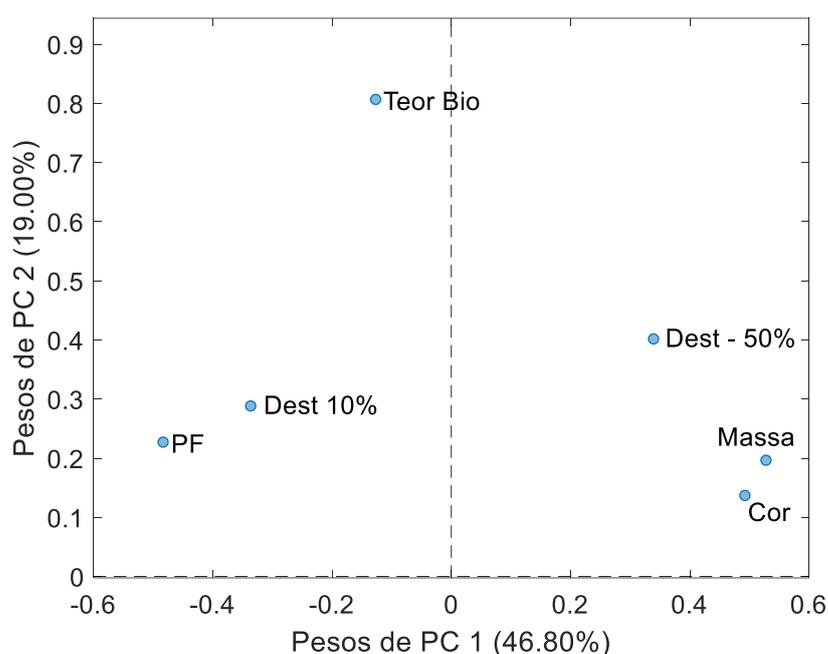


Figura 7 - Distribuição dos pesos de PC1 x PC2 para amostras de diesel em função do tipo e do estado onde a amostra foi coletada.

Ao analisar o gráfico de escores e pesos do diesel para PC1 x PC2 (Figuras 6 e 7), é possível observar 2 agrupamentos maiores distintos e separados. Nos escores negativos de PC1 têm-se as amostras de diesel S10 e nos escores positivos têm-se as amostras de diesel S500. Os pesos positivos de PC1 mostram que as variáveis cor, temperatura da destilação correspondente ao volume de 50% evaporado e massa específica são as principais responsáveis por essa distinção. Quanto à cor, deve-se ao fato de o diesel S10 ser, na maioria dos casos, amarelo e o S500 ser, majoritariamente, vermelho. As cores correspondentes são os números “2” e “6” (Tabela 2). Logo, quanto mais escuro for o

diesel, mais positivo estará em PC1. Já para massa específica, conclui-se que as amostras de diesel S500 analisadas apresentaram massa específica maior que o diesel S10. A especificação do ensaio de massa específica já previa valores maiores para S500 (vide Tabela 1). Notou-se que as variáveis cor e massa específica são muito correlacionadas e capturam, aproximadamente, a mesma informação.

O grupo de amostras de óleo diesel S500 que se encontra disperso na parte negativa de PC2 revelou a grande variação na faixa de temperatura para a destilação correspondente aos 10% do volume evaporado. Algumas dessas amostras são aquelas com altos valores de Q e T² (Figura 5). Já para os pesos positivos de PC2, as amostras têm elevado teor de biodiesel, sendo a variável que mais influenciou no comportamento observado.

Ao analisar o gráfico de escores e pesos de PC3 x PC4 (Figuras 8 e 9), foi possível visualizar um agrupamento de diesel S500 comum mais disperso na parte negativa de PC4 e positiva de PC3. Para PC3, a variável que encontra melhor correlação foi a temperatura da destilação no volume de 10%. A mesma diagonal foi observada em PC1 x PC2, porém a PC3 evidenciou melhor a variação das temperaturas na faixa de 10%, sendo que elas variaram desde 66,4 ° C a 309° C para as amostras localizadas nas extremidades da diagonal. Para PC4, a variável relativa à destilação 50% teve grande influência, bem como o ponto de fulgor (PF). O local de coleta das amostras não influenciou o comportamento das amostras nos gráficos abaixo.

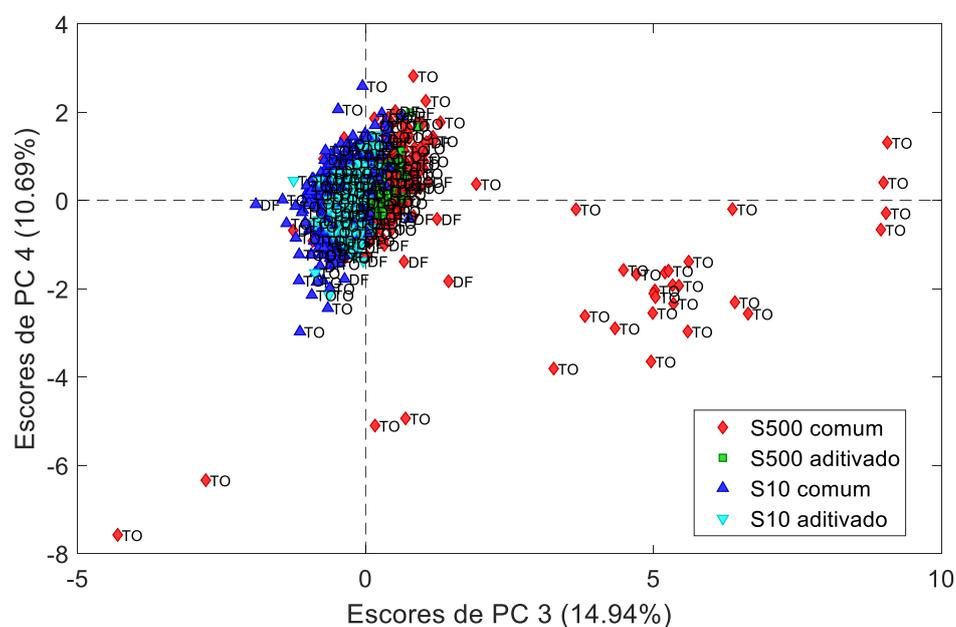


Figura 8 – Distribuição dos escores de PC3 x PC4 para amostras de diesel em função do tipo e do estado onde a amostra foi coletada.

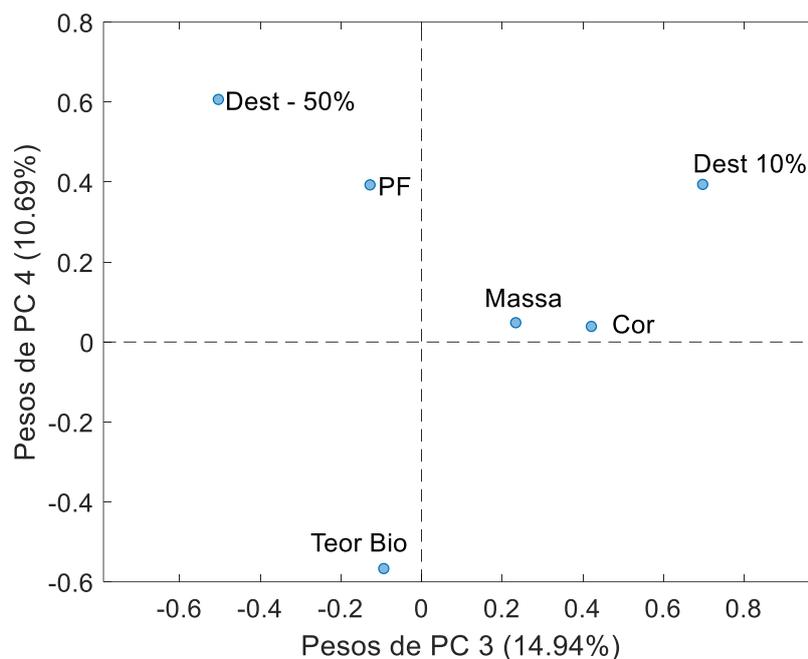


Figura 9 - Distribuição dos pesos de PC3 x PC4 para amostras de diesel em função do tipo e do estado onde a amostra foi coletada.

5.2 Modelagem por PCA para as amostras de gasolina

Para as amostras de gasolina, a modelagem por PCA com 4 componentes capturou 86,52% da variância total dos dados (Tabela 5).

Tabela 5 – Percentual de variância explicada por cada componente principal para o modelo PCA construído para amostras de gasolina.

Componente Principal	% Variância explicada	% Variância total explicada
1	50,90	50,90
2	13,32	64,22
3	12,89	77,12
4	9,40	86,52

As amostras com altos valores de Q e T² tiveram variações não esperadas nas variáveis originais (Tabela 3). Para gasolina, em comparação com diesel, foram observadas poucas amostras com valores de T² e Q superiores aos limites de 95% de confiança. Em sua maioria, as amostras de excediam esses limites foram aquelas que não tinham

disponíveis os resultados para o ensaio de massa específica. Esses dados substituídos por NaN podem ter contribuído para o menor ajuste dessas amostras.

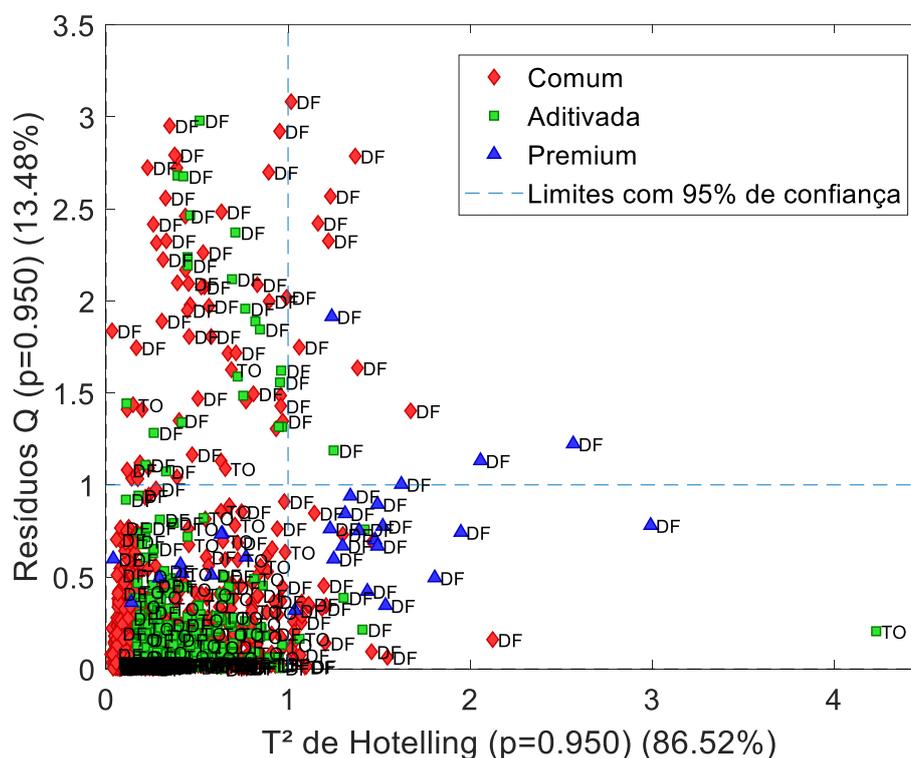


Figura 10 – Gráfico dos parâmetros T² e Q para as amostras de gasolina em função do tipo e do estado onde a amostra foi coletada.

Os gráficos bivariados de escores e pesos foram traçados para as amostras de gasolina. A legenda dos gráficos de escores separa as amostras por tipo de gasolina (comum, aditivada ou Premium) e a UF em que a amostra foi coletada.

Ao analisar o gráfico de escores e pesos de PC1 x PC2 (Figuras 11 e 12), é possível observar 2 agrupamentos separados por PC1. Nos escores negativos de PC1 têm-se as amostras de gasolina, em sua maioria, oriundas do Tocantins e nos escores positivos de PC1 têm-se as amostras de gasolina oriundas dos postos do Distrito Federal. Com o gráfico de pesos foi possível observar que as variáveis massa específica e todos os parâmetros relacionados à destilação foram as principais variáveis responsáveis por essa distinção. Quanto à massa específica e PFE, por exemplo, as amostras do DF apresentaram resultados maiores do que as amostras do TO. Uma possível explicação é o fato de os 2 estados terem seus combustíveis processados em diferentes refinarias, logo, o refino dos produtos derivados de petróleo não é o mesmo e isso influencia nas características físico-químicas do produto final.

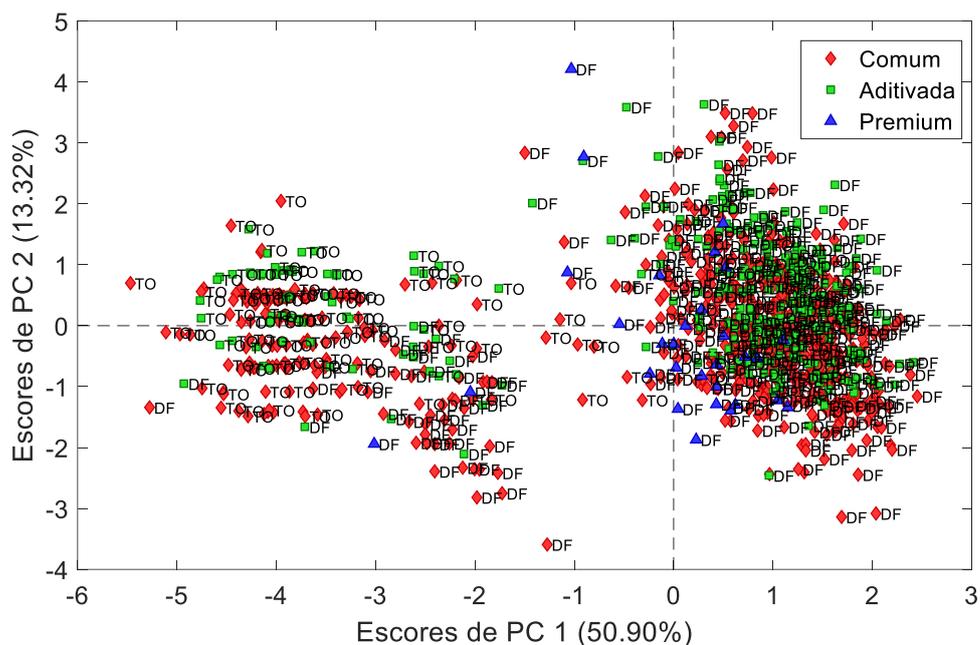


Figura 11 - Distribuição dos escores de PC1 x PC2 para amostras de gasolina em função do tipo e do estado onde a amostra foi coletada.

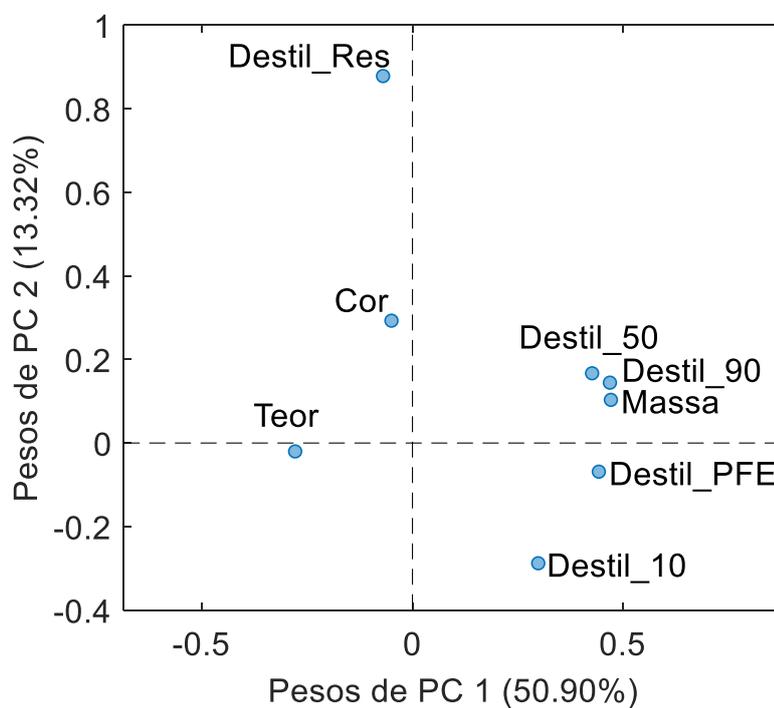


Figura 12 - Distribuição dos pesos de PC1 x PC2 para amostras de gasolina em função do tipo e do estado onde a amostra foi coletada.

As gasolinas Premium estão variando pouco em relação ao eixo de PC1, isso se deve ao fato de todas as amostras desse tipo terem sido coletadas no estado do DF. Para os pesos de PC2, a variável que guardou maior relação foi o resíduo da destilação. Contudo, PC2 não apresentou nenhum padrão visível no gráfico de escores.

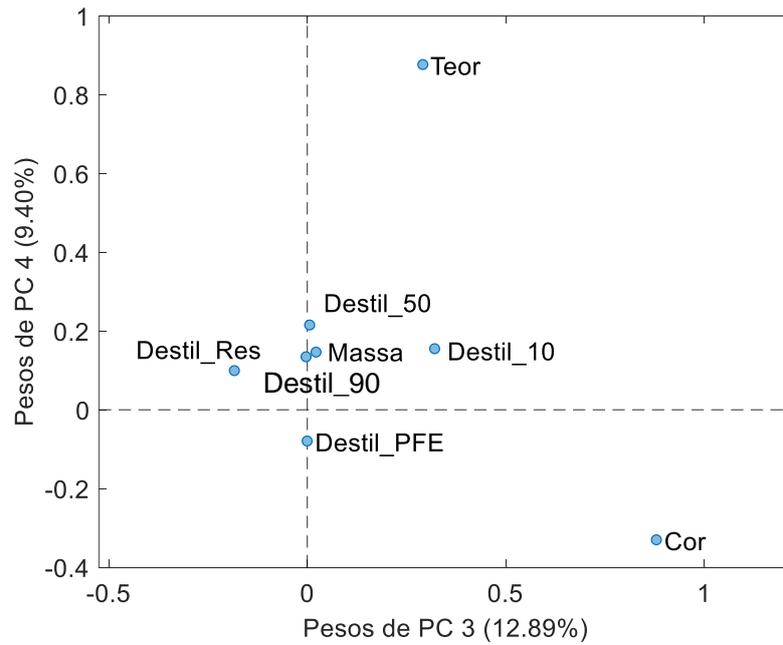


Figura 14 - Distribuição dos pesos de PC3 x PC4 para amostras de gasolina em função do tipo e do estado onde a amostra foi coletada.

Além dos ensaios, também foi possível buscar semelhanças entre as amostras por data, distribuidora, bairro, Unidade Federativa e conformidade em relação à especificação da ANP, em ambos os combustíveis. Não houve relação direta para o diesel, mas no caso da gasolina observou-se um padrão onde a PC1 teve correlação com a separação por estado onde a amostra foi coletada.

6. CONSIDERAÇÕES FINAIS

O modelo de PCA se mostrou eficiente para fazer a análise exploratória dos combustíveis. Desse modo, pelas análises dos gráficos de escores e pesos foi possível observar padrões de similaridade entre as amostras, principalmente por tipo de combustível. Pela análise dos escores das amostras de diesel, observou-se o diferente perfil químico entre as amostras de S10 e S500, já as amostras de gasolina se separam por local de coleta, neste caso a análise exploratória comprovou a existência de diferentes características químicas entre as amostras de gasolina coletadas no DF e TO. O uso do método de PCA permitiu a identificação das variáveis que mais contribuíram com o comportamento dos gráficos, sendo elas os ensaios de massa específica, cor e destilação para ambos os combustíveis. O teor de etanol foi uma variável importante para diferenciar os 3 tipos de gasolina estudados: comum, aditivada e Premium.

Em linhas gerais, a análise de componentes principais se mostrou como uma ferramenta muito útil para avaliar a qualidade de combustíveis de um laboratório, como no caso do CPT. Diante da grande quantidade de amostras analisadas pelo Programa de Monitoramento de Qualidade de Combustíveis, poucas amostras foram consideradas anômalas, consoante com as exigências da ANP, que exige as mesmas especificações para um mesmo combustível independente do estado onde a amostra é coletada. Essas amostras corresponderam aproximadamente a 0,01% para cada combustível, de acordo com o gráfico dos parâmetros Q e T². Além disso, foi possível obter conhecimentos e aplicar uma técnica quimiométrica ao nível de graduação, a PCA foi eficiente para fazer a análise simultânea de uma grande quantidade de dados químicos e apresentou resultados esperados de acordo com as especificações e normas adotadas para cada combustível.

7. REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Boletim do monitoramento de qualidade dos Combustíveis; Disponível em: http://www.anp.gov.br/images/QUALIDADE/BOLETIM/2019_FEV_Boletim_PMQC.pdf. Acessado em março de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Dados abertos. Disponível em: <http://www.anp.gov.br/dados-abertos-anp>. Acessado em março de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Etanol. Disponível em: <http://www.anp.gov.br/producao-de-biocombustiveis/etanol>. Acessado em março de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Gasolina. Disponível em: <http://www.anp.gov.br/petroleo-derivados/155-combustiveis/1855-gasolina>. Acessado em março de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Histórico. Disponível em <http://www.anp.gov.br/qualidade-de-produtos/historico>. Acessado em março de 2019

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Legislação do óleo diesel combustível. Disponível em: <http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2013/dezembro&item=ramp-50--2013>. Acessado em maio de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Legislação da gasolina combustível. <http://legislacao.anp.gov.br/?path=legislacaoanp/resolanp/2013/outubro&item=ramp-40--2013&export=pdf>. Acessado em maio de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Qualidade dos produtos. Disponível em: <http://www.anp.gov.br/qualidade-de-produtos>. Acessado em março de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); Refino. Disponível em: <http://www.anp.gov.br/producao-de-derivados-de-petroleo-e-processamento-de-gas-natural/petroleo/refino>. Acessado em março de 2019.

AGÊNCIA NACIONAL DO PETRÓLEO (ANP); RESOLUÇÃO ANP Nº 8, DE 9.2.2011 - DOU 10.2.2011. Disponível em: <http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2011/fevereiro&item=ramp-8--2011>. Acessado em março de 2019.

CARVALHO, F.; FILHO, H. Artigo. **química nova**, v. 37, n. 1, p. 33–38, 2014.

DE SOUSA, A. M. Educação. **química nova**, v. 35, n. 1, p. 223–229, 2012.

FERREIRA, F. et al. Artigo. **química nova**, v. 36, n. 9, p. 1281–1287, 2013.

GONÇALVES, M. A. *et al.* Artigo. v. 36, n. 3, p. 393–399, 2013.

LEITE, D. C. *et al.* *Can Brazil replace 5 % of the 2025 gasoline world demand with ethanol ?* v. 34, p. 655–661, 2009.

OLIVEIRA, F. C. C. *Biodiesel: Possibilidades e Desafios*. 2008.

PAULA, D.; SOARES, M. *THE OIL INDUSTRY IN BRAZIL : EVALUATION OF HISTORIC*. p. 134–143, 2012.

PESQUISA, E. DE; EPE, E. *Balanço Energético Nacional Relatório Síntese | ano base 2017* |. 2018.

RASMUSSEN, A. K. S. *Analytical Methods*. *Analytical methods*, p. 2812–2831, 2014.

TRYGG, J. *Understanding food consumption patterns in Europe with principal component analysis (PCA)*. *Homepage of Chemometrics*. p. 1–3, 2003.