

CINTHIA RAQUEL DE SOUZA

**DETERMINAÇÃO DO NÚMERO DE CETANO NO ÓLEO
DIESEL CONVENCIONAL ATRAVÉS DE
ESPECTROSCOPIA DE RMN de ^1H ASSOCIADA À
CALIBRAÇÃO MULTIVARIADA.**

**Dissertação apresentada ao programa de
Pós-Graduação em Química como parte
dos requisitos para a obtenção do título de
Mestre em Química, área de concentração
QUÍMICA ORGÂNICA.**

Orientador: Prof. Dr. Andersson Barison

Curitiba/PR

2010

CINTHIA RAQUEL DE SOUZA

**DETERMINAÇÃO DO NÚMERO DE CETANO NO ÓLEO
DIESEL CONVENCIONAL ATRAVÉS DE
ESPECTROSCOPIA DE RMN de ^1H ASSOCIADA À
CALIBRAÇÃO MULTIVARIADA.**

**Dissertação apresentada ao programa de
Pós-Graduação em Química como parte
dos requisitos para a obtenção do título de
Mestre em Química, área de concentração
QUÍMICA ORGÂNICA.**

Orientador: Prof. Dr. Andersson Barison

Curitiba/PR

2010

DEDICATÓRIA

A *Deus*, por me dar forças, para continuar nos momentos difíceis.

A minha mãe, *Lourdes da Silva Alves* pelo apoio incondicional, pela compreensão e por todos os sacrifícios que fez por mim, durante toda minha vida.

AGRADECIMENTOS

Ao Prof^o. Dr. *Andersson Barison*, pela paciência, e por ter apoiado nos momentos mais difíceis.

A Prof^a. Dr. *Noemi Nagata*, pela contribuição dada nas análises quimiométricas, pela paciência e pela presteza.

A Prof^a. Dr. *Orliney Maciel Guimarães* e a Dr. *Francinete Ramos Campos* pelos conselhos e pela amizade.

A *REPAR* ao *João Luiz Tabora Ribas* pelo fornecimento de todas as amostras empregadas neste trabalho.

Aos colegas do Laboratório de RMN, *Aline, Caroline, Cláudio e Pardal*, por todos os momentos em que me ofereceram ajuda, pela amizade e pela ótima convivência.

Aos queridos amigos *Marco Antônio Medrado Xavier* e *Irecilse Drongek*, pelo apoio, pela amizade incondicional e pelo enorme incentivo em todas as esferas da minha vida.

A Dr. *Cynthia Maria Mattar Bernardelli Dias* e ao Sr. *Gilmar Afonso Kaminski*, pela compreensão e pelo apoio.

LISTA DE ABREVIATURAS E SÍMBOLOS

ANP	- Agência Nacional do Petróleo Gás Natural e Biocombustíveis
ASTM	- <i>American Society for Testing and Materials</i>
δ	- Deslocamento químico em partes por milhão (ppm)
FID	- Free Induction Decay
Hz	- Hertz
NC	- Número de cetano
PLS	- Partial Least Squares
R^2	- Coeficiente de correlação linear
REPAR	- Refinaria Presidente Getulio Vargas
RMN	- Ressonância Magnética Nuclear
RMN de ^1H	- Ressonância Magnética Nuclear de Hidrogênio
RMSEC	- <i>Root Mean Squared Error of Calibration</i>
RMSECV	- <i>Root Mean Square Error of Cross Validation</i>
TMS	- Tetrametilsilano
VL	- Variável Latente

LISTA DE TABELAS

Tabela 3.1 - Volumes de T-23 e U16, utilizados para se preparar 100 μ L de amostra, de acordo com cada número de cetano, utilizados para calibração do motor a combustão.....	23
Tabela 4.1 – Atribuição dos deslocamentos químicos de RMN de ^1H da mistura dos padrões	29
Tabela 4.2 – Atribuição dos deslocamentos químicos de RMN de ^1H do óleo diesel convencional. ...	35
Tabela 4.3 - Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 6,75 – 8,00 ppm, para os espectros obtidos a partir da mistura dos padrões.....	39
Tabela 4.4 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5 – 2,75 ppm e 6,75 – 8,00 ppm, para os espectros obtidos a partir da mistura dos padrões.	39
Tabela 4.5 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5 - 2,75 ppm, para os espectros obtidos a partir da mistura dos padrões.	40
Tabela 4.6 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5 – 8,00 ppm, para os espectros obtidos a partir da mistura dos padrões.....	40
Tabela 4.7 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 6,75 – 8,00 ppm, para os espectros obtidos a partir do óleo diesel convencional.	41
Tabela 4.8 - Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5 – 2,75 ppm e 6,75 – 8,00 ppm, para os espectros obtidos a partir do óleo diesel convencional.	41
Tabela 4.9 - Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5 – 2,75 ppm, para os espectros obtidos a partir do óleo diesel convencional.	42
Tabela 4.10 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5 – 8,00 ppm, para os espectros obtidos a partir do óleo diesel convencional.	42
Tabela 4.11 – Comparação dos coeficientes de correlação linear (R^2) e RMSECV obtidos utilizando-se as diferentes regiões dos espectros. Utilizando o modo especial de integração e <i>buckets</i> de 0,01 ppm.	44
Tabela 4.12 – Comparação entre os modelos construídos utilizando a mistura de padrões.....	49
Tabela 4.13 - Comparação dos valores obtidos para os modelos construídos com diferentes valores de variáveis latentes para diesel convencional.	55

LISTA DE FIGURAS

Figura 1.1 – Estrutura <i>n</i> -hexadecano (cetano) e 2,2,4,4,6,8,8- heptametilnonano (isocetano).....	8
Figura 4.1 – Comparação dos espectros de RMN de ¹ H dos padrões T-23 e U-16.	30
Figura 4.2 – Comparação dos espectros de RMN de ¹ H dos padrões T-23 e U-16. Ampliação da região entre 0,5 e 2,00 ppm.	30
Figura 4.3 – Comparação entre os espectros de RMN de ¹ H das diferentes proporções da mistura dos padrões T-23 e U-16	32
Figura 4.4 – Comparação entre os espectros de RMN de ¹ H das diferentes proporções da mistura dos padrões T-23 e U-16. Ampliação das regiões de 2,10 - 2,33, 1,21 - 1,31 e 0,70 - 0,90 ppm.	33
Figura 4.5 – Comparação entre os espectros de RMN de ¹ H da mistura dos padrões contendo 45% de T-23 e 55% de U-16 e uma amostra de óleo diesel, ambos com número de cetano de 44,8.....	34
Figura 4.6 – Comparação entre os espectros de RMN de ¹ H da mistura dos padrões contendo 45% de T-23 e 55% de U-16 e uma amostra de óleo diesel, ambos com número de cetano de 44,8. Ampliação da região de 0,5 - 2,5 ppm	34
Figura 4.7 – Espectro de RMN de ¹ H representativo de uma amostra de óleo diesel convencional com número de cetano igual a 43,7.....	36
Figura 4.8 – Comparação entre os espectros de RMN de ¹ H de amostras de óleo diesel convencional com diferentes números de cetano.	36
Figura 4.9 – Comparação entre os espectros de RMN de ¹ H de amostras de óleo diesel convencional com diferentes números de cetano. Ampliação da região entre 0,70 e 1,32 ppm.	37
Figura 4.10 – Gráfico do número de variáveis latentes <i>versus</i> RMSECV, para as amostras obtidas a partir dos padrões T-23 e U-16.	45
Figura 4.11 – Gráfico de Leverage <i>versus</i> Resíduos de Students, utilizando 5 variáveis latentes para construção do modelo, com limite de Leverage de 0,3	46
Figura 4.12 – Gráfico utilizado para escolha do número de variáveis latentes em função dos valores de RMSECV.....	47
Figura 4.13 – Gráfico de Leverage <i>versus</i> Resíduos de Students, utilizando 4 variáveis latentes, com limite de Leverage de 0,24, com a exclusão das amostras 8, 9 e 11	48
Figura 4.14 – Número de cetano medido <i>versus</i> previsto sem validação cruzada, utilizando 4 variáveis latentes para construção do modelo.....	50
Figura 4.15 – Número de cetano medido <i>versus</i> previsto com validação cruzada, utilizando 4 variáveis latentes para construção do modelo	50
Figura 4.16 – Comparação entre o número de cetano medido pelo método de referência <i>versus</i> previsto pelo modelo de calibração construído a partir da mistura dos padrões T-23 e U-16.....	52

Figura 4.17 – Número de variáveis latentes <i>versus</i> RMSECV, para as amostras de diesel convencional com número de cetano previamente determinado pelo método ASTM D613.....	53
Figura 4.18 – Gráfico de Leverage <i>versus</i> Resíduos de Students, utilizando 4 variáveis latentes para construção do modelo, com limite de Leverage de 0,12.....	54
Figura 4.19 – Número de cetano medido <i>versus</i> previsto sem validação cruzada, utilizando 2 variáveis latentes para construção do modelo.....	57
Figura 4.20 – Número de cetano medido <i>versus</i> previsto com validação cruzada, utilizando 2 variáveis latentes para construção do modelo	57
Figura 4.21 - Comparação entre o número de cetano medido pelo método de referência <i>versus</i> previsto pelo modelo de calibração construído a partir das amostras de diesel.....	59

RESUMO

O número de cetano (NC) é um dos mais importantes indicadores da qualidade do diesel combustível, tais como, propriedades ignitórias (autocombustão) e emissões atmosféricas, entre outros. Tradicionalmente, sua determinação é realizada através da queima do combustível em um motor a explosão, calibrado com dois compostos de diferentes propriedades de atraso na ignição (ASTM D613). No entanto, este método é bastante laborioso, além de possuir elevado custo, o que dificulta sua aplicação rotineira. Estes fatos, aliados a crescente utilização de motores a diesel no Brasil e, as exigências cada vez maiores nos padrões de qualidade, implicaram na necessidade de novas metodologias analíticas para controle de qualidade de combustíveis. Ao encontro desta necessidade, a RMN é uma técnica analítica que permite analisar amostras sem a necessidade de tratamentos prévios, com maior rapidez e utilizando uma quantidade menor de amostra. Neste contexto, a RMN, associada aos métodos de calibração multivariada, tem se mostrado uma ferramenta útil em diversos outros processos de controle de qualidade. No entanto, até o momento, há poucos relatos da utilização da RMN na determinação do NC do óleo diesel. Neste trabalho foi desenvolvido um novo método para a determinação do NC do óleo diesel, para isso, dois modelos de calibração foram construídos: um a partir da misturas dos padrões, empregados na calibração do motor de acordo com o método de referência e outro a partir de amostras de diesel com NC previamente determinado pelo método de referência. O modelo de calibração multivariada obtido a partir da mistura dos padrões apesar de possuir maior coeficiente de correlação linear (R^2), menos eficiente na previsão do NC de novas amostras de diesel. Por outro lado, o modelo de calibração multivariada obtido a partir das amostras de diesel com NC determinado pelo método de referência foi capaz de prever todas as novas amostras dentro do intervalo de precisão aceitável, configurando-se mais eficiente na previsão do NC. No entanto, este modelo necessita de um conjunto de calibração com amostras de diesel com NC previamente determinadas pelo método de referência, ao passo que obtido a partir dos padrões elimina totalmente esta necessidade.

ABSTRACT

Cetane number (CN) is one of the most important parameter indicting diesel oil quality, regarding mainly ignition properties and atmospheric emissions and others. Traditionally, CN determination is performed by calibrating an explosion engine with two compounds of different delay on the engine ignition. This method is quite laborious, time and fuel consuming and high cost, which difficult its usual employment. These facts allied with the increasing demand for diesel engines in Brasil and more rigorous exigency on the quality pattern resulted in the need for developing new analytical methods to be employed in the quality control of fuel. In response to this concern, NMR spectroscopy is increasingly as a versatile and powerful analytical tool for the quality control in several areas, like petroleum and fuels, which permits the fast and direct investigation in a small sample amount without the need of any sample treatment. However, to the best of our knowledge, there are only few reports in the literature regarding the employment of NMR in CN evaluation. In this work, a new method for CN determination based on two multivariate calibration models was developed. The first one was elaborated from the standards used in the engine calibration according to reference method, while the second was constructed from diesel oils sample with CN previously determined by the reference method. Both the methods were able on the CN prevision of new diesel samples. The model obtained from the standard shows high correlation coefficient (R^2), nevertheless was less efficient on the CN prevision of new samples. On the other hand, the model obtained from diesel oils samples was able on the CN prevision of all the new samples within interval required. Although, this method needs of a calibration set of samples with CN previously determined by reference method, while the model from the standard eliminated this need.

SUMÁRIO

LISTA DE ABREVIATURAS E SÍMBOLOS.....	v
LISTA DE TABELAS	vi
LISTA DE FIGURAS.....	vii
RESUMO.....	ix
ABSTRACT	x
SUMÁRIO.....	xi
1 – INTRODUÇÃO	2
1.1 – Motores do ciclo de diesel	2
1.2 – Ressonância Magnética Nuclear.....	9
1.3 – Quimiometria.....	10
1.3.1 – Calibração Multivariada por Mínimos Quadrados Parciais (PLS).....	15
1.4 – Espectroscopia de RMN aliada a Quimiometria	16
2 - OBJETIVOS.....	20
2.1 – Geral	20
2.2 – Especificos.....	20
3 – EXPERIMENTAL	22
3.1 – Amostras.....	22
3.1.1 – Padrões T-23 e U - 16.....	22
3.1.2 – Óleo diesel convencional.....	23
3.2 – Espectros de RMN de ¹ H.....	23
3.3 – Análises quimiométricas	24
3.3.1 – Transformações (Ajustes) dos espectros de RMN de ¹ H	24
3.3.2 – Pré-processamentos	26
3.3.2.1 – Análises por Regressão de mínimos quadrados parciais (PLS)	26
3.3.2.2 – Validação Cruzada	26

4 – RESULTADOS E DISCUSSÃO	28
4.1 – Análise visual dos espectros de RMN de ^1H	28
4.1.1 – Mistura dos padrões.....	28
4.1.2 – Amostras de diesel convencional	35
4.2 – Determinação do tamanho dos <i>buckets</i> , modo de integração e região do espectro utilizada para construção dos modelos de calibração multivariada.....	37
4.3 – Análises quimiométricas das amostras obtidas a partir da mistura dos padrões	44
4.4 – Análises quimiométricas das amostras de diesel convencional.....	52
5 – CONCLUSÕES.....	61
6 – REFERÊNCIAS BIBLIOGRÁFICAS	63

Introdução

1 – INTRODUÇÃO

1.1 – Motores do ciclo Diesel e o óleo diesel

O óleo diesel é um subproduto da destilação do petróleo que historicamente, devido a sua aparência e odor, não parecia ser útil, como os demais derivados do petróleo, que eram utilizados na iluminação pública. Sua importância surgiu somente com o desenvolvimento de motores de combustão interna de ciclo Diesel. Rudolf Christian Karl Diesel (1858-1913) era um engenheiro alemão especialista em Ciclos térmicos. Em 1890, Diesel teve a idéia de aperfeiçoar o motor de combustão interna a gasolina e desenvolveu um protótipo que comprimia a mistura ar-combustível a uma pressão de aproximadamente 30 atm. Isto era conseguido num motor mono-cilindro alimentado originalmente com óleo de amendoim que aquecia a mistura e provocava a auto-ignição (Souza, 2008).

O primeiro protótipo, dos motores do ciclo de Diesel, utilizando óleo diesel, foi construído em 1893 e se destruiu porque não haviam dispositivos mecânicos que agüentassem trabalhar em altas pressões. O segundo protótipo, desenvolvido em 1896 tinha uma eficiência de 20% e operava apenas a 30 atm de pressão. A grande dificuldade, apresentada por esses motores, era injetar o combustível na câmara de combustão. Por problemas de patente, ele só pôde retomar o desenvolvimento do motor a partir de 1908, quando foram construídos motores pequenos para uso em carros e máquinas. Somente em 1927, depois de cerca de 10 anos de a morte de Diesel, a Bosch desenvolveu um sistema de injeção direta de combustível que viabilizou a produção de motores movidos a óleo diesel em escala comercial, o que gera um imenso impacto na economia mundial até os dias atuais. Dada a importância do invento do motor Diesel, o combustível para o mesmo, que tinha características autodetonantes especiais, passou a se chamar óleo diesel (Souza, 2008).

O princípio de funcionamento dos motores diesel difere significativamente dos motores de ignição por centelha que são dominantes nas aplicações de carros de passeio. Nos motores de ignição por centelha, o combustível e o ar são misturados em proporções estequiométricas bem definidas e introduzidos em uma cavidade cilíndrica (câmara de combustão), sendo então comprimidos por um pistão, seguida da ignição por uma centelha elétrica, conhecida como vela de ignição. A potência do motor é controlada pela quantidade da mistura ar-combustível que entra na câmara de combustão, utilizando-se uma válvula reguladora de fluxo, chamada de acelerador. No motor diesel, também

conhecido como motor de ignição por compressão, apenas o ar entra no cilindro através do coletor de admissão. Este ar é comprimido em altas temperaturas e pressão, quando então o combustível finamente vaporizado é pulverizado em alta velocidade no ar. Quando em contato com o ar em altas temperaturas, o combustível vaporiza-se rapidamente e, ao misturar-se com o ar, reações químicas espontâneas resultam na ignição espontânea ou auto-ignição. Nos motores do ciclo de Diesel não há necessidade de vela de ignição, entretanto, alguns motores são equipados com velas de aquecimento para proporcionar uma melhor partida sob baixas temperaturas (Knothe *e col.*, 2005).

A potência do motor é controlada pela variação do volume de combustível injetado no interior da câmara de combustão, assim não há a necessidade do acelerador. O tempo envolvido no processo de combustão deve ser precisamente controlado para proporcionar que as reações químicas sejam completas e, com isso, otimizar a eficiência do combustível além de baixas emissões de poluentes. Esta duração é a soma dos tempos de injeção acrescida de um curto período entre o início da injeção e a auto-ignição, chamado de retardamento de ignição. Assim que ocorre a auto-ignição, a porção de combustível acumulada no retardamento queima muito rapidamente, durante um período denominado combustão em pré-mistura. Enquanto o combustível que tinha sido preparado durante o retardamento da ignição é consumido, o restante do combustível queima a uma taxa que depende da relação ar/combustível. Este período é conhecido como o de combustão com mistura controlada (Knothe *e col.*, 2005).

Os motores diesel variam muito em tamanho, potência e velocidades operacionais. Podem ser projetados para operar em ciclos de dois ou quatro tempos e podem ter ou não superalimentação (Campos *e col.*, 1989). A classificação usual dos motores diesel é feita em função de sua rotação. Com uma variedade tão grande de motores diesel torna-se evidente que a escolha do combustível mais adequado não é uma tarefa simples.

As variáveis mais importantes que determinam esta escolha são:

- Tamanho e projeto do motor;
- Faixas de velocidade e cargas operacionais;
- Frequência das mudanças de velocidade e carga;
- Manutenção do motor;
- Condições atmosféricas;
- Disponibilidade e preço do combustível.

Os combustíveis normalmente utilizados nos motores Diesel, ou seja o óleo diesel, são constituídos de misturas de hidrocarbonetos, com uma faixa de pontos de ebulição variando aproximadamente entre 150 °C e 370 °C, o que corresponde aos destilados intermediário do petróleo, que destilam após o querosene e assemelham-se aos gasóleos mais leves (Campos *e col.*, 1989). A composição química do óleo diesel é muito variável no que diz respeito à distribuição desses hidrocarbonetos, que podem ser classificados em três tipos: parafinas, naftênicos e aromáticos para os produtos de destilação direta, aparecendo ainda olefinas quando o óleo diesel contém, também, produtos de craqueamento (Campos *e col.*, 1989).

A predominância de um ou de outro tipo depende do petróleo que deu origem ao combustível e do processamento e tratamento que ele sofreu na refinaria. O peso molecular médio é também variável, mas, geralmente, o óleo diesel apresenta hidrocarbonetos variando entre nove e vinte átomos de carbono. A distribuição dos hidrocarbonetos dessas quatro categorias é muito importante, porque as propriedades do combustível serão em função dela. Assim a relação carbono/hidrogênio, tão importante na combustão, é variável, sendo maior nos hidrocarbonetos aromáticos e menor nos parafínicos. Propriedades como a densidade, faixa de destilação, poder calorífico, viscosidade, ponto de fluidez, número de cetano, entre outras, variam com a distribuição dos diversos hidrocarbonetos (Campos *e col.*, 1989).

Originalmente o óleo diesel era constituído por produtos provenientes da destilação direta do petróleo, mas a crescente demanda fez com que surgissem combustíveis para motores a diesel contendo destilados craqueados, permitindo aumentar o volume desse produto sem onerar seu preço. Em alguns países, o óleo diesel contém, ainda, certos aditivos, como aumentadores do número de cetano, redutores do ponto de fluidez e supressores de fumaça (Campos *e col.*, 1989). Em sua maioria, o óleo diesel é obtido por destilação direta do petróleo na operação de topeamento. A fração de onde provém o corresponde a um produto mais pesado do que o querosene e mais leve do que o gasóleo pesado. Hoje, com desenvolvimento de vários processos de craqueamento, o combustível diesel pode conter quantidades variáveis de destilados craqueados selecionados, que contribuem para aumentar a disponibilidade do produto. Também existe um tipo mais pesado de óleo, obtido pela mistura de destilados mais viscosos com óleos combustíveis residuais. Este produto é usado em motores de baixa e média velocidade, trabalhando em regime de carga e velocidade constantes (Campos *e col.*, 1989).

No Brasil, até a década de 1980, era produzido e comercializado apenas um tipo de óleo diesel. No entanto, o rápido aumento na demanda por esse produto obrigou a Petrobras a maximizar a produção de óleo diesel em suas refinarias. Esse aumento na produção de óleo diesel foi possível incorporando-se frações mais leves da faixa da gasolina e frações mais pesadas obtidas em processos de craqueamento catalítico e térmico.

Atualmente, no Brasil, são comercializados quatro tipos específicos de óleo diesel, cuja característica principal reside no teor de enxofre contido na formulação, entretanto outros tipos de óleo diesel já estão sendo desenvolvidos utilizando tecnologias mais limpas, ou seja que geram menos resíduos e poluentes. Os principais tipos produzidos no Brasil são o óleo diesel do tipo B, que possui teor de enxofre em torno de 0,50% em massa e está disponível em todo o território nacional, exceto nas grandes metrópoles. Óleo diesel do tipo C, cujo teor de enxofre está em torno de 0,30%, disponível nas regiões metropolitanas de Belo Horizonte, Porto Alegre, Curitiba, Belém, Campinas e São José dos Campos. O tipo D que apresenta no máximo 0,20% de enxofre e deste modo é menos poluente, está disponível nas regiões metropolitanas de São Paulo, Santos, Cubatão, Salvador, Aracaju, Rio de Janeiro, Recife e Fortaleza. Existe ainda o óleo diesel marítimo, que possui ponto de fulgor em torno de 60 °C (nos demais não há limite para essa especificação), sendo produzido exclusivamente para utilização em motores de embarcações marítimas. No Brasil, o óleo diesel é desenvolvido para atender às exigências específicas dos testes de avaliação de consumo e emissão de poluentes pelos motores Diesel, sendo utilizado pelos fabricantes de motores e pelos órgãos responsáveis pela homologação do combustível (Cardoso, 2005).

A popularidade dos motores a Diesel deve-se, principalmente, à eficiência do óleo diesel como combustível (principalmente devido ao poder de autoignição) em relação à gasolina, ou mesmo com relação a outros combustíveis simples ou misturados, como o metanol, por exemplo, o que chega a conferir uma economia relativa de 25 a 45%. Além disso, o preço do diesel é inferior ao de outros combustíveis, o que se reflete na grande utilização do mesmo em caminhões, ônibus e máquinas agrícolas. Os motores Diesel apresentam ainda uma excepcional durabilidade, sendo comum um motor Diesel, para veículos de carga pesada, ter um tempo de vida superior a um milhão de quilômetros, ou seja, cerca de dez vezes mais que a durabilidade apresentada por um motor à gasolina, por exemplo. Assim, as vantagens oferecidas pelos motores a diesel, tais como durabilidade, segurança e eficiência, justificam sua utilização em vários tipos de máquinas e veículos,

apesar dos problemas relacionados com a composição química e emissões atmosféricas (Braun *e col.*, 2004).

O óleo diesel deve satisfazer diversos requisitos de qualidade, deve apresentar características que visem para garantir um bom funcionamento do motor dentre estas deve permitir boa partida, proporcionar um aquecimento uniforme e aceleração suave, proporcionar uma operação suave sem problemas de detonação, minimizar a emissão de poluentes, manter limpos injetores, câmaras de combustão e exaustão, minimizar a corrosão e o desgaste das peças, evitar excessiva diluição do óleo lubrificante, proporcionar vida longa aos filtros, e permitir máxima economia de combustível (Campos *e col.*, 1989).

As características de partida, aquecimento, aceleração suave e tendência do combustível diluir o óleo lubrificantes dos motores estão relacionadas com a volatilidade do mesmo, isto é, com sua faixa de destilação. O problema relacionado a detonação do diesel depende da sua qualidade de auto-ignição e portanto do retardo da ignição, característica avaliada pelo número de cetano e influenciada pela volatilidade do combustível. A tendência em formar fumaça depende do número de cetano e da viscosidade. Os depósitos produzidos estão relacionados com a volatilidade e o teor de enxofre, o resíduo de carbono e a estabilidade do combustível. A corrosão e o desgaste são, principalmente, função do teor de enxofre. A vida útil dos filtros está relacionada ao teor de água e sedimentos, assim como sua estabilidade. A densidade relaciona-se ao poder calorífico, deste modo afeta a economia de quilometragem por litro (Campos *e col.*, 1989).

A Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) estabelece diversas especificações que visam garantir um bom desempenho aos motores Diesel, minimizar as emissões atmosféricas e adequar o combustível aos vários tipos de motores e as suas condições de operação. A ANP através da Resolução 15/2006 estabelece as especificações do óleo diesel usado no transporte rodoviário brasileiro. Os parâmetros de qualidade devem ser avaliados e termos da aparência, composição, volatilidade, fluidez, combustão, corrosão, contaminantes, lubricidade e número de cetano (ANP, 2006).

Com a crescente demanda no uso do óleo diesel, maiores percentagens de destilados craqueados, têm sido utilizadas, que causam o empobrecimento das propriedades de ignição o que pode induzir há alguns problemas como a ignição do motor em baixas temperaturas. Deste modo os combustíveis a diesel necessitam melhorar suas propriedades de ignição. Já que estas podem ser avaliadas em termos do número de

cetanos, o aumento no número de cetano é um método de melhora nas propriedades ignitórias do óleo diesel (Robbins *e col.*, 1951).

O número de cetano, que é uma medida da qualidade de ignição do óleo diesel, é determinada por um ensaio em motor de combustão interna de ciclo Diesel, conforme especificado pela norma ASTM D613. A qualidade de ignição, que em certa medida caracteriza a cinética de combustão, é quantificada através da medida do atraso na ignição, ou seja, o período entre o momento da injeção e o início da combustão (ignição) do combustível. A escala de número de cetano é baseada nas características de ignição de dois hidrocarbonetos puros, sendo um de cadeia linear, o *n*-hexadecano (Cetano) e seu isômero de cadeia ramificada, o 2,2,4,4,6,8,8-heptametilnonano (Isocetano), ambos com 16 átomos de carbono em sua estrutura (Figura 1). O cetano possui alta qualidade de ignição e, conseqüentemente, um pequeno atraso na ignição ao qual é atribuído um número de cetano de 100, enquanto que o isocetano possui uma baixa qualidade de ignição e, portanto, um longo retardo na ignição, ao qual é atribuído um número de cetano igual a 15. Desta forma, o número de cetano reflete a capacidade do combustível entrar em auto-ignição, pois quanto maior for o número de cetano, menor será o retardo na ignição e, por conseguinte, melhor será sua capacidade de queima. Logo, a qualidade de ignição do óleo diesel depende diretamente de sua composição química.

Apesar de não ser um indicativo direto da qualidade do combustível um óleo diesel com maior número de cetano, ao entrar em ignição mais rapidamente, tem mais tempo para queima durante o ciclo de expansão do motor, o que leva a menores níveis de emissão de gases poluentes, em especial de óxidos de nitrogênio (NO_x) e de material particulado. Valores inferiores a 40 podem causar fumaça na descarga, aumento de consumo, perda de potência e aumento de ruído (ASTM D613; Speight, 2007; Obert, 1971). O número de cetano para motores Diesel, em geral, situa-se a faixa de 40 a 60 (Obert, 1971). No Brasil, os combustíveis diesel comercializados apresentam normalmente número de cetano variando entre 45 e 51, sendo que valor mínimo especificado pela resolução ANP 15/2006 é de 42.

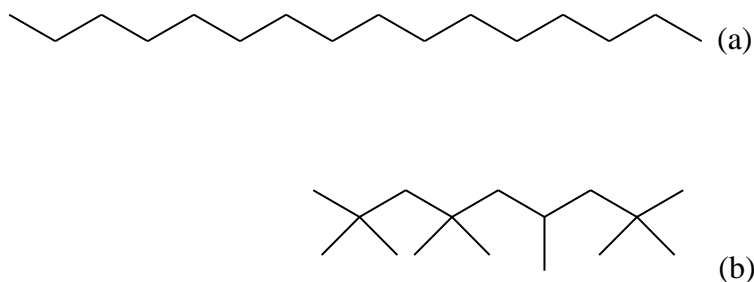


Figura 1.1 - Estruturas do *n*-hexadecano, conhecido como cetano (a) e seu isômero 2,2,4,4,6,8,8-heptametilnonano, conhecido como isocetano (b).

Deste modo, um combustível com um número de cetano alto possui um atraso de ignição menor e entra em autocombustão logo depois de ser injetado um motor. Com base nisto, os motores utilizados na determinação do número de cetano, são previamente calibrados através de quantidades pré-definidas, equivalente à percentagem em volume da mistura do cetano com o isocetano (ASTM D613; Speight, 2007), de acordo com a seguinte expressão:

$$NC = (\% \text{ cetano}) + 0,15 (\% \text{ isocetano})$$

Após a calibração do motor à amostra a ser avaliada é injetada no motor e seu atraso de ignição é comparado com aquele observado para mistura de cetano e isocetano (ASTM D613).

O método tradicional de determinação do número de cetano (ASTM D613) apresenta diversas desvantagens, o que inclui uma elevada quantidade de amostra (cerca de 1L), alto consumo de tempo para execução, em geral algumas horas, e baixa reprodutibilidade (Ghosh, 2008).

O número de cetano pode ser estimado através do índice de cetano calculado (ICC). Este método alternativo é fundamentado em cálculos baseados em equações que envolvem outras características do combustível. Em geral, o ICC é calculado a partir da densidade e temperaturas da curva de destilação de recuperação de 10%, 50% e 90% do produto. A norma ASTM D976 apresenta uma equação, que apresenta somente duas variáveis para calcular o ICC sendo também conhecida como equação de dois parâmetros, de acordo com a seguinte expressão:

$$ICC = 454,74 - 1641,416D + 774,74D^2 - 0,554T_{50} + 97,803(\log T_{50})^2$$

Onde, D é a densidade a $15\text{ }^{\circ}\text{C}$ (g/cm^3) e T_{50} é a temperatura de destilação de 50% do produto ($^{\circ}\text{C}$).

Desta forma, através do ICC, as refinarias durante a destilação do petróleo, são capazes de produzir óleo diesel com número de cetano aproximadamente superior a 42, visando atender a resolução ANP 15/2006. No entanto, as agências reguladoras exigem a determinação precisa do número de cetano através do método estabelecido pela norma ASTM D613.

1.2 - Ressonância Magnética Nuclear

A espectroscopia de ressonância magnética nuclear (RMN) é um das ferramentas analíticas mais poderosas para a elucidação e identificação estrutural de compostos. A RMN estabeleceu-se como uma técnica de imenso potencial e de grande utilidade para a análise dos produtos petrolíferos, especialmente da classe de combustíveis (Kapur *e col.*, 2001). É baseada na aplicação de forte campo magnético e pulsos de radiofrequência ao núcleo dos átomos dos compostos sob investigação. Para núcleos com número de prótons e/ou nêutrons ímpares, como ^1H e ^{13}C , a presença de um campo magnético externo fará com que seus spins nucleares possuam dois estados de energia e que os mesmos precessem ao redor do eixo do campo magnético aplicado. A absorção da energia do pulso de radiofrequência permitirá a promoção de estados de spin de baixa energia para de alta energia. Subseqüentemente ocorre a emissão de energia através de um processo de relaxação, no qual os spins voltam ao estado inicial de energia. Durante esse processo é possível medir a velocidade de precessão dos spins nucleares ao redor do eixo do campo magnético aplicado, o qual é dependente do ambiente químico em que o núcleo se encontra, e é denominado deslocamento químico (Silverstein, 2005).

A principal característica desta técnica é que o deslocamento químico medido pela espectroscopia de RMN é fortemente dependente do ambiente químico em que núcleos investigados se encontram. Além disso, é uma técnica de execução relativamente rápida e altamente versátil, sem restrição de ponto de ebulição da amostra, e altamente quantitativa. Desde seu surgimento em 1945, a RMN tem passado por grandes transformações, tais como o desenvolvimento de ímãs supercondutores, que permitiram obter campos magnéticos cada vez mais intensos, e com isso aumentar a sensibilidade e a resolução da técnica, a transformada de Fourier que possibilitou o uso da técnica pulsada e com isso acumular FIDs (Free Induction Decay), permitindo a observação de núcleos de baixa

sensibilidade, sondas com gradiente de campo, as quais permitiram melhorar significativamente a qualidade dos experimentos e mais recentemente as sondas criogênicas, que aumentaram muito a sensibilidade da técnica. Todas estas sofisticações acumuladas ao longo do tempo possibilitaram ampliar cada vez mais o campo da atuação da RMN, tornando a elucidação estrutural apenas uma das inúmeras aplicações da técnica.

1.3 - Quimiometria

Na ciência as idéias freqüentemente surgem bem antes da tecnologia necessária para explorá-las. Foi o que aconteceu em todos os ramos da Química onde a teoria indicava que a resolução de determinado problema exigiria muitos cálculos. Na química quântica, por exemplo, metodologias como Hartree-Fock, interação de configurações e teoria de perturbação Möller Plesset, apesar de concebidas nas primeiras décadas do século 20, somente puderam ser postas em prática muitas décadas depois, quando os recursos computacionais necessários se tornaram disponíveis. Da mesma forma, os métodos de matemática e estatística multivariada demoraram a ser usados, não só nas ciências naturais como em outras áreas, pela falta de tecnologia computacional capaz de suportar a pesada carga de cálculos necessários na maioria das aplicações de interesse prático. A importante análise de componentes principais, proposta por Pearson, já em 1901, e desenvolvida por Hotelling trinta anos depois, somente nos anos 70 começou a ser usada como método de classificação em química, sendo posteriormente introduzida na calibração multivariada, onde hoje é quase indispensável.

Um exemplo ainda mais dramático da distância entre a concepção e a aplicação prática é o da transformada de Fourier, que foi formulada na primeira metade do século XIX, mas levou cerca de um século e meio até se incorporar à instrumentação analítica (Neto *e col.*, 2006). A transformada de Fourier foi inclusive responsável por revolucionar a espectroscopia de RMN. O emprego da transformada de Fourier permitiu a utilização da técnica pulsada, e com isso adquirir de uma forma extremamente rápida os espectros de RMN. Tornando possível, por exemplo, a obtenção de espectros de RMN de ^{13}C em amostra diluídas (Claridge, 1999).

O uso de computadores para analisar dados químicos cresceu drasticamente nos últimos vinte anos, em parte devido aos recentes avanços em "hardware" e "software". Por outro lado, a aquisição de dados principalmente na área de química analítica, atingiu um ponto bastante sofisticado com o interfaceamento de instrumentos e técnicas aos

computadores produzindo uma enorme quantidade de informação, muitas vezes complexa e variada (Ferreira *e col.*, 1999).

Uma das características mais interessantes dos modernos instrumentos é o número das variáveis que podem ser medidas em uma única amostra. Um exemplo notável é a intensidade de absorção em mil ou mais comprimentos de onda que é rotineiramente registrada em um único espectro de infravermelho, ultravioleta etc. ou frequências de ressonância nos espectros de RMN. De posse de tal quantidade de dados, a necessidade de ferramentas novas e mais sofisticadas para tratá-los e extrair informações relevantes cresceu muito rapidamente, dando origem a Quimiometria, que é uma área especificamente destinada à análise de dados químicos de natureza multivariada (Ferreira *e col.*, 1999).

A quimiometria é um ramo da química relativamente novo, mas já possui um imenso impacto no campo da espectroscopia, de modo que atualmente uma extensão de programas utilizados em análises quimiométricas são integrados aos laboratórios de espectroscopia e aos processos de instrumentalização com um padrão, como por exemplo o software AMIX nos espectrômetros Bruker. A quimiometria já é amplamente utilizada em diversos processos de controle de qualidade (Geladi, 2003).

Os métodos utilizados na quimiometria, a princípio, foram desenvolvidos em outras disciplinas que com a aplicação e pesquisas voltadas para o tratamento de dados químicos, acabou dando origem a uma nova área dentro da química analítica. Isso se deu após a segunda metade dos anos 60, com o surgimento de métodos instrumentais computadorizados para a análise química o que promoveu a geração de uma grande quantidade de dados. Até este período, os químicos baseavam suas decisões em uma pequena quantidade de dados que, na maioria das vezes, eram obtidos de forma lenta e dispendiosa. A partir dos anos 60, com a grande quantidade de dados obtidos de maneira rápida e com menor esforço, foi preciso analisar todos esses dados e extrair as informações relevantes. Foi então, que teve início à pesquisa e a utilização dos métodos matemáticos e estatísticos que acabaram resultando nessa nova área conhecida como quimiometria (Otto, 1999)

A quimiometria foi introduzida por um número de grupos de pesquisa, principalmente na área de química analítica e físico-química orgânica. As principais razões para o crescimento no uso da quimiometria foram duas: (1) a introdução da instrumentalização, podendo desta maneira obter respostas multivariadas para uma mesma amostra e (2) a maior viabilidade na utilização de microcomputadores. Este crescimento advém do fato, de que a utilização da estatística univariada não era mais suficiente para

descrever e criar modelos de experimentos. No início a quimiometria foi descrita como algo que ia além das limitações da estatística univariada, usando planejamento experimental, classificação multivariada e calibração multivariada, usando métodos derivados da matemática, estatística e informática. A quimiometria continua evoluindo e a sua definição possivelmente deve ser modificada de tempos em tempos, de forma que sejam incluídos os avanços obtidos (Geladi, 2003).

Existem três marcos importantes na evolução da quimiometria: (1) o uso de planejamento experimental é o mais importante aspecto da quimiometria, mas é amplamente baseado no trabalho da estatística e de sociedades tecnométricas, (2) o uso da quimiometria para substituir as rigorosas leis em físico-química orgânica através de correlações, o que levou a quimiometria a ser utilizada no desenvolvimento de drogas, relações quantitativas entre estrutura e atividade, e tem mostrado coincidir com genômica, proteômica, bioinformática, metabolômica etc. e (3) o uso da quimiometria em química analítica instrumental levou a melhoria em processos analíticos e industriais. A quimiometria tem se expandido a cada especialidade possível utilizando espectros, imagens e / ou técnicas de separação (Geladi, 2003).

Com o avanço da instrumentação, uma quantidade cada vez maior de dados vem sendo gerada e em diversos níveis de complexidade. Isso requer e possibilita o desenvolvimento de novos procedimentos para a análise de dados, possibilitando a determinação de propriedades de interesse dentro de sistemas relativamente complexos. A relação entre medidas instrumentais e valores da propriedade de interesse correspondentes realizados em padrões é feita através de uma série de operações matemáticas definidas como calibração (Valderrama, 2009).

O processo de calibração pode ser definido como uma série de operações que estabelecem, sob condições específicas, uma relação entre medidas instrumentais e os valores correspondentes à propriedade de interesse realizada em padrões. Entre os métodos de calibração existentes, sem dúvida, os mais difundidos são os métodos univariados, onde se tem apenas uma medida (resposta) instrumental para cada uma das amostras de calibração. Esses métodos são relativamente fáceis de serem aplicados, validados e se encontram descritos na literatura em vários trabalhos. Porém, a aplicação de modelos univariados é restrita a situações em que a grandeza que é medida diretamente no sistema (ex.: absorvância de um composto) é livre de interferentes que possam provocar desvios entre sua relação linear com a propriedade de interesse (Braga *e col.*, 2004).

A calibração multivariada é a aplicação de maior sucesso proveniente da combinação entre a quimiometria e dados espectrais. A calibração consiste em construir uma relação entre um produto químico ou biológico desejado, ou propriedade física da amostra a diversas respostas, por exemplo os comprimentos de onda em um espectro de infravermelho ou frequências de ressonância em um espectro de RMN. Os métodos de calibração existentes podem ser divididos quanto à complexidade ou dimensionalidade dos dados, em calibração de ordem zero, primeira e segunda ordem. Métodos de calibração de primeira ordem fazem uso de um vetor de medidas instrumentais para cada amostra. As calibrações de primeira ordem geram uma matriz unidimensional de dados $n \times i$. Para análise de “n” amostras têm-se “i” variáveis. Este o caso de espectros de RMN unidimensionais. Esses métodos possibilitam as análises mesmo na presença de interferentes, desde que esses interferentes estejam presentes nas amostras de calibração (vantagem de primeira ordem), determinações simultâneas e análises sem resolução. Dentre os modelos de calibração multivariada de primeira ordem os mais empregados são a Regressão por Componentes Principais (PCR - *Principal Component Regression*) e Regressão por Mínimos Quadrados Parciais (PLS - *Partial Least Squares*) (Valderrama e col., 2009).

As calibrações de segunda ordem são construídas para métodos que geram uma matriz de dados bidimensionais por amostra. Para a análise de “n” amostras, têm-se “i” e “j” variáveis, gerando uma matriz $n \times i \times j$. Este é o caso da RMN bidimensional. Esses métodos têm a grande vantagem de permitirem a determinação de espécies de interesse na presença de interferentes, mesmo que estes interferentes não tenham sido incluídos nas amostras de calibração (vantagem de segunda ordem). Além disso, o perfil espectral de cada composto linearmente independente, presente na amostra, pode ser estimado com dados de segunda ordem. Diversos modelos de calibração multivariada para dados de segunda ordem vêm sendo utilizados, tais como: Análise de Fatores Paralelos (PARAFAC - *Parallel Factor Analysis*), Análise de Fatores Paralelos com Dependência Linear (PARALIND - *Parallel Factor Analysis with Linear Dependencies*), Método Generalizado de Aniquilação de Posto (GRAM - *Generalized Rank Annihilation Method*), Regressão por Mínimos Quadrados Parciais N-dimensional (N-PLS), Resolução de Curvas Multivariadas (MCR - *Multivariate Curve Resolution*), Mínimos Quadrados Bilineares (BLLS - *Bilinear Least Squares*), PLS desdobrado (*Unfolded PLS*) e PCR desdobrado (*Unfolded PCR*) (Valderrama e col., 2009).

Calibrações de terceira ordem requerem um cubo de dados por amostra, ou seja para cada “n” amostras, têm-se “i”, “j” e “k” variáveis, gerando uma matriz $n \times i \times j \times k$. Este é o caso da RMN tridimensional. Contudo, um dos fatores que ainda restringe a aplicação desse tipo de calibração é a elaborada instrumentação necessária para a aquisição dos dados. Uma vantagem deste tipo de calibração é que como se tem um cubo de dados por amostra, o perfil característico de cada ordem pode ser unicamente determinado para cada espécie na amostra. Entretanto, maiores vantagens para esse tipo de calibração ainda são desconhecidas (Valderrama *e col.*, 2009).

Um dos mais importantes problemas na química analítica é a estimativa do número e concentração das espécies em misturas através de espectros ou cromatogramas. Uma série de técnicas estatísticas têm sido utilizadas para se desenvolver metodologias multivariadas e para extrair informação dos espectros e cromatogramas, com o objetivo de identificar as espécies presentes e determinar quantitativamente as concentrações de algumas ou todas elas. A aplicabilidade de cada metodologia depende do conjunto de dados (informação experimental) submetidos à análise. Do ponto de vista da química analítica estes métodos multivariados de uma forma geral podem ser classificados em três categorias: (1) quando se tem informação espectral de todas as espécies existentes na amostra; (2) quando a informação dos espectros das espécies químicas são incompletos ou parciais e (3) quando não existe nenhuma informação sobre a composição química das amostras (Scarminio *e col.*, 1998)

Em métodos de Calibração Multivariada duas ou mais respostas instrumentais são relacionadas com a propriedade de interesse. Esses métodos possibilitam análises mesmo na presença de interferentes, desde que estejam presentes nas amostras de calibração, determinações simultâneas, análises sem resolução etc. Isso faz com que, modelos de Calibração Multivariada sejam uma alternativa quando métodos univariados não podem ser aplicados. Diversos modelos de calibração multivariada vêm sendo utilizados, tais como Regressão Linear Múltipla (MLR), Regressão por Componentes Principais (PCR) e Regressão por Mínimos Quadrados Parciais (PLS). Esses modelos têm apresentado ótimos resultados, com aplicações em diversas áreas. Contudo, sua aceitação e implementação têm sofrido restrições em alguns setores, como as áreas farmacêutica e médica, devido às exigências de validação. Essa validação é feita com base no cálculo e análise das chamadas figuras de mérito, que certificam que o modelo multivariado proposto é confiável e atende às especificações impostas pela indústria e órgãos de fiscalização nacionais e internacionais como “United States Pharmacopeia” (USP), “International Conference on

Harmonisation's" (ICH), "American Society for Testing and Materials" (ASTM), Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO), Agência Nacional de Vigilância Sanitária (ANVISA) etc. (Braga *e col.*, 2004).

1.3.1 - Calibração Multivariada por Mínimos Quadrados Parciais (PLS)

A regressão por mínimos quadrados parciais (PLS) vem sendo usada para estabelecer uma relação matemática (modelo de calibração) entre os valores químicos de referência e os dados espectrais para uma determinada propriedade medida. Pré-tratamentos matemáticos aplicados aos dados espectrais são, em geral, necessários para facilitar a interpretação dos resultados e, também, melhorar a qualidade de previsão dos modelos de calibração (Morgano *e col.*, 2007).

O método de regressão dos Mínimos Quadrados Parciais (PLS-Partial Least Squares) foi desenvolvido por Herman Wold e colaboradores, entre as décadas de 70 e 80 (Wold *e col.*, 2001) e é o método normalmente mais empregado em análises de calibração multivariada. Neste método, os sinais multivariados, como, por exemplo, os valores das diferentes frequências de ressonância nos espectros de RMN (variáveis x) e concentrações das amostras (variáveis y), são utilizados para estabelecer um modelo de regressão linear. Para a construção do modelo de calibração, os dados são dispostos na forma de matrizes: Matriz X e Matriz Y que contêm os dados de variáveis independentes x e dependentes y , respectivamente (Barthus *e col.*, 2007).

Antes da aplicação de um modelo construído, o mesmo deve ser validado com o objetivo de testar sua capacidade preditiva. A validação consiste em testar o modelo prevendo concentrações de amostras (de preferência não usadas em sua construção), para estabelecer se ele de fato irá refletir o comportamento do analito de interesse. Durante a etapa de validação dois fatores devem ser levados em consideração: (1) o número de variáveis latentes utilizados para construção do modelo e (2) a presença de amostras anômalas (*outliers*) (Ferreira *e col.*, 1999).

Para a determinação do número de variáveis latentes que serão empregadas na modelagem, o ideal seria a utilização de um conjunto-teste de dados. Entretanto, na maioria das vezes isto não é possível, pois pode ser um processo demorado e caro. Uma alternativa prática, e que funciona bem, é o método de validação cruzada. A validação cruzada é uma metodologia utilizada para a escolha do número de componentes principais

baseada na avaliação da magnitude dos erros de previsão de um dado modelo de calibração. Esta avaliação é feita pela comparação das previsões das concentrações previamente conhecidas. Uma das formas de validação cruzada amplamente utilizada é o sistema *leave-one-out*, que consiste na retirada de um elemento (amostra) do conjunto de calibração, o qual é utilizado como elemento de previsão. O processo é repetido n vezes (onde n é o número de amostras). Este procedimento de validação cruzada é ideal para a construção de modelos com poucas amostras (Ferreira *e col.*, 1999).

A detecção de "outliers" é tão importante quanto à determinação do número de componentes principais que serão empregadas no modelo. Ao verificar a qualidade do conjunto de calibração, deve-se assegurar de que as amostras formam um conjunto homogêneo, removendo-se aquelas amostras que são solitárias. Para a detecção de "outliers", usam-se duas grandezas complementares: "leverage" e resíduos de Students (Ferreira *e col.*, 1999).

A "leverage" é uma medida da influência de uma amostra no modelo de regressão. Um valor de "leverage" pequeno indica que a amostra em questão influencia pouco na construção do modelo de calibração. Por outro lado, se as medidas experimentais de uma amostra são diferentes das outras do conjunto de calibração, ela provavelmente terá uma alta influência no modelo, que pode ser negativa. Em geral, estas amostras solitárias estão visíveis no gráfico de "scores". A "leverage" pode ser interpretada geometricamente como a distância de uma amostra ao centróide do conjunto de dados e é calculada segundo a equação (Ferreira *e col.*, 1999).

$$\text{Limite de leverage} = 3LV's/n$$

Onde, $LV's$ é o número de variáveis latentes e n é o número de amostras do conjunto de calibração.

1.4 - Espectroscopia de RMN aliada a Quimiometria

A química analítica está saindo dos laboratórios. A análise clínica, por exemplo, necessita de monitorização não invasiva e rápida, e a indústria precisa de controle de qualidade rápido de seu processo produtivo. Em geral, a química analítica evolui para a remoção de resíduos de produtos químicos pelo uso de miniaturização instrumental e

emprego métodos não destrutivos. Especialmente na indústria, onde não se pode esperar horas para que uma análise laboratorial responda se o processo produtivo está satisfatório ou se está fora de controle. O que a indústria necessita é a resposta rápida e análise não destrutiva da amostra. Neste contexto, a combinação de técnicas espectroscópicas e da quimiometria é ideal para solucionar essa situação. A exatidão e a precisão em testes laboratoriais podem muitas vezes ser sacrificados para que se obtenham respostas rápidas que geralmente são usadas para monitorar e controlar o processo produtivo continuamente (Geladi *e col.*, 2004).

Ao encontro desta necessidade, quimiometria é uma ferramenta estatístico-matemática valiosa, que em combinação com diferentes métodos analíticos, possibilita a análise de um grande número de variáveis (respostas) em uma única amostra (Breton, 2002; Beebe *e col.*, 1998).

Neste contexto, os dados obtidos através da RMN, são muito atrativos para as análises quimiométricas, devido à grande quantidade de informações contida nestes, por exemplo, em um único espectro de RMN de ^1H . Dispensando assim, a necessidade de outras medidas analíticas para aumentar a quantidade de informações a respeito do sistema em estudo. Em resumo, a grande vantagem da união entre quimiometria e a RMN, deve-se ao fato de uma ser uma poderosa ferramenta de análise de dados complexos, enquanto que a outra possibilita a obtenção de um grande número de informações (variáveis) em uma única medida, de uma forma simples e rápida. Por esta razão, estudos envolvendo a aplicação em conjunto destas ferramentas vem aumentando significativamente (Barison, 2005).

Nos últimos anos diversas pesquisas vêm sendo realizadas com a finalidade de desenvolver um método mais rápido para determinar o número de cetano de amostras de óleo diesel, porém sem muito sucesso. Dentre as ferramentas analíticas disponíveis para identificação e quantificação de compostos em misturas pode-se destacar a Cromatografia Gasosa de Alta Resolução acoplada à Espectrometria de Massas (CGAR-EM), a Espectroscopia na Região do Infra-Vermelho (IV) e do Ultra-Violeta (UV) e a Ressonância Magnética Nuclear (RMN) (Faria *e col.*, 2007). Entretanto, há poucos relatos na literatura da descrição de métodos que utilizem RMN para esta finalidade. Por outro lado já tem sido demonstrado o potencial da RMN na determinação do teor de biodiesel no diesel convencional através de RMN de ^1H aliado à calibração multivariada, e na avaliação da qualidade desta mistura (Monteiro *e col.*, 2009a; Monteiro *e col.*, 2009b). A maioria dos trabalhos utilizando a espectroscopia de RMN aliada a ferramentas estatísticas, para

controle de qualidade em combustíveis, tem sido relatada para a previsão da RON em amostras de gasolina. Em contraste, há poucos relatos na literatura sobre trabalhos semelhantes sobre o óleo diesel (Kapur *et al.*, 2001).

Objetivos

2 - OBJETIVOS

2.1 - Geral

- Desenvolver uma nova metodologia para determinar o número de cetano a partir de espectros de RMN ^1H e calibração multivariada.

2.2 - Específicos

I – Construir um modelo de calibração a partir dos padrões de T-23 e U-16 utilizados para calibração do motor de acordo com a norma ASTM D613;

II – Determinar o número de cetano das amostras de óleo diesel através do modelo construído a partir dos padrões;

III – Construir modelos de calibração com amostras de óleo diesel, com número de cetano previamente determinado pela REPAR, através do método de referência;

IV – Determinar o número de cetano de amostras de diesel através do modelo construído a partir das amostras de óleo diesel, com número de cetano previamente determinado.

Experimental

3 – EXPERIMENTAL

3.1 – Amostras

3.1.1 - Padrões T-23 e U-16

Os padrões T-23 e U-16 são utilizados para calibração do motor a combustão, onde é realizado o teste para determinação do número de cetano em amostras de diesel, de acordo com a norma ASTM D613. Estes padrões foram fornecidos pela REPAR, e devido à patente detida por uma empresa norte americana, não se sabe a composição exata dos mesmos. No entanto, pelos seus respectivos espectros de RMN de ^1H , pode-se perceber que os padrões T-23 e U-16, tratam-se dos compostos cetano e isocetano respectivamente, possivelmente com a adição de outras substâncias.

Conforme a norma ASTM D613, valores de número de cetano são atribuídos a partir de misturas em proporções definidas dos dois padrões. Sendo assim, foram preparadas amostras de 100 μL cada, com número de cetano no intervalo de 39,2 a 50,5 (Tabela 3.1), através da mistura de volumes específicos de T-23 e U-16.

Tabela 3.1 – Volumes de T-23 e U16, utilizados para se preparar 100 μL de amostra, de acordo com cada número de cetano, utilizados para calibração do motor a combustão.

Número de cetano	Volume de T-23 (μL)	Volume de U-16 (μL)
39,2	35	65
39,7	36	64
40,3	37	63
40,9	38	62
41,4	39	61
42,0	40	60
42,6	41	59
43,1	42	58
43,7	43	57
44,3	44	56
44,8	45	55
45,4	46	54
46,0	47	53
46,5	48	52
47,1	49	51
47,7	50	50
48,2	51	49
48,8	52	48
49,4	53	47
49,9	54	46
50,5	55	45

3.1.2 - Óleo diesel convencional.

As amostras de óleo diesel convencional, sem adição de biodiesel bem como de melhoradores do número de cetano, foram gentilmente cedidas pela Refinaria Presidente Getulio Vargas - REPAR, com número de cetano previamente determinado pelo método de referencia, estabelecido pela norma ASTM D613, em um intervalo variando entre 42,0 e 46,8.

3.2 – Espectros de RMN de ^1H

Os espectros de RMN de ^1H das amostras, obtidas a partir da mistura dos padrões e das amostras de óleo diesel convencional, foram adquiridos em um espectrômetro Bruker,

modelo AVANCE 400, operando a 9,4 Tesla, observando o núcleo de ^1H a 400,13 MHz, a temperatura de aproximadamente 295 K, usando-se a seqüência de pulsos de RD-90°-AQ (zg nos equipamentos Bruker). O espectrômetro foi equipado com uma sonda multinuclear de observação direta de 5 mm com gradiente de campo no eixo z. Os espectros, mais precisamente os FIDs (Free Induction Decay ou Livre decaimento da magnetização induzida) foram medidos com 64K pontos em uma janela espectral de 3641 Hz (~9,1 ppm), acumulando-se 8 varreduras, e tempo de aquisição (AQ) e relaxação (RD) de 9,4 e 1,0 segundos, respectivamente. Tetrametilsilano (TMS) foi utilizado como referência interna (δ 0,00) para referenciar os deslocamentos químicos dos espectros de RMN de ^1H .

Para aquisição dos espectros de RMN de ^1H , 30 μL da mistura de padrões T-23 e U-16 ou de óleo diesel convencional, foram transferidos diretamente para tubos de RMN de 5 mm, com auxílio de uma micropipeta volumétrica, seguido da adição de 570 μL de CDCl_3 contendo 0,01% de TMS. Todas as medidas foram realizadas em triplicata.

3.3 – Análises quimiométricas

As análises multivariadas foram realizadas empregando-se os espectros de RMN de ^1H , com auxílio dos programas computacionais AMIX[®] da Bruker e o programa PLS-Toolbox que opera em ambiente MatLab[®].

3.3.1 – Transformações (Ajustes) dos espectros de RMN de ^1H

Para se obter os espectros de RMN utilizados nas análises quimiométricas, os FIDs foram convertidos do domínio de tempo para o domínio de freqüências pela aplicação da transformada de Fourier. Sendo assim, os espectros foram construídos com 64K pontos, sendo que antes os FIDs sofreram uma multiplicação exponencial, para alisamento dos sinais, por um fator de 0,3 Hz. A linha de base dos espectros foi automaticamente corrigida através de um ajuste polinomial, enquanto que a fase dos sinais foi corrigida manualmente. Os espectros foram também calibrados em relação ao sinal do TMS em δ 0,00, como referência interna. Estas transformações e/ou ajustes foram efetuadas com auxílio do programa computacional TopSpin[®] da Bruker, que também foi utilizado para a aquisição dos espectros.

Para submeter os espectros à análise quimiométrica, o número de variáveis foi reduzido dividindo-se os espectros em pequenos intervalos, denominado de *buckets*. A conversão dos espectros em *buckets* foi realizada com o auxílio do programa AMIX[®], desenvolvido pela Bruker especialmente para explorar dados de RMN. Para a determinação das áreas dos *buckets* diversos modos de integração foram avaliados. Após a integração, as áreas dos *buckets* foram normalizadas em relação à área total de cada espectro. As áreas dos *buckets* foram então utilizadas como variáveis de entrada nas análises quimiométricas, as quais foram realizadas utilizando o programa PLS-Toolbox que opera em ambiente MatLab[®]. Em todos os casos a região entre 1,50 e 1,65 ppm foi excluída das análises, uma vez que a água dissolvida em CDCl₃ apresenta um sinal nesta região.

O principal problema enfrentado nas análises quimiométricas de dados espectrométricos, deve-se as variações na linha de base dos espectros. Para uma análise estatística a linha de base dos espectros deve ser aproximadamente a mesma para todos os conjuntos de dados. Caso a linha de base apresente oscilações muito significativas entre as amostras, os resultados obtidos a partir da construção do modelo de calibração multivariada, utilizando estas amostras, influenciará nos resultados fornecidos pelo modelo podendo conduzir a obtenção e interpretação de resultados equivocados.

No caso da RMN há ainda um fator adicional que é a fase dos sinais, que se não forem adequadamente ajustadas acarretará em alterações nas variáveis pela variação da intensidade dos sinais de RMN, as quais não serão devidas a alterações na composição química das amostras, e sim a diferenças no deslocamento químico de determinados sinais. Desta forma, para proceder às análises quimiométricas foi necessário que antes todos os espectros fossem adequadamente ajustados, para que a área das variáveis (*buckets*) dependesse somente das variações na composição química das amostras e a nenhum outro fator. Inclui-se aqui também a calibração de todos os espectros em relação a um sinal de referência para que cada núcleo em um mesmo ambiente químico apresente mesmo deslocamento químico independente da amostra. Neste trabalho, o sinal das metilas do TMS foi utilizado como referência em δ 0,00 ppm. Além disso, os espectros sofreram uma multiplicação exponencial Lorentziana dos FIDs por um fator de 0,3 Hz. Após o ajuste dos espectros, a matriz de dados para todas as amostras foram construídas com o auxílio do programa AMIX[®] da Bruker.

3.3.2 – Pré-processamentos

3.3.2.1 – Análises por regressão de mínimos quadrados parciais (PLS)

Para a análise quimiométrica dos espectros de RMN de ^1H , todos os dados referentes ao bloco X (amostras) foram autoescalados, enquanto que os referentes ao bloco Y (número de cetano medido) foram centrados na média, os tipos de pré-processamento utilizados foram determinados após várias tentativas e combinações.

3.3.2.2 – Validação cruzada

O sistema de validação cruzada é uma metodologia adotada para a validação dos modelos de calibração, principalmente para selecionar o número de variáveis latentes que será utilizado na construção dos modelos. Nos modelos construídos neste trabalho utilizou-se o sistema *leave-one-out*, que consiste na retirada de um elemento (amostra) do conjunto de calibração, o qual é utilizado como elemento de previsão. O processo é repetido n vezes (onde n é o número de amostras). Este procedimento de validação cruzada é ideal para a construção de modelos com poucas amostras (Ferreira *e col.*, 1999).

Resultados e discussão

4 – RESULTADOS E DISCUSSÃO

4.1 – Análise visual dos espectros de RMN de ^1H

4.1.1 – Mistura de padrões

A calibração do motor para determinação do número de cetano (NC) de acordo com a norma ASTM D613 é realizada através da alimentação do motor com uma mistura de hidrocarbonetos sendo um deles de cadeia linear, o *n*-hexadecano (cetano) e seu isômero de cadeia ramificada, o 2,2,4,4,6,8,8-heptametilnonano (Isocetano), ambos com 16 átomos de carbono em sua estrutura (Figura 1). O cetano possui um pequeno período de atraso na ignição e lhe é atribuído um número de cetano de 100, enquanto que o isocetano possui um longo período de atraso na ignição e lhe é atribuído um número de cetano igual a 15.

Os espectros de RMN de ^1H da mistura dos padrões T-23 e U-16 apresentaram sinais tanto na região aromática quanto alifática (Figura 4.1). Porém, os sinais na região alifática apresentaram-se bem mais intensos do que aqueles da região aromática, característicos de hidrocarbonetos alifáticos, o que pode ser um indicativo da presença dos padrões *n*-hexadecano e 2,2,4,4,6,8,8-heptametilnonano, ou compostos relacionados (Figura 4.1). Analisando-se a região alifática dos espectros de RMN de ^1H dos padrões, observa-se um sinal bastante intenso centrado em 1,25 ppm, para o padrão T-23, o que é indicativo da presença de grupos $-\text{CH}_2$ de cadeia linear extensa, além de um único tripleto em 0,86 ppm característico de um grupo metila terminal acoplado a um grupo CH_2 (Figura 4.2), indicando que o padrão T-23 contém o composto *n*-hexadecano (Cetano). Estes sinais não são observados para o padrão U-16, ao contrário são observados diversos sinais sobrepostos na região entre 0,70 e 0,90 ppm o que é indicativo da ausência de grupos $-\text{CH}_2$ de cadeia linear e presença de grupos metila em uma cadeia mais ramificada (Figura 4.2), indicando assim que U-16 contém o composto 2,2,4,4,6,8,8-heptametilnonano (Isocetano). A completa atribuição dos deslocamentos está apresentada na Tabela 4.1.

Tabela 4.1 – Atribuição dos deslocamentos químicos nos espectros de RMN de ^1H da mistura dos padrões.

δ^a (ppm)	Hidrogênio	Descrição
0,68 - 0,92	$\text{CH}_3\text{-C}$	Grupos metilas de cadeias alquilas lineares e/ou ramificadas
1,00 - 1,70	$\text{C-(CH}_2\text{)-C}$	Grupos metilênicos
2,00 - 2,95	$\text{-OC-CH}_2\text{-}$ ou -OC-CH_3	Grupos metilênicos ou metilas α carbonilas
6,75 - 8,00	H-Ar	Hidrogênio ligado a carbono aromático

^a Valores obtidos a 400 MHz em CDCl_3 em relação ao sinal do TMS em δ 0,00, utilizado como referência interna.

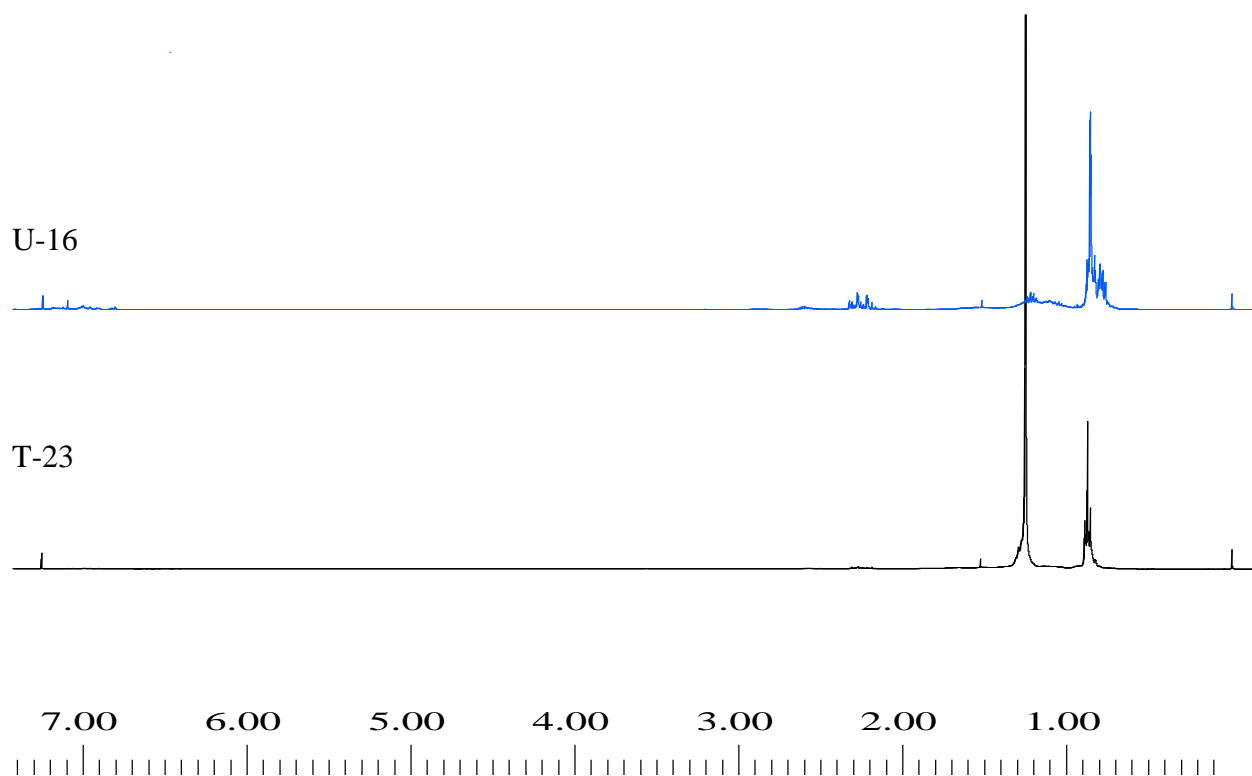


Figura 4.1 – Comparação dos espectros de RMN de ^1H dos padrões T-23 e U-16.

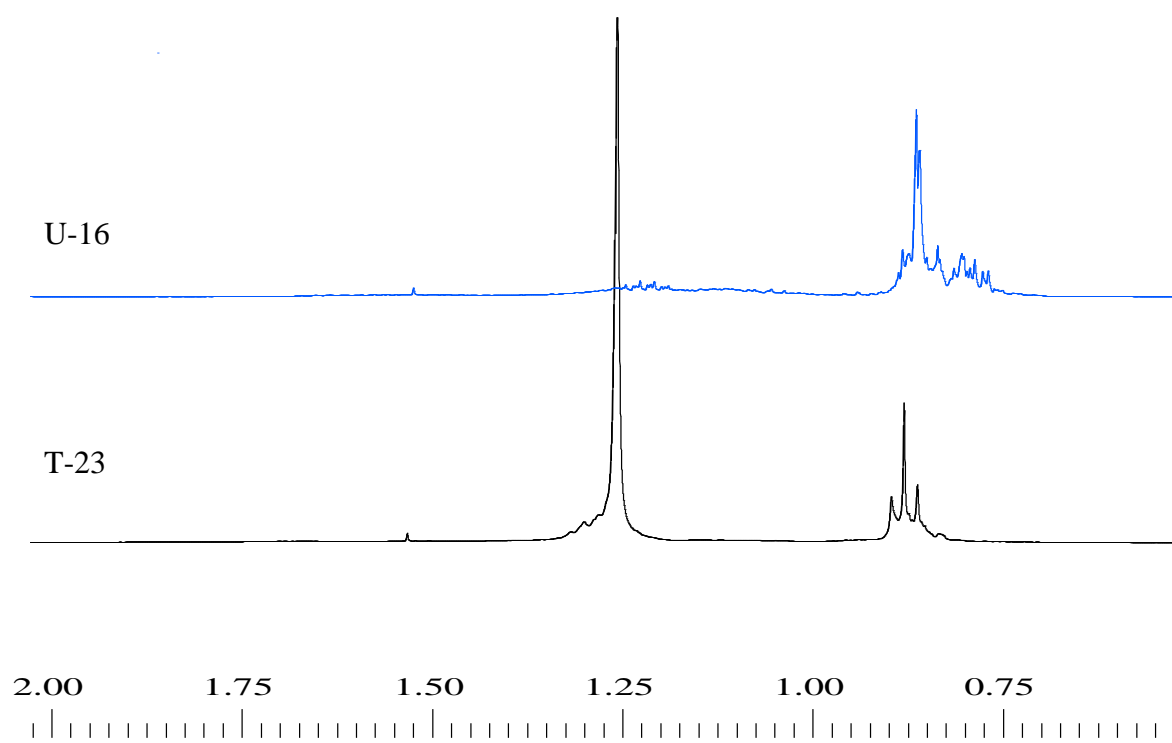


Figura 4.2 – Comparação dos espectros de RMN de ^1H dos padrões T-23 e U-16. Ampliação da região entre 0,5 - 2,0 ppm.

Através da análise visual dos espectros de RMN de ^1H de misturas em diferentes proporções dos padrões T-23 e U-16, é possível perceber que o perfil espectral vai sendo alterado conforme o número de cetano aumenta, ou seja conforme a concentração de T-23 vai sendo aumentada e a de U-16 vai sendo diminuída (Figura 4.3). Entre as principais alterações, observa-se que a intensidade do sinal centrado em 1,25 ppm referente ao padrão T-23 vai sendo intensificado conforme sua proporção na mistura aumenta o que também conduz ao aumento do número de cetano (Figura 4.3). Ao contrário os principais sinais presentes no padrão U-16 nas regiões de 0,70 a 0,90 e 2,10 a 2,70 bem como os sinais aromáticos vão diminuindo conforme sua proporção na mistura é diminuída (Figura 4.3). Desta forma, fica evidenciado que seria possível construir uma curva de calibração utilizando os próprios espectros de RMN de ^1H da mistura dos padrões T-23 e U-16 em diferentes proporções.

A presença de compostos aromáticos causa um impacto negativo no número de cetano e os compostos de referência cetano (NC=100) e isocetano (NC=15) não consideram a presença dos compostos aromáticos no óleo diesel. Portanto os compostos cetano e isocetano não podem unicamente ser considerados na calibração do motor de acordo com a norma ASTM D613. Desta forma, os padrões T-23 e U-16, provavelmente foram desenvolvidos para considerar também a presença de compostos aromáticos no óleo diesel (Santana *e col.*, 2006). Por este motivo, os valores de número de cetano apresentados na Tabela 3.1 são menores do que os valores calculados simplesmente através expressão: $\text{NC} = (\% \text{ cetano}) + 0,15(\% \text{ isocetano})$. Na Figura 4.1 pode-se observar a presença de compostos aromáticos, principalmente no padrão U-16 e a diminuição da presença deste coincide com aumento do número de cetano (Figuras 4.3 e 4.4). No entanto, devido à estreita faixa de NC investigada, essas diferenças são bastante sutis, sendo observadas somente com a ampliação das respectivas regiões nos espectros de RMN de ^1H (Figura 4.4).

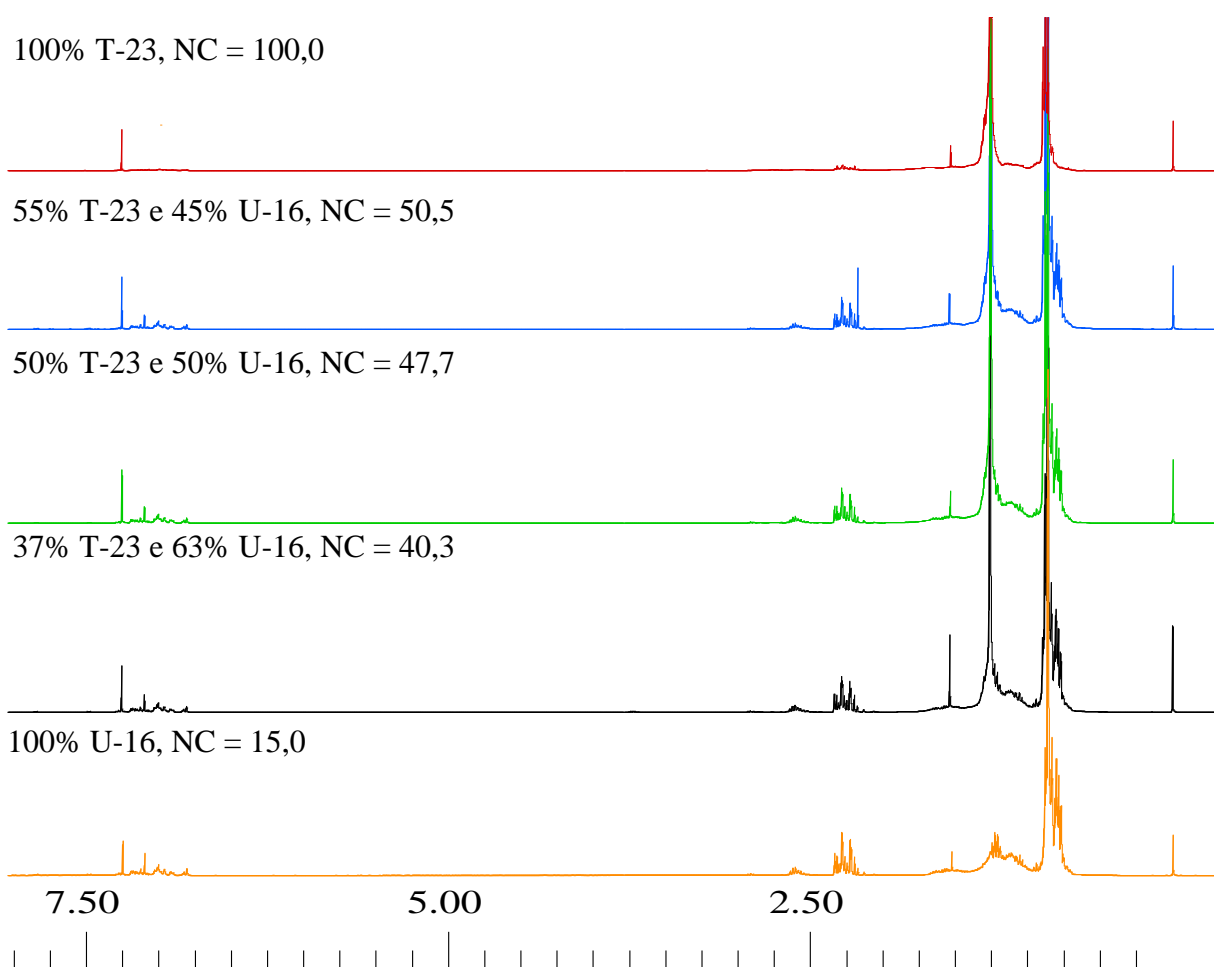


Figura 4.3 – Comparação entre os espectros de RMN de ^1H das diferentes proporções da mistura dos padrões T-23 e U-16.

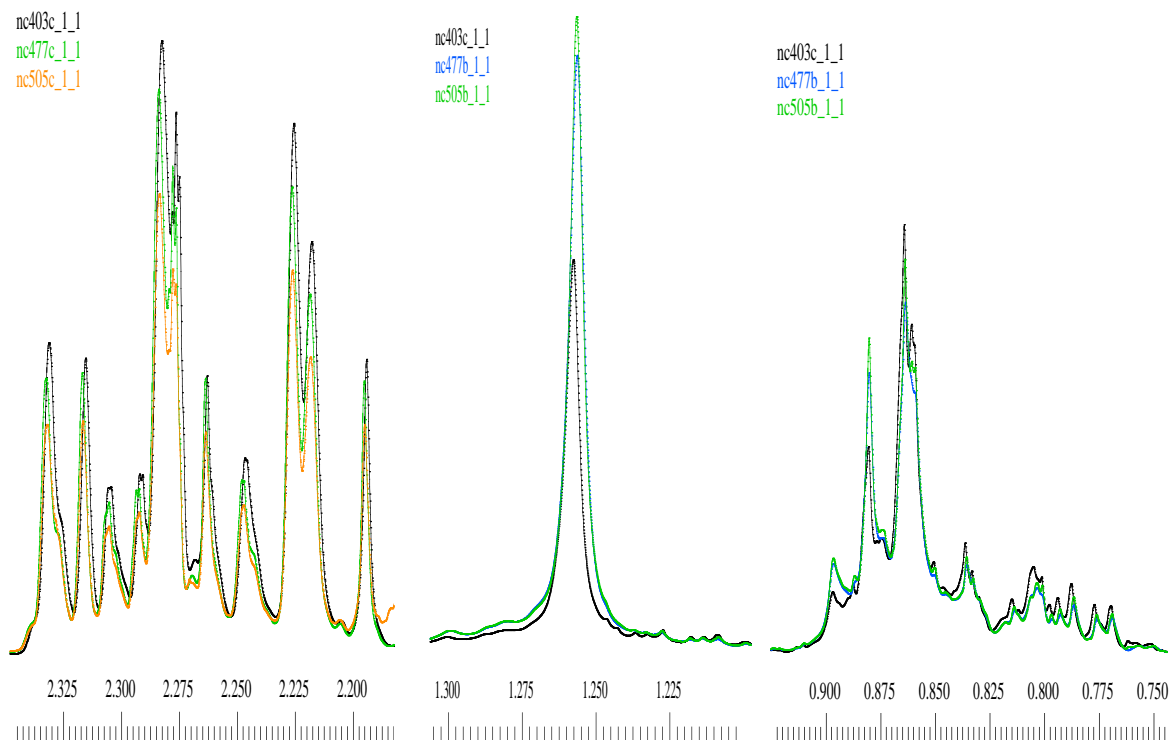


Figura 4.4 – Comparação entre os espectros de RMN de ^1H das diferentes proporções da mistura dos padrões T-23 e U-16. Ampliação das regiões 2,10 - 2,33, 1,21 - 1,31 e 0,70 - 0,90.

No entanto, a composição química da mistura dos padrões, apesar de ser semelhante, a mesma não coincide exatamente com aquela do óleo diesel, conforme pode ser observado na Figuras 4.5 e 4.6. Desta forma, a hipótese de se construir um modelo de calibração baseado na mistura dos padrões e empregá-lo na previsão do número de cetano de amostras de óleo diesel foi investigada.

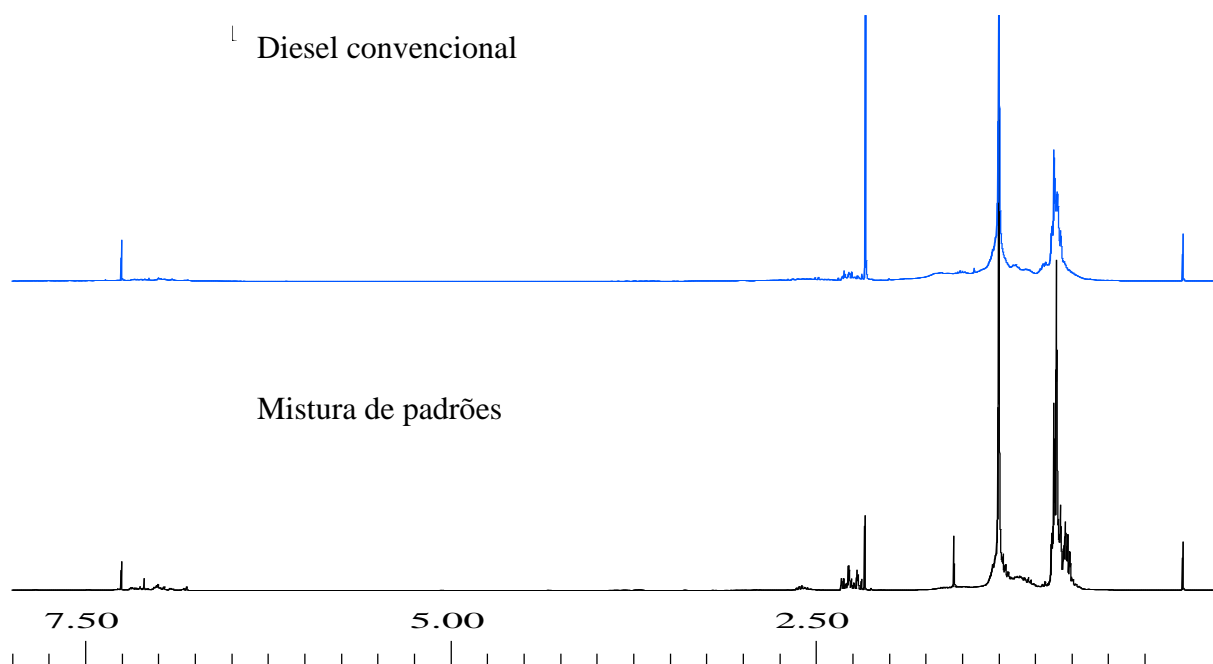


Figura 4.5 – Comparação entre os espectros de RMN de ^1H da mistura dos padrões contendo 45 % de T-23 e 55 % de U-16 e uma amostra de óleo diesel, ambos com número de cetano de 44,8.

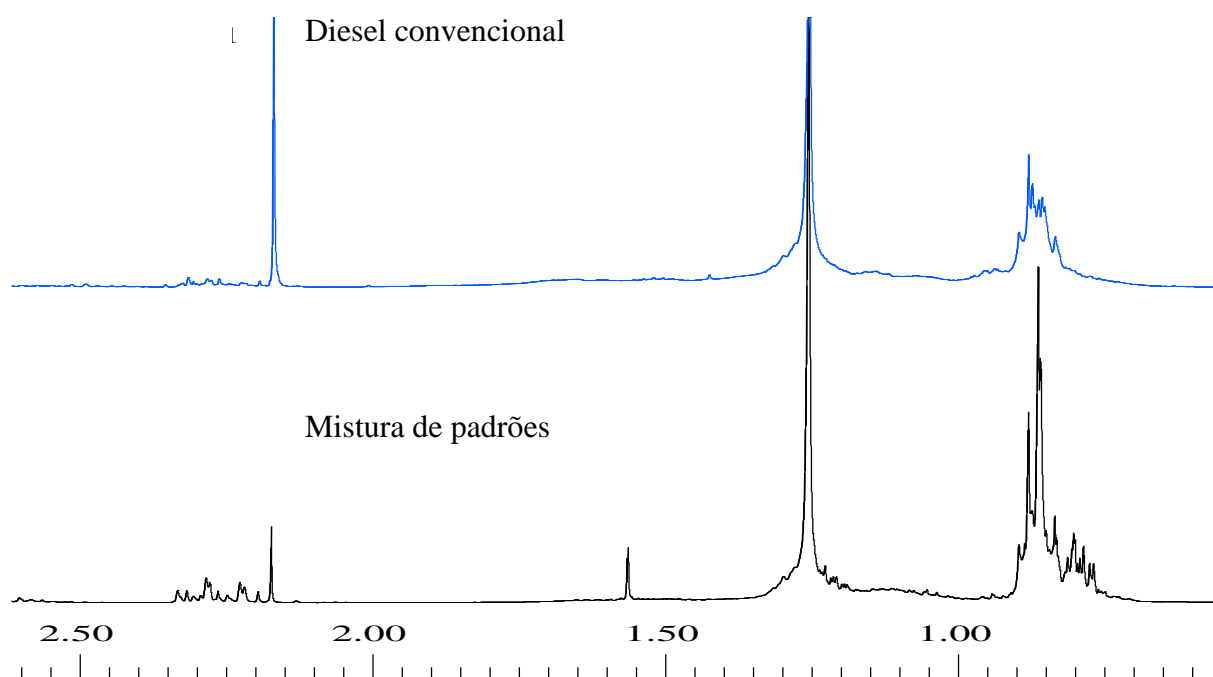


Figura 4.6 – Comparação entre os espectros de RMN de ^1H da mistura dos padrões contendo 45 % de T-23 e 55 % de U-16 e uma amostra de óleo diesel, ambos com número de cetano de 44,8. Ampliação da região de 0,5 - 2,5 ppm.

4.1.2 – Amostras de óleo diesel convencional

Os espectros de RMN de ^1H das amostras de óleo diesel convencional apresentam predominantemente sinais na região alifática, característicos de hidrocarbonetos alifáticos, mas também são observados sinais na região aromática em menor proporção (Figura 4.7). A completa atribuição dos deslocamentos está apresentada na Tabela 4.2.

Tabela 4.2 – Atribuição dos deslocamentos químicos nos espectros de RMN ^1H do óleo diesel convencional.

δ^a (ppm)	Hidrogênio	Descrição
0,68 - 1,00	$\text{CH}_3\text{-C}$	Grupos metilas de cadeias alquílicas lineares ou ramificadas
1,00 - 1,78	$\text{C-(CH}_2\text{)-C}$	Grupos metilênicos
2,07 - 2,74	$\text{-OC-CH}_2\text{-}$ ou -OC-CH_3	Grupos metilênicos ou metilas α carbonilas
6,75 - 8,00	H-Ar	Hidrogênio ligado a carbono aromático

^a Valores obtidos a 400 MHz em CDCl_3 , em relação ao sinal do TMS em δ 0,00, utilizado como referência interna.

A análise visual dos espectros de RMN de ^1H das amostras de óleo diesel com diferentes números de cetano, conforme previamente determinado pelo método especificado pela norma ASTM D613 pela REPAR, permite verificar apenas diferenças sutis no perfil espectral entre as amostras (Figura 4.8 e 4.9). Essas pequenas diferenças são indicativo de que seria possível construir um modelo de calibração multivariada, com auxílio de ferramentas estatísticas.

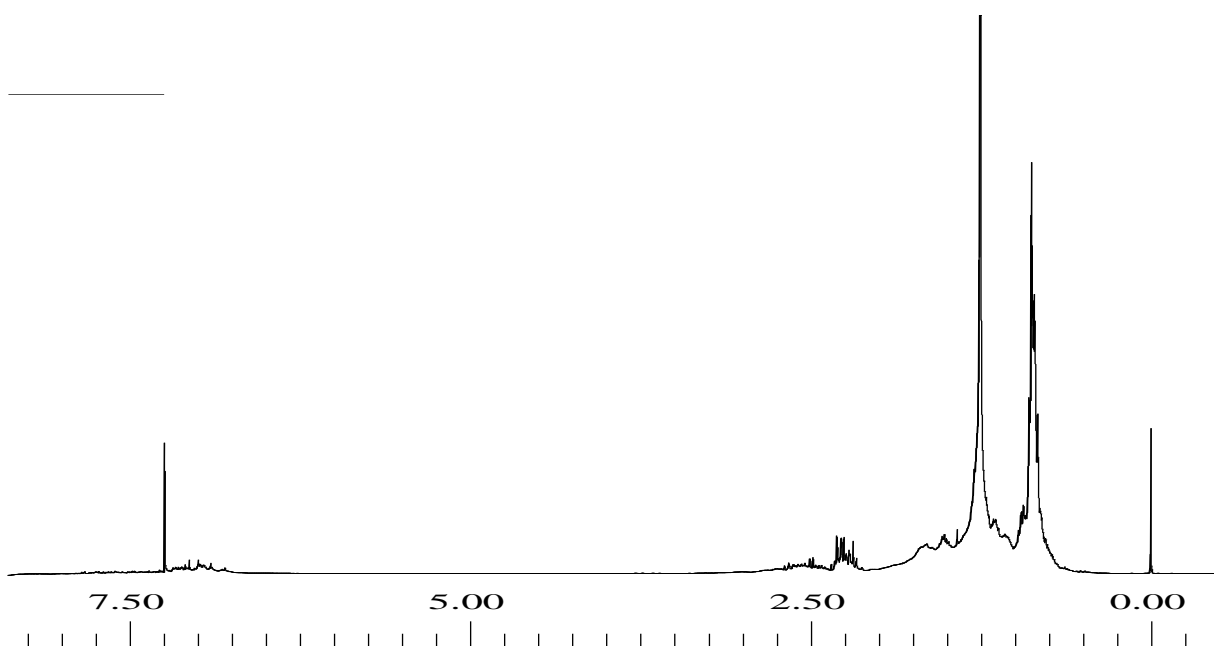


Figura 4.7 – Espectro de RMN de ^1H representativo de uma amostra de óleo diesel convencional com número de cetano igual a 43,7.

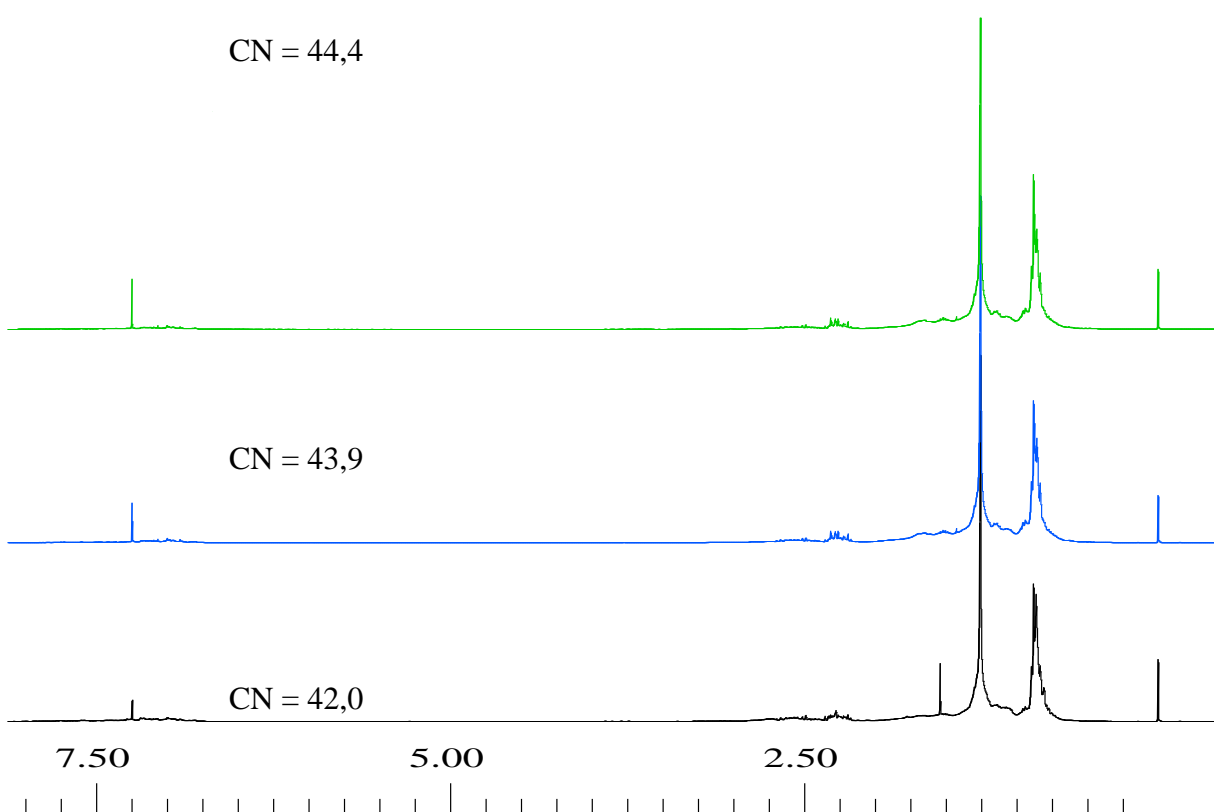


Figura 4.8 – Comparação entre os espectros de RMN de ^1H de amostras de óleo diesel convencional com diferentes números de cetano.

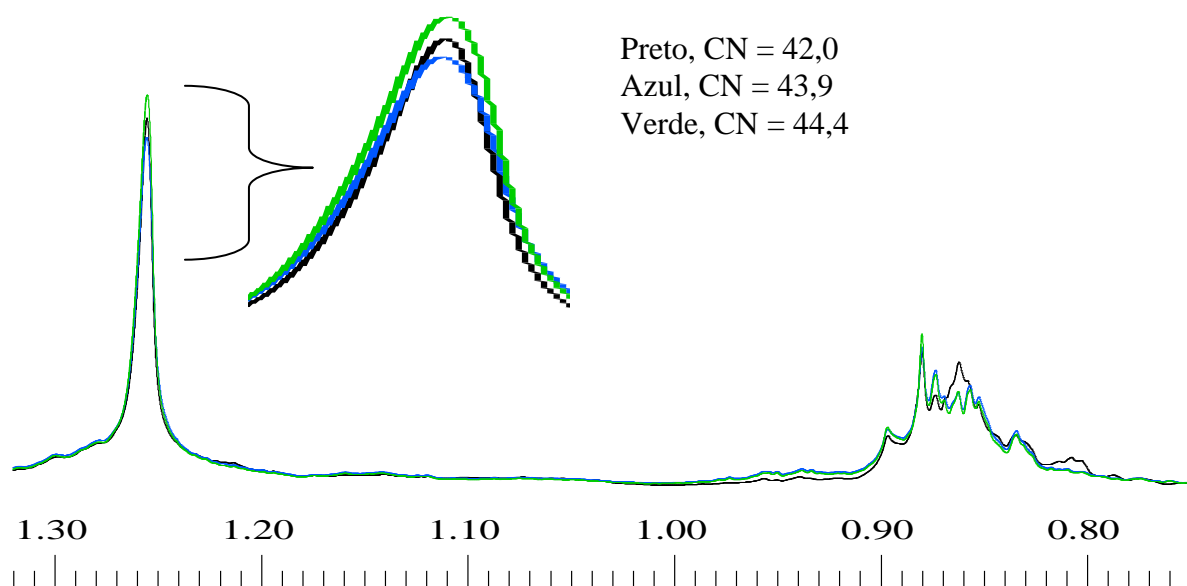


Figura 4.9 – Comparação entre os espectros de RMN de ^1H de amostras de óleo diesel convencional com diferentes números de cetano. Ampliação da região entre 0,70 - 1,32 ppm.

4.2 - Determinação do tamanho dos *buckets*, modo de integração e região do espectro utilizada para construção dos modelos de calibração multivariada

Para aumentar a eficiência das análises quimiométricas de espectros de RMN unidimensionais, os mesmos são convertidos em *buckets*, sob os quais a área é calculada, os quais são as variáveis nas análises quimiométricas, enquanto que os espectros são as amostras. A determinação das variáveis de entrada para construção dos modelos de calibração multivariada foram realizadas em conjunto com auxílio do programa AMIX® da Bruker (Tabelas 4.3 a 4.10).

Todos os modos de integração disponíveis no programa AMIX®, relativos a soma das intensidades dos sinais, soma das intensidades absolutas, intensidades positivas e modo especial de integração utilizados na determinação da área dos *buckets* foram avaliados (Tabelas 4.3 a 4.10). O programa AMIX® também fornece um modo de integração utilizando as intensidades negativas, entretanto, à utilização deste modo de integração excluiria todos os sinais com intensidade positiva no espectro, prejudicando a construção do modelo de calibração e, portanto, não foi utilizado (Manual do AMIX 2004).

Os *buckets* são as variáveis de entrada utilizadas para construção dos modelos de calibração no PLS Toolbox que opera em ambiente MatLab®. Desta forma, os espectros de RMN foram inicialmente convertidos em *buckets* e o tamanho destes está diretamente

relacionado à largura dos sinais e à sobreposição dos mesmos nos espectros de RMN. O emprego de *buckets* estreitos é útil para distinguir sinais muito próximos nos espectros. Porém, se forem excessivamente estreitos, haverá vários *buckets* para representar um mesmo sinal, e, portanto desnecessário, ou seja, não há ganho em informação. Além disso, o emprego de *buckets* muito estreitos não serão eficientes para minimizar os efeitos das pequenas variações que ocorrem nos deslocamentos químicos de uma amostra para outra, devido a alterações na temperatura, pH e falta de homogeneidade do campo magnético. Por outro lado, se forem utilizados *buckets* muito largos haverá um mesmo *buckets* para representar dois ou mais sinais e conseqüentemente poderá haver perda de informação semelhante a uma perda de resolução (Barison, 2005). Neste trabalho foram avaliados *buckets* de, 0,01, 0,05, 0,1 e 0,5 ppm (Tabelas 4.3 a 4.10).

Na tentativa de se verificar qual região dos espectros de RMN seriam mais importantes na construção de modelos de calibração multivariados, diversas regiões dos espectros foram consideradas separadamente, entre elas: 6,75 - 8,00 ppm; 0,5 - 2,75 e 6,75 - 8,00 ppm; 0,5 - 2,75 ppm, bem como todo o espectro de RMN de ^1H , ou seja de 0,5 - 8,00 ppm, para as amostras obtidas a partir da mistura dos padrões e para as amostras de diesel convencional. Em todos os casos foi excluída a região entre 1,50 - 1,65 ppm, uma vez que a água em CDCl_3 apresenta um sinal nesta região.

A avaliação dos modelos foi realizada através do coeficiente correlação linear (R^2) bem como do RMSECV (Root Mean Square Error of Cross Validation – Raiz quadrada da média dos quadrados dos erros da validação cruzada), resultantes da construção de diversos modelos de calibração multivariada. Os valores de coeficiente de correlação linear (R^2) e RMSECV, são apresentados nas Tabelas 4.3 a 4.6, para amostras obtidas a partir da mistura dos padrões T-23 e U-16 e nas Tabelas 4.7 a 4.10 para as amostras de diesel com NC previamente determinado.

Tabela 4.3 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 6,75-8,00 ppm, para os espectros obtidos a partir da mistura dos padrões.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,764 (2,0453)	0,712 (2,1448)	0,724 (2,0683)	0,526 (2,5772)
Soma das intensidades absolutas	0,744 (2,0139)	0,649 (2,2881)	0,647 (2,287)	0,191 (3,0018)
Intensidades positivas	0,719 (2,1336)	0,686 (2,1963)	0,700 (2,1616)	0,552 (2,5175)
Intensidades negativas	*	*	*	*
Modo especial	0,810 (1,7664)	0,722 (2,1093)	0,680 (2,2081)	0,545 (2,5354)

Tabela 4.4 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5-2,75 ppm e 6,75-8,00 ppm, para os espectros obtidos a partir da mistura dos padrões.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,912 (1,2565)	0,891 (1,4102)	0,861 (1,6666)	0,930 (1,1013)
Soma das intensidades absolutas	0,916 (1,2293)	0,894 (1,3902)	0,855 (1,7199)	0,943 (0,98989)
Intensidades positivas	0,911 (1,2601)	0,889 (1,4244)	0,857 (1,6975)	0,933 (1,0783)
Intensidades negativas	*	*	*	*
Modo especial	0,933 (1,0761)	0,917 (1,19820)	0,911 (1,2335)	0,918 (1,1862)

Tabela 4.5 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5-2,75 ppm, para os espectros obtidos a partir da mistura dos padrões.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,918 (1,2155)	0,921 (1,1858)	0,877 (1,572)	0,934 (1,0713)
Soma das intensidades absolutas	0,917 (1,2235)	0,920 (1,1949)	0,876 (1,5783)	0,934 (1,0681)
Intensidades positivas	0,918 (1,2192)	0,921 (1,1903)	0,877 (1,5748)	0,934 (1,0697)
Intensidades negativas	*	*	*	*
Modo especial	0,931 (1,0941)	0,915 (1,2078)	0,912 (1,2301)	0,893 (1,34469)

Tabela 4.6 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5-8,00 ppm, para os espectros obtidos a partir da mistura dos padrões.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,922 (1,1713)	0,922 (1,1631)	0,896 (1,3795)	0,943 (0,99692)
Soma das intensidades absolutas	0,922 (1,1969)	0,930 (1,126)	0,902 (1, 3783)	0,944 (0,9892)
Intensidades positivas	0,917 (1,2094)	0,925 (1,141)	0,899 (1, 3727)	0,942 (1,0009)
Intensidades negativas	*	*	*	*
Modo especial	0,932 (1,0906)	0,917 (1,2002)	0,912 (1,232)	0,920 (1,1741)

Tabela 4.7 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 6,75-8,00 ppm, para os espectros obtidos a partir do óleo diesel convencional.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,769 (0,80449)	0,694 (0,90615)	0,599 (1,0075)	0,445 (1,1259)
Soma das intensidades absolutas	0,855 (0,65136)	0,819 (0,72066)	0,708 (0,89013)	0,093 (1,2676)
Intensidades positivas	0,210 (2,3792)	0,717 (0,89333)	0,465 (1,8294)	0,596 (1,0078)
Intensidades negativas	*	*	*	*
Modo especial	0,822 (0,7242)	0,805 (0,74625)	0,769 (0,80249)	0,368 (1,1722)

Tabela 4.8 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5-2,75 ppm e 6,75-8,00 ppm, para os espectros obtidos a partir do óleo diesel convencional.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,870 (0,62209)	0,853 (0,65988)	0,842 (0,70795)	0,674 (0,93018)
Soma das intensidades absolutas	0,865 (0,63959)	0,852 (0,66733)	0,842 (0,70401)	0,714 (0,8887)
Intensidades positivas	0,824 (0,74592)	0,597 (1,6209)	0,699 (1,1624)	0,694 (0,90628)
Intensidades negativas	*	*	*	*
Modo especial	0,933 (0,45063)	0,907 (0,53285)	0,891 (0,5702)	0,853 (0,65522)

Tabela 4.9 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5-2,75 ppm, para os espectros obtidos a partir do óleo diesel convencional.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,818 (0,73919)	0,796 (0,78772)	0,793 (0,79762)	0,711 (0,88438)
Soma das intensidades absolutas	0,809 (0,75629)	0,784 (0,80883)	0,781 (0,81976)	0,712 (0,88316)
Intensidades positivas	0,814 (0,74545)	0,791 (0,79603)	0,788 (0,80697)	0,711 (0,88376)
Intensidades negativas	*	*	*	*
Modo especial	0,926 (0,47463)	0,868 (0,63472)	0,871 (0,61813)	0,834 (0,69163)

Tabela 4.10 – Valores para coeficiente de correlação linear (R^2) e RMSECV obtidos utilizando-se a região entre 0,5-8,00 ppm, para os espectros obtidos a partir do óleo diesel convencional.

Tipo de integração	Tamanho dos <i>buckets</i>			
	0,01	0,05	0,1	0,5
	R^2 (RMSECV)			
Soma das intensidades	0,873 (0,61641)	0,859 (0,64805)	0,870 (0,63499)	0,724 (0,8662)
Soma das intensidades absolutas	0,871 (0,6187)	0,866 (0,60083)	0,870 (0,61981)	0,733 (0,85554)
Intensidades positivas	0,731 (1,0624)	0,399 (3,4306)	0,391 (3,4914)	0,247 (8,8002)
Intensidades negativas	*	*	*	*
Modo especial	0,928 (0,46622)	0,908 (0,52811)	0,893 (0,56348)	0,857 (0,6474)

A análise das Tabelas 4.3 a 4.6, para as amostras obtidas a partir da mistura dos padrões, permitiu observar que os valores mais elevados para R^2 e mais baixos para RMSECV, são obtidos utilizando *buckets* de 0,5 ppm e a soma das intensidades absolutas, entretanto a utilização de *buckets* tão largos leva a representação de vários sinais por um

mesmo *bucket*. Por outro lado, foi observado que há uma tendência tanto nas amostras obtidas a partir da mistura de padrões quanto no óleo diesel convencional no aumento dos valores para R^2 e diminuição de RMSECV utilizando *buckets* menores aliados ao método especial de integração. O modo especial de integração do programa AMIX[®] da Bruker calcula as áreas dos *buckets* através de um algoritmo que independe do ajuste da fase e da linha de base dos espectros, sendo necessária somente a calibração dos espectros (Barison, 2005; Manual do AMIX 2004). Desta forma são eliminadas todas as subjetividades provenientes da manipulação dos espectros, pelo analista. Analisando a tabela 4.3, para amostras obtidas a partir da mistura dos padrões, pode-se observar que os melhores valores de coeficiente de correlação linear e RMSECV foram obtidos utilizando o modo especial de integração aliado aos menores valores de *buckets*. Estas observações se repetem quando são analisadas as tabelas 4.4 a 4.10. Desta forma, optamos por utilizar este método para construção de todos os modelos.

Em relação às regiões dos espectros avaliados, foi evidenciado melhores coeficientes de correlação linear e RMSECV, quando a região de 0,5 a 2,75 ppm esteve incluída na construção dos modelos de calibração, tanto para as amostras obtidas a partir da mistura dos padrões quanto para as amostras de diesel convencional (Tabelas 4.3 a 4.10 e resumidas na Tabela 4.11). Ao compararmos os modelos construídos com e sem a exclusão da região dos compostos aromáticos, observamos valores de R^2 e de RMSECV bastante próximos, evidenciando a baixa influência da região acima de 2,75 ppm. Desta forma, fica evidente que os compostos alifáticos são os principais responsáveis pelo número de cetano e que os compostos aromáticos possuem pouca influência na construção dos modelos de calibração multivariada. Com base nestes resultados não há motivos para que alguma região dos espectros não fosse considerada, ao contrario se todo espectro for utilizado aumentasse a representatividade da amostra e, portanto, todos os modelos de calibração multivariada foram construídos considerando-se todo o espectro, com exceção da região do sinal da água entre 1,50 – 1,65 ppm, utilizado-se *buckets* de 0,05 ppm, em detrimento aos *buckets* de 0,01 ppm, que apresentaram melhores valores de coeficiente de correlação linear e RMSECV, para evitar que um mesmo sinal seja representado por vários *buckets*.

Tabela 4.11 – Comparação dos coeficientes de correlação linear (R^2) e RMSECV obtidos utilizando-se as diferentes regiões dos espectros. Utilizando o modo especial de integração e *buckets* de 0,05 ppm.

	Região do espectro			
	6,75 - 8,00	0,5 - 2,75 e 6,75 - 8,00	0,5 - 2,75	0,5 - 8,00
	Mistura dos padrões T-23 e U-16			
R^2 (RMSECV)	0,722 (2,1093)	0,917 (1,19820)	0,931 (1,2078)	0,917 (1,2002)
	Diesel de NC previamente determinado			
	0,805 (0,74625)	0,907 (0,53285)	0,868 (0,63472)	0,908 (0,52811)

4.3 – Análises quimiométricas das amostras obtidas a partir da mistura dos padrões

Os espectros de RMN de ^1H , das amostras obtidas a partir da mistura dos padrões T-23 e U-16, após serem convertidos em *buckets* com auxílio do programa AMIX[®] da Bruker, foram submetidos à análise por PLS, utilizando o programa PLS Toolbox que opera em ambiente MatLab[®]. Inicialmente foi avaliado o tipo de pré-processamento através de inúmeras tentativas avaliando-se o menor valor de RMSECV (Root Mean Square Error of Cross Validation – Raiz quadrada da média dos quadrados dos erros da validação cruzada) obtido, aliado ao menor número de variáveis latentes (Tabela 4.12). O melhor resultado foi obtido utilizando-se o auto-escalamento para os valores associados ao bloco X (amostras = espectros) e valores centrados na média para o bloco Y (número de cetano), e regressão dos dados do tipo NIPLS e utilizando o sistema *leave-one-out* para validação cruzada, ideal para modelos com número pequeno de amostras (Tabela 4.12).

O número de variáveis latentes foi determinado analisando o gráfico de variáveis latentes *versus* RMSEC (sem o emprego da validação cruzada) e RMSEVC (com o emprego da validação cruzada), para cada um dos modelos construídos. Neste trabalho a escolha do número de variáveis latentes foi baseada na validação cruzada, através do valor de RMSECV. Em geral, a utilização de um número excessivo de variáveis latentes pode levar a uma situação de super-ajuste do modelo, fazendo com que algumas variáveis

latentes sejam utilizadas, apenas, para descrever anomalias apresentadas pelo conjunto de calibração ou ruídos instrumentais (Zamora *e col.*, 2005).

Vemos que de acordo com o gráfico de número de variáveis latentes *versus* RMSECV, que a utilização de 12 variáveis latentes nos proporcionaria os menores valores de RMSECV (Figura 4.10). Entretanto utilizando um valor tão alto de variáveis latentes introduziríamos erros associados aos valores de concentração bem como a inclusão de ruídos no modelo de calibração. Deste modo analisando o gráfico da figura 4.10, observamos que utilizando 5, 6 ou 7 variáveis latentes que o incremento nos valores de RMSECV aumenta de forma pouco significativa e há uma diminuição importante no número de variáveis latentes.

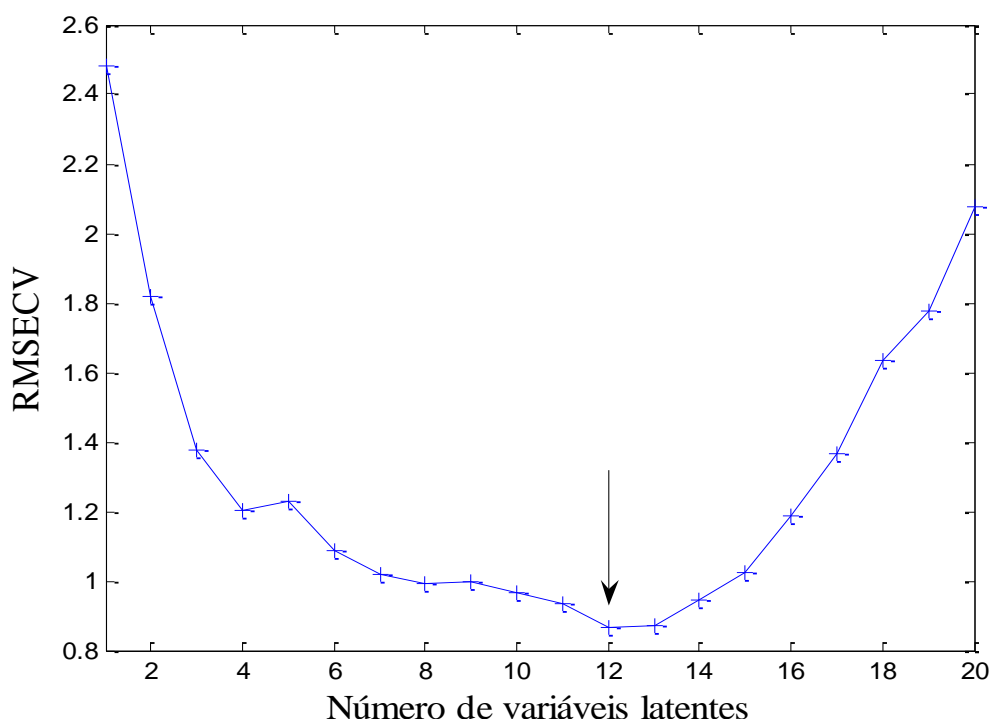


Figura 4.10 - Gráfico do número de variáveis latentes *versus* RMSECV, para as amostras obtidas a partir dos padrões T-23 e U-16.

Partindo deste raciocínio avaliamos a existência de amostras anômalas (*outliers*) ao conjunto de calibração, através do gráfico de Leverage *versus* Resíduos de Students, utilizando 5 variáveis latentes, conforme Figuras 4.11. Observamos na Figura 4.11 que as amostras 8, 9 e 11 apresentam altos valores de leverage. Deste modo, optamos por excluir estas amostras do modelo e verificamos que houve uma diminuição no número de

variáveis latentes de 12 para 10 ou de 5 para 4, conforme anteriormente discutido, bem como a diminuição no valor de RMSECV, conforme Figura 4.12 e Tabela 4.12.

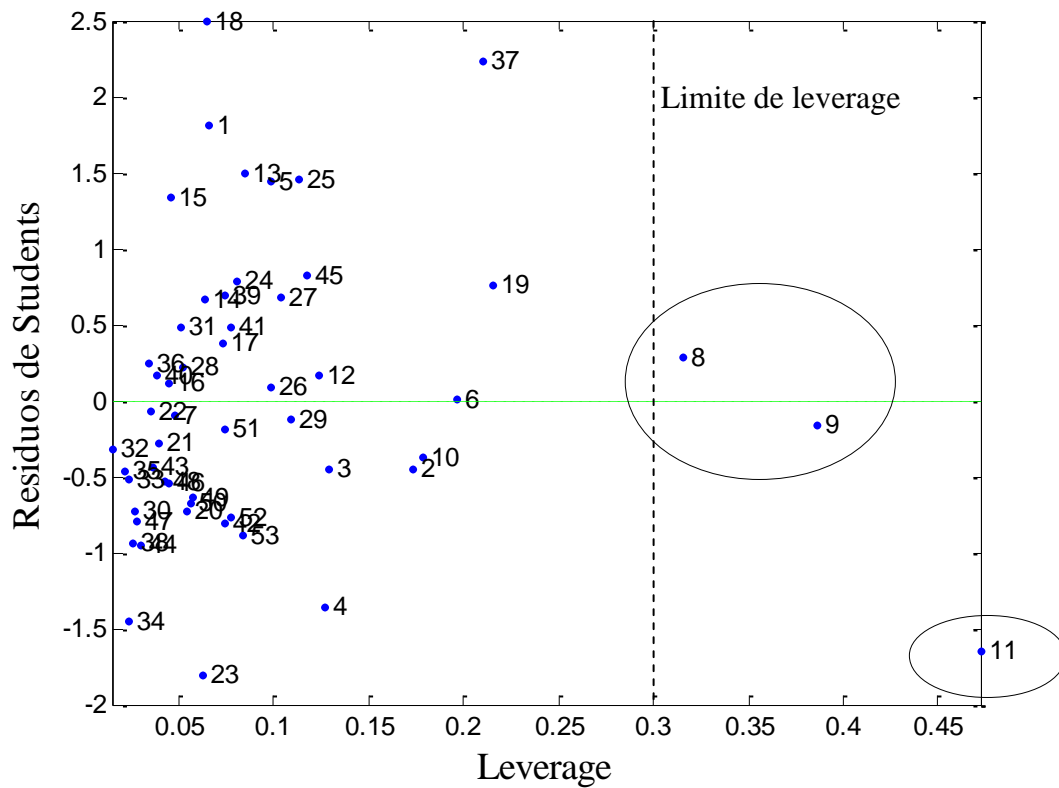


Figura 4.11 - Gráfico de Leverage *versus* Resíduos de Students, utilizando 5 variáveis latentes para construção do modelo, com limite de Leverage de 0,3.

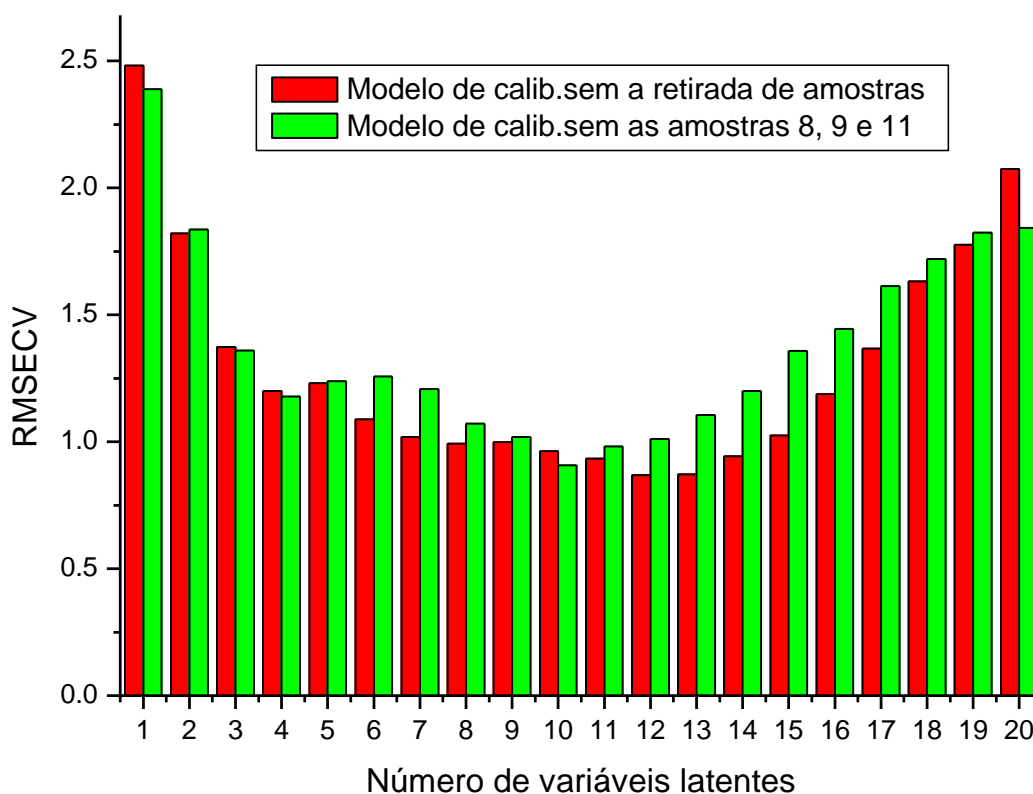


Figura 4.12. Gráfico utilizado para escolha do número de variáveis latentes em função dos valores de RMSECV.

Uma vez removido as amostras 8, 9 e 11, houve o surgimento de novas amostras anômalas (*outliers*), conforme pode ser observado na Figura 4.13. No entanto, como nenhuma destas amostras apresentou simultaneamente alto valor de leverage ($3VL/N$) e resíduos de Students (limite $\pm 2,5$), não foi necessário a remoção das mesmas do modelo de calibração. A Tabela 4.12 apresenta uma comparação entre alguns dos modelos construídos. Entre todos os modelos construídos, observamos que o modelo 4 (Tabela 4.12) forneceu menores valores de RMSEC e RMSECV com uma diminuição muito pequena do coeficiente de correlação linear (R^2).

As figuras 4.14 e 4.15 mostram graficamente o modelo 4 (Tabela 4.12) de calibração multivariada, obtidos sem e com validação cruzada, respectivamente. Nestes gráficos pode ser observado que há uma dependência linear entre número de cetano e o perfil dos espectros de RMN de 1H das amostras obtidas a partir da mistura dos padrões. O alto coeficiente de correlação linear indica que o modelo possui grande capacidade preditiva de novas amostras, pelo menos de mistura dos padrões T-23 e U-16.

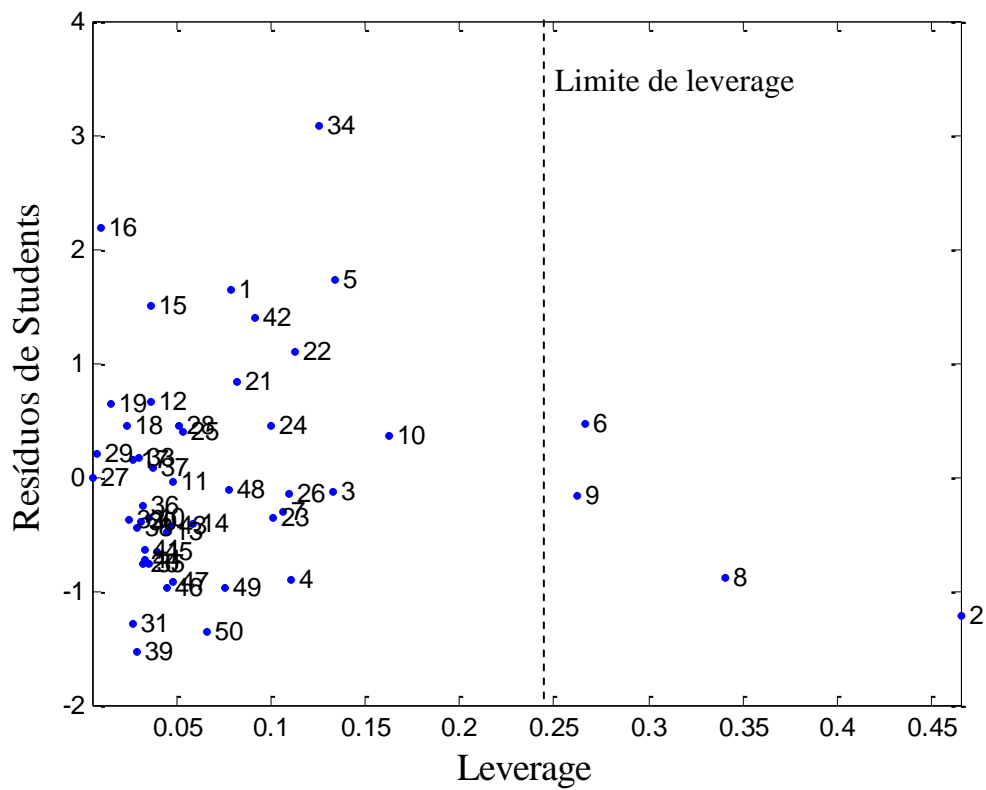


Figura 4.13. Gráfico de Leverage *versus* Resíduos de Students, utilizando 4 variáveis latentes para construção do modelo, com limite de Leverage de 0,24, com a exclusão das amostras 8, 9 e 11.

Tabela 4.12 – Comparação entre os modelos construídos utilizando a mistura de padrões.

Modelo	Transformada/ Pré-processamento	VL(s)	Anomalias	RMSEC	RMSECV	R²(cal)	R²(val)	Observações (amostras retiradas)
1	Auto-escalados, centrado na média, NIPLS, leave-one-out	5	8, 9 e 11 (alto leverage)	0,81859	1,2309	0,962	0,917	
2	Auto-escalados, centrado na média, NIPLS, leave-one-out	6	9 e 11 (alto leverage)	0,64438	1,0883	0,976	0,937	
3	Auto-escalados, centrado na média, NIPLS, leave-one-out	7	8, 9 e 11 (alto leverage)	0,56748	1,0186	0,982	0,944	
4	Auto-escalados, centrado na média, NIPLS, leave-one-out	4	2,6,8 e 9 (alto leverage) 34 (alto resíduo)	0,92139	1,1779	0,950	0,919	8, 9 e 11
5	Auto-escalados, centrado na média, NIPLS, leave-one-out	5	2 e 8 (alto leverage) 34 (alto resíduo)	0,79548	1,2385	0,963	0,916	8, 9 e 11
6	Auto-escalados, centrado na média, NIPLS, leave-one-out	6	2 e 8 (alto leverage) 34 (alto resíduo)	0,60825	1,2564	0,978	0,918	8, 9 e 11
7	Auto-escalados, centrado na média, NIPLS, leave-one-out	7	2, 6 e 8 (alto leverage)	0,56272	1,2077	0,982	0,924	8, 9 e 11

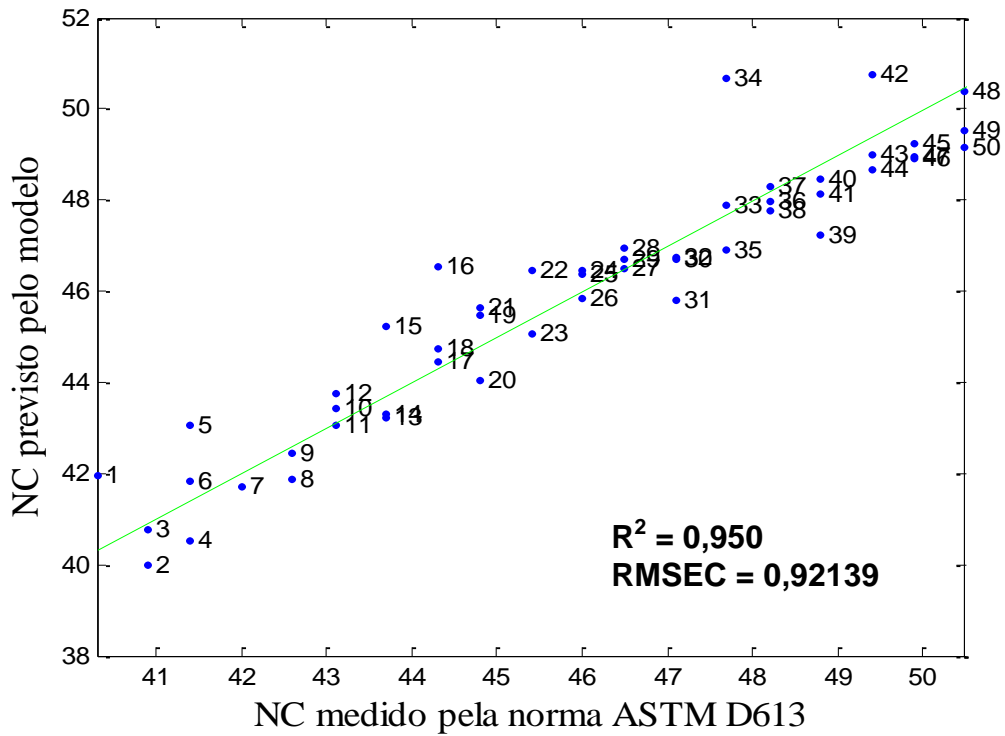


Figura 4.14 - Número de cetano medido *versus* previsto sem validação cruzada, utilizando 4 variáveis latentes para construção do modelo.

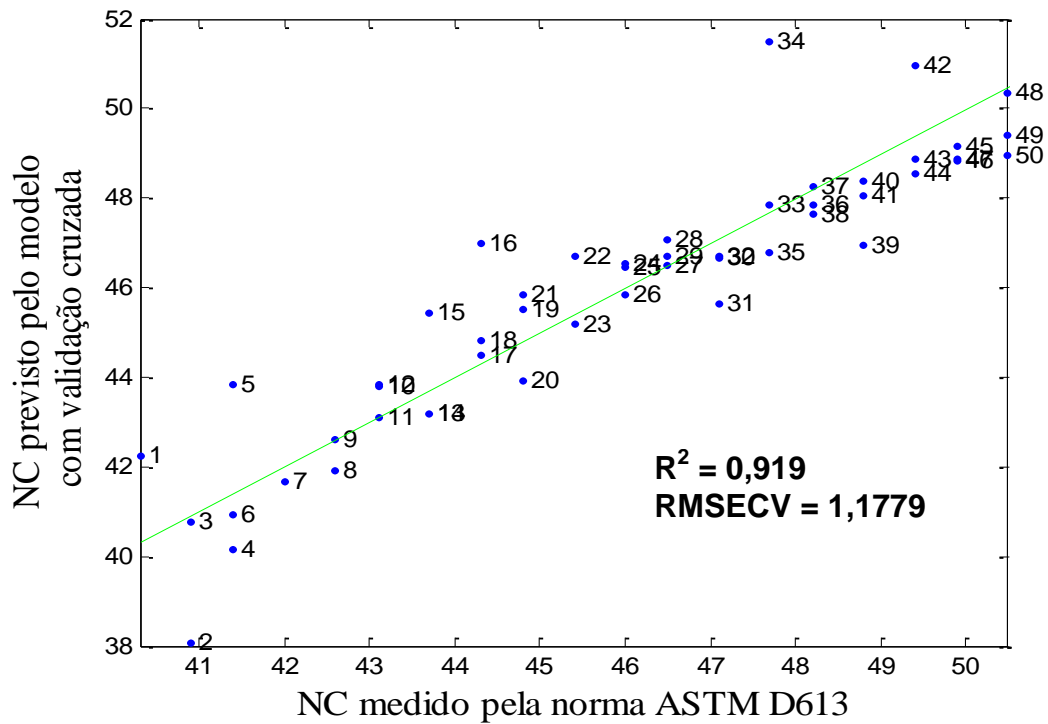


Figura 4.15 - Número de cetano medido *versus* previsto com validação cruzada, utilizando 4 variáveis latentes para construção do modelo.

O modelo de calibração 4 (Tabela 4.12) com validação cruzada e um coeficiente de correlação linear (R^2) de 0,919 (Figura 4.15) foi então utilizado para a previsão do número de cetano de cerca de 10 amostras de óleo diesel convencional com número de cetano previamente determinado pela REPAR através do método especificado pela norma ASTM D613.

O emprego do modelo de calibração sobre os espectros de RMN de ^1H permitiu prever o número de cetano das amostras com razoável precisão (Figura 4.16). Uma vez que número de cetano fornecido pelo método de referência conforme a norma ASTM D613 possui um desvio padrão de $\pm 2,0$, somente poucas amostras foram previstas pelo modelo fora do limite de $\pm 2,0$, conforme mostra a Figura 4.16. Desta forma fica evidente que um modelo de calibração pode ser construído a partir da mistura dos padrões T-23 e U-16 e que o mesmo pode ser empregado na previsão do número de cetano do óleo diesel convencional. Além disso, não é necessário, o uso de amostras de óleo diesel, com número de cetano previamente determinado pelo método de referência, para a construção de um modelo de calibração. Sendo assim, elimina-se a necessidade do emprego de um motor a combustão. Estes resultados demonstram a versatilidade da espectroscopia de RMN, que poderá vir a ser empregada também para esta finalidade. O custo de um espectrômetro de RMN é similar ao de um equipamento utilizado para a determinação do número de cetano pelo método de referência. No entanto, ao passo que o equipamento para a determinação do número de cetano tem somente esta utilidade, um espectrômetro de RMN em uma refinaria de petróleo tem inúmeras outras utilidades.

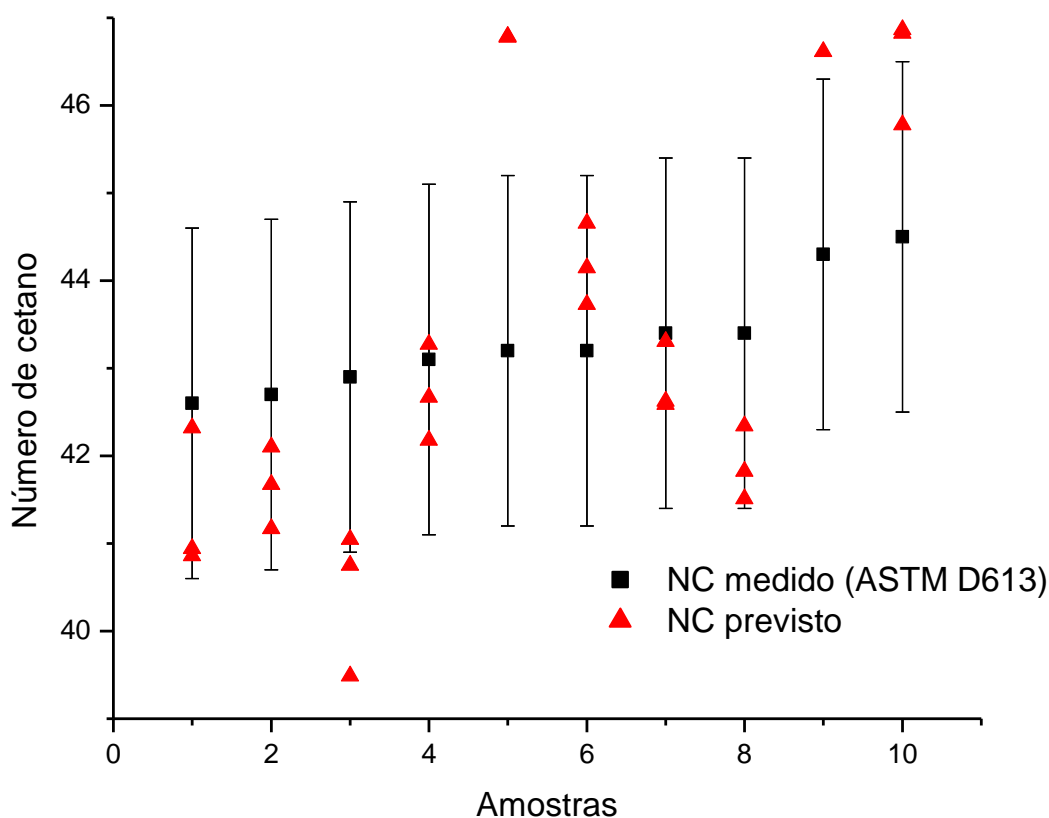


Figura 4.16 – Comparação entre o número de cetano medido pelo método de referência *versus* previsto pelo modelo de calibração construído a partir da mistura dos padrões T-23 e U-16. As barras indicam o desvio padrão de $\pm 2,0$ NC.

4.4 – Análises quimiométricas das amostras de óleo diesel

A construção de modelos de calibração a partir das amostras de óleo diesel convencional com número de cetano previamente determinado pela REPAR através do método estabelecido pela norma ASTM D613 foi realizada com os mesmos ajustes utilizados na análise da mistura dos padrões. Da mesma forma que na análise da mistura dos padrões, diversos pré-processamentos foram testados antes de realizar a construção do modelo de calibração (Tabela 4.13). Nestes experimentos, foi avaliado o valor de RMSECV obtido, aliado ao menor número de variáveis latentes. Após as diversas tentativas foram obtidos melhores resultados utilizando-se o auto-escalamento para os valores associados ao bloco X (amostras), e valores centrados na média para o bloco Y (número de cetano), regressão dos dados do tipo NIPLS e sistema *leave-one-out* para

validação cruzada, ideal para modelos com número pequeno de amostras. Portanto os mesmos tipos de pré-processamento utilizados na análise da mistura dos padrões.

O número de variáveis latentes foi determinado analisando o gráfico de variáveis latentes *versus* RMSEVC conforme Figura 4.17. Deste modo analisando o gráfico da figura 4.17, observamos que utilizando 4 a 7 variáveis latentes observam-se os menores valores de RMSECV. Portanto, a utilização de 4 variáveis latentes propicia o menor valor de RMSECV, além de minimizar a introdução de erros associados a concentração e inclusão de ruídos no modelo.

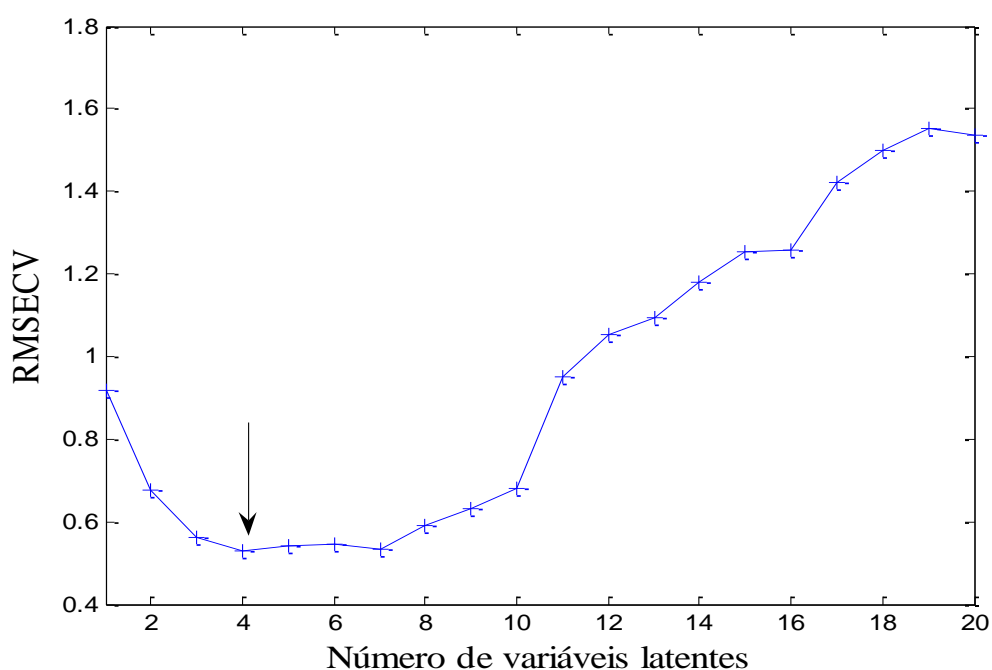


Figura 4.17 - Número de variáveis latentes *versus* RMSECV, para as amostras de diesel convencional com número de cetano previamente determinado pela norma ASTM D 613.

Após a determinação do número ideal de variáveis latentes, utilizamos o gráfico de Leverage *versus* Resíduos de Students para verificação da existência de amostras anômalas (*outliers*) ao conjunto de calibração, conforme Figura 4.18. Observamos a existência de 6 amostras anômalas ao conjunto de calibração, ou seja, com altos valores de Leverage e/ou Resíduos de Students. Desta forma, optamos então em excluí-las em conjunto e em separado para observar a ocorrência de diminuição no número de variáveis latentes concomitante a melhoria no valor de RMSECV. Os valores obtidos com a exclusão das amostras estão descritos na Tabela 4.13, sendo que esta tabela também fornece os valores

obtidos na construção do modelo sem a exclusão da referidas amostras anômalas, para comparação entre os diversos modelos.

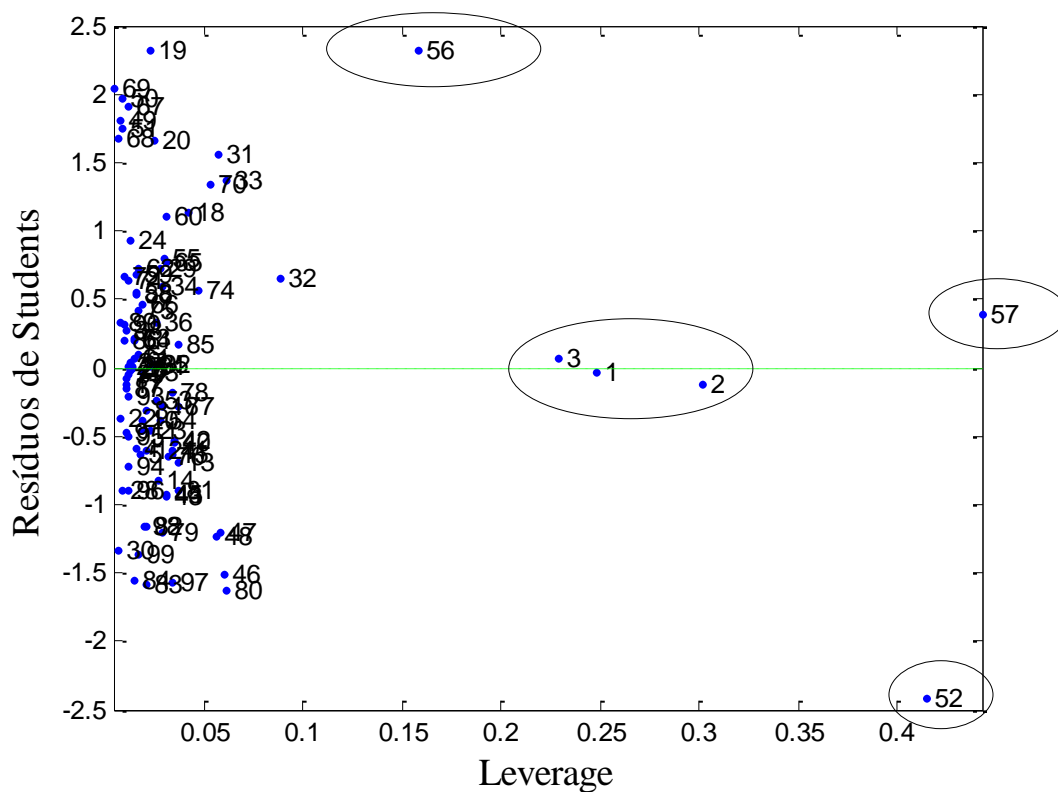


Figura 4.18 - Gráfico de Leverage *versus* Resíduos de Students, utilizando 4 variáveis latentes para construção do modelo, com limite de Leverage de 0,12.

Tabela 4.13 - Comparação dos valores obtidos para os modelos construídos com diferentes valores de variáveis latentes para o diesel convencional.

Modelo	Transformada/ Pré-processamento	VL(s)	Anomalias	RMSEC	RMSECV	R ² (cal)	R ² (val)	Observações (amostras retiradas)
1	Auto-escalados, centrado na média, NIPLS, leave-one-out	4	1,2,3,52,56,57	0,40575	0,52811	0,946	0,908	
2	Auto-escalados, centrado na média, NIPLS, leave-one-out	5	1,2,3,52 (alto leverage, alto resíduo),56,57	0,3611	0,54272	0,958	0,905	
3	Auto-escalados, centrado na média, NIPLS, leave-one-out	6	1,2,3,52 e 56 (alto leverage, alto resíduo),57	0,32222	0,54495	0,966	0,905	
4	Auto-escalados, centrado na média, NIPLS, leave-one-out	3	50, 54 e 55 (alto leverage)	0,42407	0,52923	0,938	0,903	1, 2 e 3
5	Auto-escalados, centrado na média, NIPLS, leave-one-out	4	50 (alta leverage e resíduo) 54 e 55 (alto leverage)	0,38355	0,55162	0,950	0,896	1,2 e 3
6	Auto-escalados, centrado na média, NIPLS, leave-one-out	5	50 e 54 (alta leverage e resíduo), 55 (alta leverage)	0,33616	0,57138	0,962	0,892	1,2 e 3
7	Auto-escalados, centrado na média, NIPLS, leave-one-out	2	1, 2 e 3 (alto leverage)	0,4996	0,53305	0,915	0,903	52, 56 e 57
8	Auto-escalados, centrado na média, NIPLS, leave-one-out	3	1, 2 e 3 (alto leverage)	0,42325	0,4859	0,940	0,920	52, 56 e 57
9	Auto-escalados, centrado na média, NIPLS, leave-one-out	4	1, 2 e 3 (alto leverage)	0,3626	0,44984	0,956	0,932	52, 56 e 57
10	Auto-escalados, centrado na média, NIPLS, leave-one-out	5	1, 2 e 3 (alto leverage)	0,32603	0,43498	0,965	0,937	52, 56 e 57
11	Auto-escalados, centrado na média, NIPLS, leave-one-out	6	1, 2 e 3 (alto leverage)	0,29038	0,41656	0,972	0,942	52, 56 e 57
12	Auto-escalados, centrado na média, NIPLS, leave-one-out	4	Sem anomalias	0,34058	0,43801	0,958	0,930	1, 2, 3, 52, 56 e 57
13	Auto-escalados, centrado na média, NIPLS, leave-one-out	5	Sem anomalias	0,30969	0,42388	0,966	0,933	1, 2, 3, 52, 56 e 57

Entre os diversos modelos, observamos que o modelo 7 (Tabela 4.13) obtido com a exclusão somente das amostras 52, 56 e 57, que apresentaram altos valores de leverage e resíduos de Students, apresentou um valor de RMSECV de 0,53305 o que representa um aumento em relação ao modelo em que estas três amostras não foram excluídas de 0,00494. Entretanto obteve-se a diminuição duas variáveis latentes para construção do modelo, enquanto que o coeficiente de correlação linear permaneceu praticamente inalterado. Inicialmente, sem a exclusão destas três amostras obtivemos um valor de RMSECV menor utilizando quatro variáveis latentes, já com a exclusão das amostras 1, 2 e 3, que também apresentaram altos valores de leverage, obtivemos uma diminuição de uma variável latente com um valor de RMSECV bastante semelhante ao obtido quando excluídas as amostras 52, 56 e 57, com diminuição de duas variáveis latentes (Tabela 4.13).

As figuras 4.19 e 4.20 mostram graficamente o modelo 7 (Tabela 4.13) de calibração multivariada, obtidos sem e com validação cruzada, respectivamente. Nestes gráficos pode ser observado que também há uma dependência linear entre o número de cetano e o perfil espectral das amostras de óleo diesel. O alto coeficiente de correlação linear indica que o modelo possui grande capacidade preditiva de novas amostras de óleo diesel.

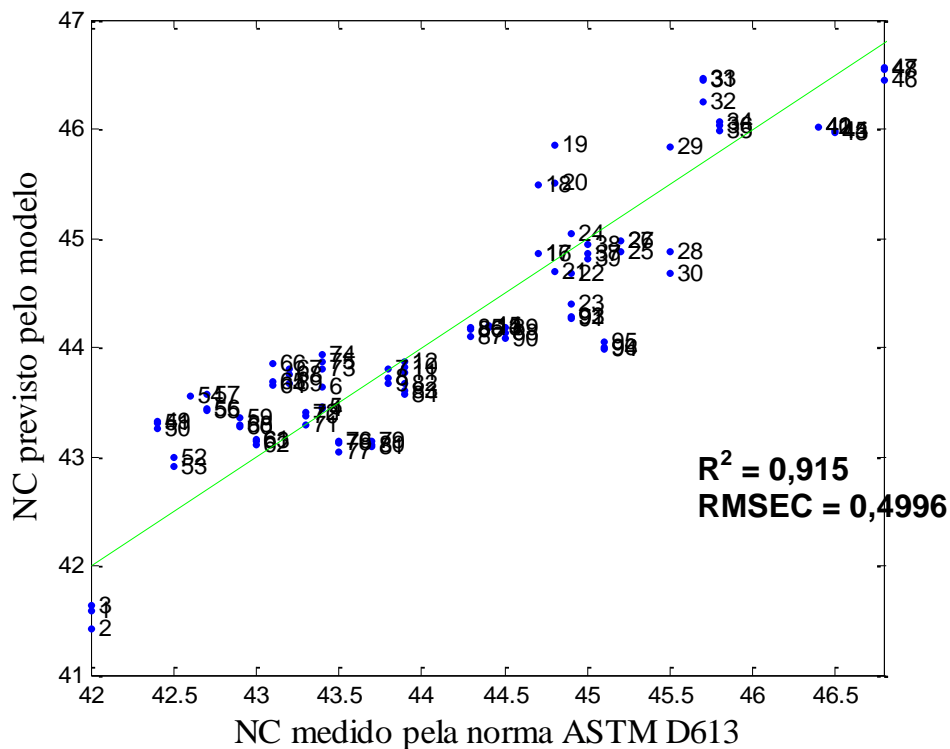


Figura 4.19 - Número de cetano medido *versus* previsto sem validação cruzada, utilizando 2 variáveis latentes para construção do modelo.

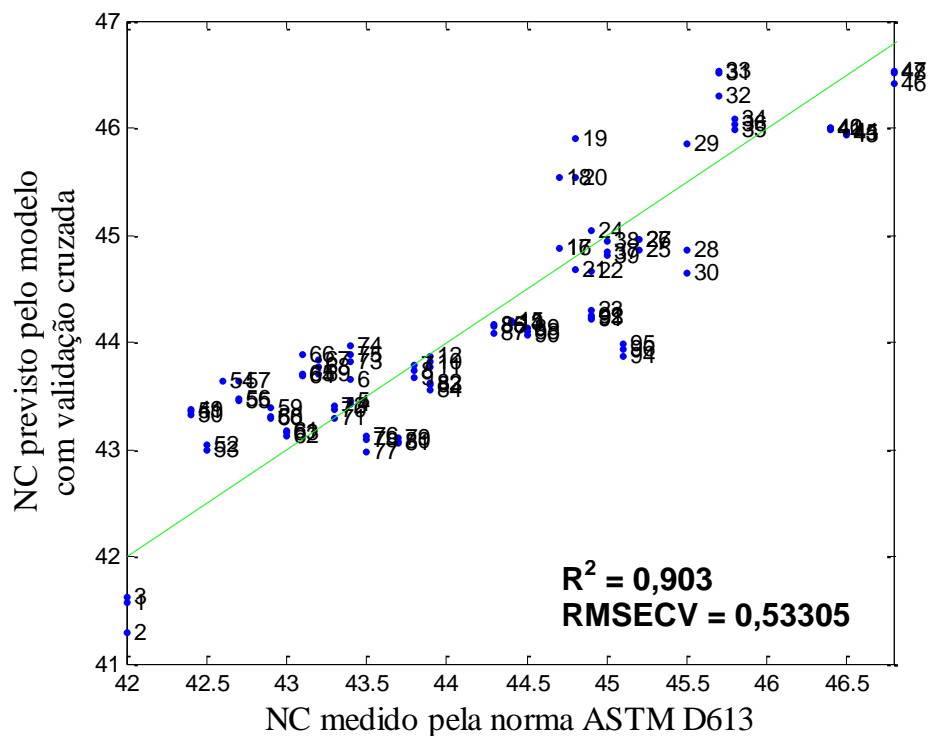


Figura 4.20 - Número de cetano medido *versus* previsto com validação cruzada, utilizando 2 variáveis latentes para construção do modelo.

O modelo de calibração obtido com validação cruzada e a exclusão de somente três amostras e correlação linear (R^2) de 0,903 (Figura 4.20 e Tabela 4.13) foi então utilizado para a previsão do número de cetano de cerca de 10 amostras, com número de cetano previamente determinado pelo método de referência (ASTM D613). A aplicação do modelo de calibração sobre os espectros de RMN de ^1H permitiu prever o número de cetano das amostras, com alta precisão (Figura 4.21). Novamente, como número de cetano fornecido pelo método de referência possui um desvio padrão de $\pm 2,0$, todas as amostras foram previstas pelo modelo dentro do limite de $\pm 2,0$, conforme mostra a Figura 4.21. Portanto, o modelo construído a partir das amostras de óleo diesel apresentou-se mais preciso na previsão de novas amostras que o modelo de calibração multivariada obtido a partir da mistura dos padrões T-23 e U-16. No entanto, este procedimento necessita de um conjunto de calibração com número de cetano previamente determinado pelo método de referência, não eliminando assim a necessidade do equipamento utilizado pelo método de referência. Porém, um único equipamento, seria suficiente para fornecer conjuntos de amostras de calibração para todo o mundo.

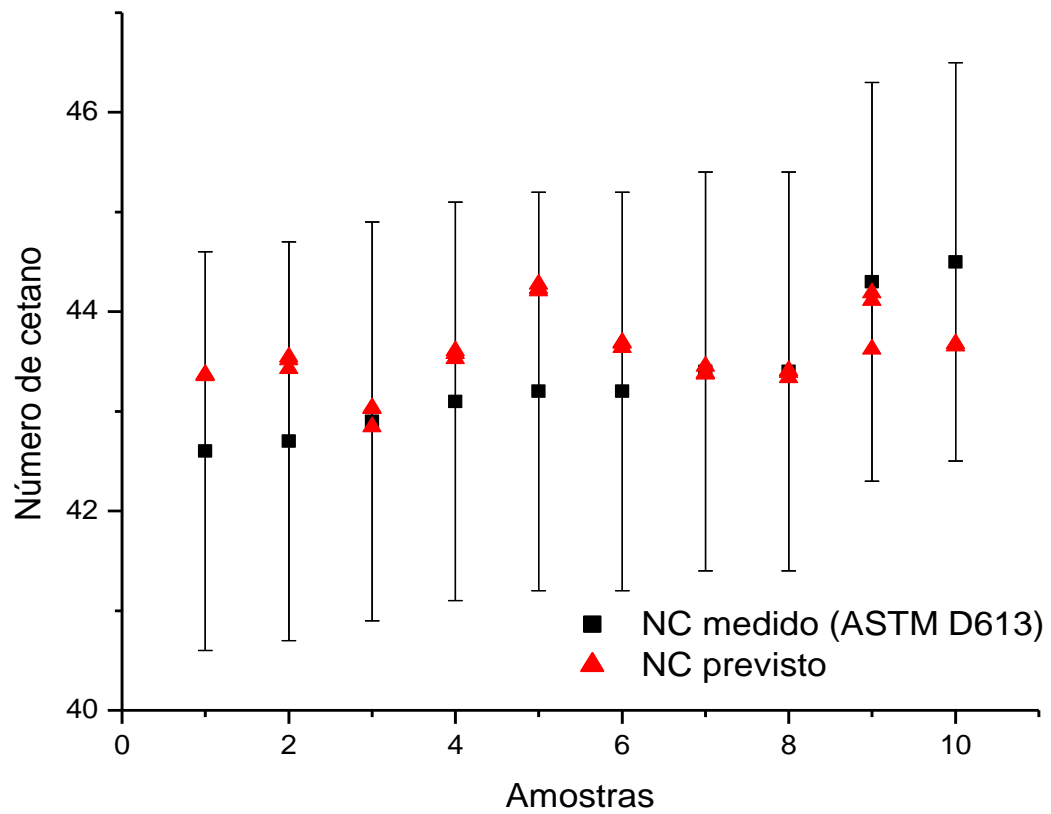


Figura 4.21 – Comparação entre o número de cetano medido pelo método de referência *versus* previsto pelo modelo de calibração construído a partir das amostras de diesel. As barras indicam o desvio padrão de $\pm 2,0$ NC.

CONCLUSÕES

5 - CONCLUSÕES

A espectroscopia de RMN juntamente com a quimiometria mostraram ser ferramentas extremamente úteis no controle de qualidade de combustíveis, especificamente para a determinação do número de cetano do óleo diesel. Estes resultados demonstram a espectroscopia de RMN, poderá vir a ser utilizada para uma nova finalidade entre, as inúmeras outras a qual já é empregada na investigação do petróleo. Através de modelos de calibração multivariados construídos a partir de padrões hidrocarbonetos comumente utilizados nas refinarias de petróleo para determinação do número de cetano, bem como de amostras com número de cetano previamente determinado pelo método de referência, foi possível determinar o número de cetano de amostras de óleo diesel.

O modelo de calibração multivariada obtido a partir dos padrões T-23 e U-16 foi capaz de prever corretamente o número de cetano de 85 % de amostras de óleo diesel. Sua principal vantagem está no fato de não ser necessário, um conjunto de calibração com amostras de óleo diesel com número de cetano previamente determinado pelo método de referência, para a construção de um modelo de calibração, eliminando totalmente a necessidade do emprego de um motor a combustão. Este fato deve-se, provavelmente a diferenças na composição química entre os padrões T-23 e U-16 e o óleo diesel. Portanto, o seu emprego na determinação do número de cetano em amostras de diesel ainda depende de novas investigações científicas.

Por outro lado, o modelo de calibração multivariada obtido a partir das amostras de óleo diesel com número de cetano previamente determinado pelo método de referência, foi capaz de prever corretamente o número de cetano de todas as novas amostras dentro do intervalo de precisão aceitável, configurando-se mais eficiente na previsão do número de cetano. No entanto, este modelo necessita de um conjunto de calibração com amostras de óleo diesel com número de cetano previamente determinadas pelo método de referência, ao passo que obtido a partir da mistura dos padrões T-23 e U-16 elimina totalmente esta necessidade.

Referências bibliográficas

6 – REFERÊNCIAS BIBLIOGRÁFICAS

1. Agência Nacional do Petróleo Gás Natural e Biocombustíveis. Resolução 15, 2006.
2. ASTM D613 Cetane Number Diesel.
3. BARISON, A. Emprego da RMN HR-MAS e análises quimiométricas no reconhecimento e avaliação de soja (*Glycine max*) geneticamente modificada. Tese de doutorado, Universidade Federal de São Carlos, 2005.
4. BARTHUS, R.C.; MAZO, L.H.; POPPI, R.J. Aplicação de métodos de calibração multivariada para determinação simultânea de riboflavina (VB2), tiamina (VB1), piridoxina (VB6) e nicotinamida (VPP). *Química Nova*, 30, nº 7, 1638-1643, 2007.
5. BEEBE, K.; PELL, R.; SEAHOLTS, N. B. *Chemometric: A practical guide*. Ed. John Wiley & Sons, New York, 1998.
6. 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. *Química Nova*, 27, nº 6, 1004-1011, 2004.
7. BRAUN, S.; APPEL, L. G.; SCHMAL, M. A poluição gerada por máquinas de combustão interna movidas à diesel - a questão dos particulados. Estratégias atuais para a redução e controle das emissões e tendências futuras. *Química Nova*, 27, nº.3, 472-482, 2004.
8. BRERETON, R. G. *CHEMOMETRICS: Data Analysis for the Laboratory and Chemical Plant*, John Wiley & Sons 2002.
9. CAMPOS, A.C.; LEONTSINIS, E. *Petróleo e derivados*. JR Ed. Técnica Ltda, Rio de Janeiro, 1989.
10. CARDOSO, L.C. *Petróleo: do poço ao posto*. Ed. Qualitymark, Rio de Janeiro, 2005.
11. CLARIDGE, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*. Ed. Pergamon, Amsterdam, 1999.
12. FERREIRA, M.M.C.; ANTUNES, A.M.; MELGO, M.S.; VOLPE, P.L.O. Quimiometria I: calibração multivariada, um tutorial. *Química Nova*, 22, nº 5, 724-731, 1999.
13. GELADI, P. *Chemometrics in spectroscopy. Part 1. Classical Chemometrics*. *Spectrochimica Acta Part B*, 58, 767-782, 2003.

14. GELADI, P.; SETHSON, B.; NYSTRÖM, J.; LILLHONGA, T.; LESTANDER, T.; BURGER, J. Chemometrics in spectroscopy. Part 2. Examples. *Spectrochimica Acta Part B*, 59, 1347-1357, 2004.
15. GHOSH, P. Predicting the effect of cetane improvers on diesel fuels. *Energy and Fuels*, 22, 1073-1079, 2008.
16. KAPUR, G.S.; ECKER, A.; MEUSINGER, R. Establishing quantitative structure-property relationships (QSPR) of diesel samples by proton-NMR and linear regression (MLR) analysis. *Energy and Fuels*, 15, 943-948, 2001.
17. KNOTHE, G.; GERPEN, J.V.; RAMOS, L.P. *Manual do Biodiesel*. Ed. Edgard Blucher, São Paulo, 2005.
18. Manual do programa AMIX[®] de análises quimiométricas. Bruker Biospin, Karlsruhe, 2004.
19. MONTEIRO, M.R.; AMBROZIN, A.R.P.; FERREIRA, A.G. Determination of biodiesel blend levels in different diesel sample by ¹H NMR. *Fuel*, 88, 691-696, 2009.
20. MONTEIRO, M.R.; AMBROZIN, A.R.P.; SANTOS, M.S.; BOFFO, E.F.; FILHO, E.R.P.; LIÃO, L.M.; FERREIRA, A.G. Evaluation of biodiesel-diesel blends quality using ¹H NMR and chemometrics. *Talanta*, 78, 600-664, 2009.
21. MORGANO, A.S.; FARIA, C.G.; FERRÃO, M.F.; FERREIRA, M.M.C. Determinação de açúcar total em café cru por espectroscopia no infravermelho próximo e regressão por mínimos quadrados parciais. *Química Nova*, 30, nº 2, 346-350, 2007.
22. NETO, B.B.; SCARMINIO, I.S.; BRUNS, R.E. 25 Anos de quimiometria no Brasil. *Química Nova*, 29, nº 6, 1401-1406, 2006.
23. OBERT, E.F. *Motores de combustão interna*. Ed. Globo, Porto Alegre, 1971.
24. OTTO, M. *Chemometrics*. Weinheim, Wiley, 1999.
25. ROBBINS, W. E.; AUDETTE, R. R.; REYNOLDS, N. E. *SAE Q. Trans.* 1951, 5, 404-417.
26. SANTANA, R. C.; DO, P.T.; SANTIKUNAPORN, M.; ALVAREZ, W. E.; TAYLOR, J. D.; SUGHRUE, E. L.; RESASCO, D. R. Evaluation of different reaction strategies for the improvement of cetane number in diesel fuels. *Fuel*, 85, 643-656, 2006.

27. SCARMINIO, I.S.; ISHIKAWA, D.N.; BARRETO, W.J.; PACZKOWSKI, E.L.; ARRUDA, I.C. Calibração multivariada para sistemas com bandas sobrepostas através da análise de fatores do tipo Q. *Química Nova*, 21, nº 5, 590-596, 1998.
28. SILVERSTEIN, R.M.; WEBSTER, F. X.; KIEMLE, D. *Spectrometric identification of organic compounds*. 7th ed. New York, John Wiley & Sons, 2005.
29. SOUZA, T.B. Revisão da equação de cálculo de índice de cetano para as características do diesel comercializado no Paraná. Dissertação de mestrado, Universidade Federal do Paraná, 2008.
30. SPEIGHT, J.G. *The Chemistry and Technology of Petroleum*, forth edition, CRC Press, Boca Raton, 2007.
31. VALDERRAMA, P. Avaliação das figuras de mérito em calibração multivariada na determinação de parâmetros de controle de qualidade em indústria alcooleira por espectroscopia de infravermelho próximo. Dissertação de mestrado, Universidade Estadual de Campinas, 2005.
32. VALDERRAMA, P.; BRAGA, J.W.B.; POPPI, R.J. Estado da arte e figuras de mérito em calibração multivariada. *Química Nova*, 32, nº 5, 1278-1287, 2009.
33. WOLD, S.; TRYGG, J.; BERGLUNG, A.; ANTTI, H. Some recent developments in PLS modeling. *Chemometrics and intelligent laboratory systems*, 58, 131-150, 2001.
34. ZAMORA, P.P.; CORDEIRO, G.; NAGATA, N. Utilização de regressão multivariada para avaliação espectrofotométrica da demanda química de oxigênio em amostras de relevância ambiental. *Química Nova*, 28, No. 5, 838-841, 2005.