

2

Revisão Bibliográfica

Historicamente, o grande impulso das aplicações da espectroscopia no infravermelho como ferramenta analítica ocorreu durante a segunda guerra mundial. Nessa ocasião, ela foi usada no controle de qualidade em várias indústrias químicas. Após o desenvolvimento da instrumentação da cromatografia, a espectroscopia no infravermelho restringiu-se, basicamente, a aplicações qualitativas, mas, com a consolidação das técnicas quimiométricas, este quadro vem sendo revertido nos últimos anos. A calibração multivariada, em especial, viabilizou o desenvolvimento de novas metodologias de análise, envolvendo técnicas espectroscópicas com resultados tão bons quanto os obtidos por cromatografia (3). Além disso, a união entre a calibração multivariada e a espectroscopia vem diminuindo o tempo e o custo das análises.

A espectroscopia vibracional é um poderoso instrumento para a determinação da composição de gasolinas. Por exemplo, espectros infravermelho e Raman podem fornecer informações sobre a razão entre os grupos metil e metileno numa amostra e, conseqüentemente, estimar o seu grau de ramificação. Além disso, esses espectros podem ser utilizados para a detecção da presença de aditivos como benzeno, tolueno e etanol.

Existem diferentes metodologias disponíveis para determinação de compostos orgânicos em misturas complexas como a gasolina, sendo a espectrometria na região do infravermelho que vem demonstrando ser a mais adequada para laboratórios industriais, devido a sua simplicidade e a quantidade de informação fornecida por ela.

O espectro de uma gasolina é dependente de sua composição e esta, por sua vez, vai determinar suas propriedades físico-químicas. É possível, portanto, construir modelos quantitativos com os quais seja possível estimar MON e RON, por exemplo, através de espectros de gasolinas. Desta forma, novas metodologias de avaliação desses parâmetros foram desenvolvidas, com base em analisadores de composição que operam na faixa do infravermelho próximo (NIR), cujo objetivo principal é possibilitar a realização de análises de campo pelo uso de analisadores portáteis (8).

Entretanto, várias razões podem ser listadas para que o infravermelho médio (MIR) seja a região do espectro escolhida para medidas diretas de gasolinas. Entre elas estão a simplicidade, técnicas bem conhecidas, disponibilidade na maioria dos laboratórios industriais e alta sensibilidade (31).

Vários artigos já foram publicados descrevendo o uso da espectroscopia na região do infravermelho próximo para a determinação de propriedades de gasolina e outros combustíveis (32,33). A faixa de operação usual da espectroscopia NIR é de 0,9 a 2,5 μm , enquanto que espectros na região do infravermelho médio são geralmente obtidos entre 2,5 e 25 μm (4000 – 400 cm^{-1}). Um espectro NIR é o resultado dos segundo e terceiro sobretons e bandas de combinação das frequências fundamentais associadas a certos grupos funcionais que produzem a medida direta na região do infravermelho médio. Como a espectroscopia MIR é baseada em medidas das ressonâncias fundamentais características, esta produz bandas específicas, usualmente estreitas e bem definidas. Dessa forma, devido a maior riqueza de informações na região do infravermelho médio, vários trabalhos já publicados indicaram que correlações que usam dados derivados de FTIR para conteúdo de hidrocarbonetos aromáticos, razão carbono-hidrogênio, índice de refração, calor de combustão, viscosidade, e densidade de combustíveis, eram superiores àquelas usando dados derivados de infravermelho próximo (18,34,35,36).

Uma exigência natural para o desenvolvimento de correlações entre propriedades de gasolinas e seus espectros na região do MIR é que estas propriedades sejam dependentes de características químicas estruturais da amostra, associadas com bandas ativas nesta região. Desse modo, os valores medidos para essas propriedades podem ser relacionados às intensidades espectrais.

Para ilustrar algumas das relações entre intensidades espectrais e propriedades de combustíveis, os seguintes argumentos podem ser considerados: altos valores de densidade específica estão associados com altas concentrações de hidrocarbonetos aromáticos ou naftênicos; baixos valores de densidade específica estão associados com altas concentrações de parafinas de cadeia aberta; um aumento na concentração de hidrocarbonetos aromáticos implica num aumento no número de octanas e num decréscimo do número de cetanas; gasolinas “oxigenadas” podem conter álcoois ou éteres e as ligações C-O-C e C-O-H possuem bandas características na região do infravermelho. Assim, várias propriedades de combustível são determinadas pela presença de grupos atômicos específicos, pelo grau de ramificação dos hidrocarbonetos

saturados, pela razão *normal-iso* parafinas, pelo conteúdo de aromáticos e pela razão entre estes vários constituintes. Todas essas características químicas são refletidas, em maior ou menor grau, nos espectros infravermelho dos diferentes compostos (18).

Uma variedade considerável de trabalhos já foi publicada sobre o desenvolvimento de novas metodologias de análise de combustíveis, usando-se espectroscopia vibracional, seja MIR, NIR ou Raman, acoplada a algum tipo de tratamento estatístico multivariado.

ICHIKAWA e colaboradores (37) estimaram os valores de MON de um conjunto de 74 amostras de gasolinas, utilizando um modelo de regressão linear. Esse modelo foi construído a partir dos valores de absorção em apenas cinco comprimentos de onda na região de estiramento C-H ($3150\text{--}2800\text{ cm}^{-1}$) do espectro infravermelho de 50 amostras de calibração. Os resultados obtidos apresentaram erro de até ± 1 unidade de MON.

Em um trabalho mais recente, ANDRADE e colaboradores (13) estimaram os valores de MON e RON de um conjunto de 310 amostras de naftas provenientes de reformas catalíticas, utilizando modelos de regressão PLS, construídos usando-se a região do espectro infravermelho de $1600\text{ a }600\text{ cm}^{-1}$. Os primeiros resultados apontaram para a existência de dois subconjuntos de amostras, divididos de acordo com o teor de aromáticos, e revelados através da diferença nas intensidades das bandas em $728\text{ e }768\text{ cm}^{-1}$, típicas da vibração C-H fora do plano. Melhores resultados foram obtidos com a construção de um modelo de predição para cada um dos subconjuntos observados. Os valores de precisão foram considerados excelentes para os dois modelos, mas a exatidão teve comportamento distinto para cada um dos dois grupos. Para amostras com baixo teor de aromáticos nenhum tipo de tendência foi observado, mas para as amostras com alto teor de aromáticos, os resultados apresentaram alguma tendência.

Mesmo o álcool não sendo tão utilizado no exterior quanto no Brasil, o seu teor e o de outros oxigenados presentes nas gasolinas aparecem como importantes propriedades estudadas. GARCIA e colaboradores (25) determinaram os teores de metanol e MTBE (methyl *tert*-butyl éter) de 10 amostras, utilizando um modelo de regressão PLS. Esse modelo foi construído usando-se a região de estiramento C-O para os dois compostos ($981\text{--}1115\text{ cm}^{-1}$), a partir de 18 amostras de calibração. Os resultados para o conjunto teste, que incluía amostras contendo de 0,40 a 4,45% de metanol, apresentaram desvio padrão de 0,10 para esse componente.

IOB e colaboradores (24) determinaram os teores dos álcoois C_1 a C_4 e de MTBE de um conjunto de 38 amostras de gasolinas, utilizando um modelo de regressão PLS. Esse modelo foi construído usando-se a região de estiramento C-O ($1300-810\text{ cm}^{-1}$) do espectro infravermelho de 32 amostras de calibração. Os resultados obtidos apresentaram desvio padrão relativo igual a 6,3% para predição do teor de etanol.

Outros trabalhos publicados concentraram-se em comparar diferentes técnicas de regressão, como regressão linear múltipla (MLR), regressão por componentes principais (PCR) e PLS, entre outras. ANDRADE e colaboradores realizaram esse tipo de estudo para a predição de propriedades de gasolina (17) e GARRIDES e colaboradores para a predição de propriedades de querosene (31), ambos pertencentes ao mesmo grupo de trabalho. Nos dois trabalhos, a regressão PLS foi a que apresentou melhores resultados.

FODOR e colaboradores (18,38) publicaram trabalhos bastante completos, envolvendo a comparação entre dois diferentes equipamentos FTIR, Nicolet e Perkin Elmer, e entre três tipos diferentes de células, célula de transmissão (KBr), célula de refletância total atenuada (ATR) plana e célula ATR do tipo túnel, a mesma que foi utilizada neste trabalho. As amostras foram coletadas em diversas regiões do país (EUA) durante as quatro estações do ano, somando um total de 800 amostras, que foram divididas nos conjuntos de calibração, validação e teste. Além disso, diversas propriedades foram estudadas, tais como: MON, RON, composição química, temperaturas de destilação, viscosidade e calor de combustão.

Os trabalhos acima mencionados apresentaram resultados com excelente qualidade em termos de exatidão e precisão para quase todas as propriedades estudadas. Além disso, os autores fazem uma ressalva quanto à utilização da célula ATR do tipo túnel, por causa da possibilidade de haver contaminação cruzada entre uma análise e outra e porque o tempo de análise é maior, visto que a etapa de limpeza da célula é bastante demorada. A utilização da célula ATR plana foi sugerida como a que fornece melhores resultados.

Trabalhos mais recentes (publicados a partir de 1999) sobre métodos de regressão multivariada, principalmente aplicados à espectroscopia NIR, apresentam novas abordagens, tais como: a utilização de novos métodos de regressão como alternativa ao PLS (17,39,40); a utilização de técnicas estatísticas de seleção de variáveis, como redes neurais e algoritmos genéticos (42); e estudos sobre a manutenção das calibrações FTIR-PLS ao longo do tempo (43,44).

Nesse sentido, HAALAND e MELGAARD (39) publicaram um artigo propondo uma extensão do método dos mínimos quadrados clássicos (PACLS) como alternativa para análises quantitativas a partir de dados espectrais, afirmando obter resultados de predição similares ou melhores que o PLS para diversos sistemas testados.

VIGNEAU e colaboradores (40) também propuseram um novo método de regressão aplicado a dados espectrais, chamado por eles de modelos de regressão através da análise dos componentes principais restrita (RM-CPCA), como uma alternativa para dados com problemas de multi-colinearidade. Os autores classificaram o novo método como promissor, mas afirmam que estudos complementares devem ser realizados.

Métodos de regressão lineares e não lineares foram utilizados para prever os valores de MON de gasolinas provenientes de reforma catalítica, a partir dos espectros de transmissão na região do infravermelho médio, em trabalho publicado por ANDRADE e colaboradores (17). A existência e a influência de não-linearidades foram investigadas comparando-se a habilidade preditiva de modelos lineares (PCR e PLS) e não-lineares, sejam paramétricos (PLS-polinomial, LWR) ou não-paramétricas (*genetic inside neural network*). Os resultados apresentados apontam como vantagem dos métodos não-lineares o fato destes modelarem a informação de maneira mais efetiva, ou seja, mais informação é explicada com uma quantidade menor de fatores ou variáveis latentes utilizada, embora a qualidade das predições seja bastante similar a dos métodos lineares.

Um estudo comparando-se o desempenho relativo de diferentes métodos de calibração multivariada em função do tamanho do conjunto dos dados de calibração, realizado por HUANG e colaboradores (41), mostrou que o desempenho da predição, em termos de exatidão e robustez, depende do tamanho do conjunto de dados utilizado.

Autores brasileiros, COSTA FILHO e POPPI (42), publicaram um trabalho sobre a aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no infravermelho médio para a determinação simultânea de glicose, maltose e frutose em amostras sintéticas. A conclusão dos autores foi que os modelos obtidos a partir das variáveis selecionadas pelo algoritmo genético eram mais robustos, ou seja, apresentavam pequenas variações no desvio padrão dos erros, durante a previsão de novas amostras. Além disso, eles concluíram que a seleção de variáveis por si só pode dispensar a utilização de pré-tratamentos sobre os dados.

MACHO e LARRECHI (43) e SETAREHDAN e colaboradores (44) estudaram algumas estratégias para a manutenção de modelos quantitativos multivariados, obtidos a partir de espectros NIR. Ambos os trabalhos concluíram ser possível manter a habilidade preditiva desses modelos ao longo do tempo.