

## Aplicação de MARS- “Multivariate Adaptive Regression Splines” em metrologia para seleção de variáveis na determinação de biodiesel

Werickson Fortunato de Carvalho Rocha<sup>1\*</sup> (PQ) e Cleber Nogueira Borges<sup>2</sup> (PQ)  
\*wfrocha@inmetro.gov.br

<sup>1</sup>INMETRO-Instituto Nacional de Metrologia, Normalização e Qualidade Industria-DQUIM- Divisão de Metrologia Química, CEP 25250-020, Av. N. S. das Graças 50 - Xerém - Duque de Caxias - RJ, Brasil.

<sup>2</sup>UEMG - Universidade do Estado de Minas Gerais - Campus de Frutal - Av. Professor Mário Palmério nº 1000 - Bairro Universitário - Frutal, MG, CEP 38200-000.

Palavras Chave: MARS, Metrologia Química Seleção de variáveis.

### Introdução

MARS – *Multivariate Adaptive Regression Splines* – é uma técnica de regressão multivariada não paramétrica com seleção de variáveis utilizada em Quimiometria. Quando aplicada em conjuntos de dados espectrais visa selecionar as melhores regiões (sinais) de uma matriz de dados de forma que esta esteja mais correlacionada com a propriedade de interesse<sup>1</sup>. MARS é utilizado tanto na classificação como na regressão e possui um número crescente de aplicações em muitas áreas como economia, ciência e tecnologia. É muito útil para problemas de elevada dimensão e mostra uma grande promessa para funções de ajuste não linear.

MARS pode ser entendida como uma extensão da regressão linear múltipla. Constrói a relação entre as variáveis independentes e respostas a partir de um conjunto de coeficientes e funções de base, denominadas *hinge functions*, dirigida unicamente por esses dados.

O objetivo deste trabalho é divulgar a utilização do método quimiométrico MARS para selecionar variáveis para construção de modelos de calibração multivariada. Para ilustrar a utilização do MARS, foi utilizado de dados obtidos pela técnica de espectroscopia no infravermelho próximo de blendas contendo biodiesel de amendoim e diesel.

### Resultados e Discussão

Foram utilizadas 49 amostras de misturas biodiesel/diesel cujas concentrações estão na faixa de 2% (V/V) a 90% (V/V) em recipientes de 5 mL. Todas as amostras foram agitadas durante 1 minuto para garantir a homogeneização.

As amostras foram divididas em dois conjuntos: 30 amostras para calibração e 20 para a validação. As análises foram realizadas no espectrômetro GX Spectrum da Perkin-Elmer com acessório de refletância difusa, com uma resolução de 4 cm<sup>-1</sup> e 10 varreduras por espectro. O programa computacional ‘R-package’ foi utilizado para a

construção do modelo juntamente com o pacote ‘earth’ para calibração multivariada referida.

Os resultados obtidos encontram-se na figura 1. Nesta figura é possível notar que o modelo MARS desenvolvido foi capaz de ajustar muito bem os dados, pois apresentou erro quadrático médio de previsão próximo de 1%. No que refere-se ao ajuste do modelo pode-se notar a presença de resíduos que se distribuem de forma aleatória tanto para as amostras de calibração como para as amostras de validação.

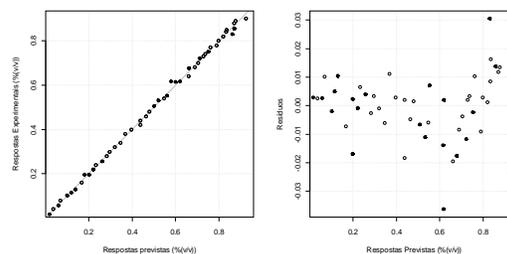


Figura 1: a) Valores mensurados versus valores previstos pelo modelo MARS. ● Amostras de calibração e ○ amostras de validação; b) Resíduos do modelo.

### Conclusões

O presente trabalho demonstra a importância na seleção de variáveis para a obtenção do modelo de calibração multivariada. O MARS permitiu, neste caso, a construção do modelo de calibração multivariada com apenas 2 comprimentos de onda.

Os resultados encontrados sugerem que o modelo MARS é eficaz para a determinação de biodiesel em diesel, pois apresenta valor de RMSEP apenas de 1,3% (v/v). Logo, o trabalho demonstra uma nova ferramenta quimiométrica para seleção de variáveis.

<sup>1</sup> Friedman, J.H.; “Multivariate Adaptive Regression Splines (with discussion)”, *Annals of Statistics*, 1991, Vol 19, N° 1, pg 1-141. url: <http://www.salfordsystems.com/doc/MARS.pdf>