

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ALIMENTOS
CURSO DE ENGENHARIA DE ALIMENTOS

LUCAS VEIGA SABINO

**QUANTIFICAÇÃO DO PERCENTUAL DE CAFÉ ARÁBICA EM
MISTURAS COM CAFÉ ROBUSTA USANDO FTIR, NIRS E MODELOS
DE DOIS ESTÁGIOS**

TRABALHO DE CONCLUSÃO DE CURSO

CAMPO MOURÃO

2015

LUCAS VEIGA SABINO

**QUANTIFICAÇÃO DO PERCENTUAL DE CAFÉ ARÁBICA EM
MISTURAS COM CAFÉ ROBUSTA USANDO FTIR, NIRS E MODELOS
DE DOIS ESTÁGIOS**

Trabalho de conclusão de curso de graduação, apresentado ao Curso Superior de Engenharia de Alimentos do Departamento Acadêmico de Alimentos, da Universidade Tecnológica Federal do Paraná – UTFPR, Câmpus Campo Mourão, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Alimentos.

CAMPO MOURÃO



TERMO DE APROVAÇÃO

QUANTIFICAÇÃO DO PERCENTUAL DE CAFÉ ARÁBICA EM MISTURAS COM CAFÉ ROBUSTA USANDO FTIR, NIRS E MODELOS DE DOIS ESTÁGIOS

Por

LUCAS VEIGA SABINO

Esse trabalho de conclusão de curso foi apresentado às 14:00 do dia 02 de dezembro de 2015, como requisito parcial para a obtenção do Título de Bacharel em Engenharia de Alimentos. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Dr. Paulo Henrique Março
Orientador

Prof. Dra. Angela Maria Gozzo
Membro da banca

Prof. Dr. Augusto Tanamati
Membro da banca

O documento original e assinado pela banca examinadora encontra-se na Coordenação do Curso de Engenharia de Alimentos da UTFPR-CM.

AGRADECIMENTOS

Agradeço ao meu professor orientador Dr. Evandro Bona por toda a inspiração, didática, conhecimento, compromisso e dedicação para o sucesso do presente estudo.

Ao professor Dr. Paulo Henrique Março por toda a colaboração, a qual foi crucial para o desenvolvimento deste trabalho.

À Luiza Mariano Leme por toda a ajuda com as análises de laboratório. À minha grande amiga Thais Agante pelo auxílio com a revisão dos textos. Em especial, à professora e amiga Maria Lucia Biazoti, por contribuir para minha formação.

À minha namorada Ana Raíza Yoshioka por todo o carinho e paciência.

Agradeço principalmente aos meus pais, Francisco Sabino e Silvana Veiga Sabino, pois sem eles nada disto seria possível.

RESUMO

SABINO, Lucas V. **Quantificação do percentual de café arábica em misturas com café robusta usando FTIR, NIRS e modelos de dois estágios.** 2015. 43 f. Trabalho de conclusão de curso – Departamento Acadêmico de Engenharia de Alimentos, Universidade Tecnológica Federal do Paraná. Campo Mourão, 2015.

O café é uma das bebidas mais aceitas e apreciadas por diversos países no mundo. As principais espécies são a *Coffea arabica* (arábica) e a *Coffea canephora* (robusta), as quais apresentam uma composição química muito distinta. Sendo que a espécie *C. arabica*, destaca-se por proporcionar bebida de qualidade e aroma superior ao *C. canephora*, cuja bebida, considerada neutra, destina-se aos *blends* e à indústria de café solúvel, favorecida pelo menor preço. As misturas de café, ou *blends*, são muito utilizadas quando se deseja manter uma uniformidade de sabor no produto, assim, foi testado o desenvolvimento de metodologias analíticas confiáveis para indicar a quantidade de cada tipo de café em uma mistura. Os espectros das misturas de arábica e robusta foram obtidos através de espectrofotometria de infravermelho próxima (NIRS) e média (FTIR). Como método de primeiro estágio, foram utilizados a Análise de Componentes Principais (PCA) e os Mínimos Quadrados Parciais. A partir dos resultados, pôde-se concluir que os espectros de FTIR conduziram a melhores resultados que os espectros NIR. O método de redução de dimensionalidade de dados mais satisfatório foi o PLS. Já a rede que apresentou melhor desempenho na criação do modelo de previsão foi a rede do tipo MLP, a qual, de maneira geral, obteve menores resultados de erro absoluto médio e maiores valores de coeficiente de correlação.

Palavras-chave: FTIR, NIRS, ACP, PLS, redes neurais artificiais.

ABSTRACT

SABINO, Lucas V. **Percentage of arabica coffee determination in blends with robusta coffee using FTIR in tandem with NIRS and two-stage models.** 2015. 44 f. Trabalho de conclusão de curso – Departamento Acadêmico de Engenharia de Alimentos, Universidade Tecnológica Federal do Paraná. Campo Mourão, 2015.

Coffee is one of the most acceptable and appreciated drinks around the world. The main species are *Coffea arabica* and *Coffea canephora*, which shows a very different chemical composition. The specie *C. Arabica* stands out for providing a drink with higher quality and aroma, if compared with *C. Canephora*. The last specie is considered a neutral drink, for blends and for the industry of instant coffee, because of his low cost. Blends are very used when it's necessary to maintain a uniform flavor. Thus, the main objective of this work was to develop a reliable analytic method to indicate the amount of each type of coffee in mixture. Spectra of samples were obtained by Infrared Spectroscopy with Fourier Transform (FTIR) and Near Infrared Spectroscopy (NIRS). In the first stage, it was used Principal Component Analysis (PCA) and the Partial Least Square (PLS). The result shows that the spectra obtained by FTIR gave better results than the NIR spectra. The method for data dimensionality reduction more satisfactory was PLS. The ANN with the best performance was MLP, which, in general, had lower results of Mean Absolute Error (MAE) and higher correlation coefficient.

Keywords: *FTR, NIRS, ACP, PLS, artificial neural networks.*

LISTA DE FIGURAS

Figura 1. Fluxograma da metodologia do trabalho.	16
Figura 2. Representação de um perceptron de múltiplas camadas (MLP).....	21
Figura 3. Representação de um perceptron de múltiplas camadas (MLP).....	23
Figura 4. Interpretação geométrica do simplex para duas variáveis (a) e para três variáveis (b).....	26
Figura 5. Movimento de reflexão em um simplex para duas variáveis. Fonte: Bona (2000).....	27
Figura 6. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede MLP com PCA como primeiro estágio (FTIR).....	31
Figura 7. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede RBF com PCA como primeiro estágio (FTIR).....	32
Figura 8. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede MLP com PLS como primeiro estágio (FTIR).....	34

LISTA DE TABELAS

Tabela 1. Parâmetros a serem otimizados para a rede do tipo MLP.....	28
Tabela 2. Parâmetros a serem otimizados para a rede do tipo RBF.....	28
Tabela 3. Otimização simplex e desempenho das redes MLP (PCA).....	30
Tabela 4. Otimização simplex e desempenho das redes RBF (PCA).....	31
Tabela 5. Otimização simplex e desempenho das redes MLP (PLS).....	33
Tabela 6. Otimização simplex e desempenho das redes RBF (PLS).....	34
Tabela 7. Otimização simplex e desempenho das redes MLP (PCA).....	36
Tabela 8. Otimização simplex e desempenho das redes RBF (PCA).....	37
Tabela 9. Otimização simplex e desempenho das redes MLP (PLS).....	39
Tabela 10. Otimização simplex e desempenho das redes RBF (PLS).....	40
Tabela 11. Comparação entre a melhor rede alimentada com espectros FTIR e a melhor alimentada com espectros NIR.	42

LISTA DE ABREVIATÖES

CP – Componentes Principais

FTIR – Espectroscopia de Infravermelho Médio com transformada de Fourier (do inglês – *Fourier Transform Mid-Infrared Spectroscopy*)

MLP – Perceptron de Múltiplas Camadas (do inglês – *Multilayer Perceptron*)

MSC – Correção do Espalhamento Multiplicativo (do inglês – *Multiplicative Scatter Correction*)

NIRS – Espectroscopia de Infravermelho Próximo (do inglês – *Near Infrared Spectroscopy*)

PCA – Análise de Componentes Principais (do inglês *Principal Component Analysis*)

PLS – Mínimos Quadrados Parciais (do inglês – *Partial Least Squares*)

RNA – Redes Neurais Artificiais

RBF – Rede de Função de Bases Radiais (do inglês – *Radial Basis Function*)

SUMÁRIO

1. INTRODUÇÃO	11
2. OBJETIVOS	15
2.1. OBJETIVO GERAL	15
2.2. OBJETIVOS ESPECÍFICOS	15
3. METODOLOGIA	16
3.1 AMOSTRAS	16
3.2. ESPECTROSCOPIA DE INFRAVERMELHO COM TRANSFORMADA DE FOURIER (FTIR)	17
3.2.1 PRÉ-PROCESSAMENTOS	17
3.3 ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS)	17
3.3.1 PRÉ-PROCESSAMENTOS	17
3.4 PRIMEIRO ESTÁGIO	18
3.4.1 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)	18
3.4.2 MÍNIMOS QUADRADOS PARCIAIS	19
3.5 SEGUNDO ESTÁGIO	20
3.5.1 PERCEPTRON DE MULTIPLAS CAMADAS (MLP)	20
3.5.2 REDES DE FUNÇÃO DE BASES RADIAIS (RBF)	23
3.6 OTIMIZAÇÃO SIMPLEX	25
3.7 IMPLEMENTAÇÃO COMPUTACIONAL	29
4. RESULTADOS E DISCUSSÕES	30
4.1 FTIR	30
4.1.2 PCA COMO PRIMEIRO ESTÁGIO	30
4.1.3 PLS COMO PRIMEIRO ESTÁGIO	32
4.2 NIR	35
4.2.1 PCA COMO PRIMEIRO ESTÁGIO	36
4.2.2 PLS COMO PRIMEIRO ESTÁGIO	38
4.3. COMPARAÇÃO ENTRE NIR E FTIR	41
5. CONCLUSÃO	43
6. REFERÊNCIAS	44

1. INTRODUÇÃO

A planta de café é originária da Etiópia, sendo a Arábia responsável pela sua propagação. O nome do café vem da palavra árabe *qahwa* (vinho), assim, o café ficou conhecido na Europa como vinho da Arábia. Os manuscritos mais antigos mencionam o café como fruto consumido em natura, no Yêmen. Com o tempo, o grão começou a ser macerado e misturado com gordura animal para facilitar seu consumo durante as viagens.

O café tornou-se de grande importância para os árabes, os quais tinham domínio na cultura do fruto e preparação da bebida. Em 1000 d.C., os árabes começaram a preparar infusões com cerejas, fervendo-as em água. Somente no século XIV, na Pérsia, o grão foi torrado e se transformou em uma bebida com aspecto mais semelhante ao atual. O hábito de tomar café foi difundido pelo mundo árabe. O Produto passou a ter tal importância que, em 1475, foi promulgada uma lei permitindo à mulher pedir divórcio, se o marido fosse incapaz de lhe prover determinada quantidade diária da bebida. Apenas em 1615 o grão, trazido por viajantes, chegou à Europa. (ABIC, 2015).

O café, no Brasil, destaca-se como produto de grande importância econômica desde a chegada das primeiras mudas, em meados do século XVIII. Por apresentar rápida adaptação ao solo e clima, se tornou um dos principais produtos de exportação. Atualmente, são 15 estados produtores, com destaque para Minas Gerais, Espírito Santo, São Paulo, Bahia, Paraná e Rondônia (MAPA, 2015). Devido à diversidade das regiões as quais o café é cultivado, diversos tipos de produtos são produzidos, de maneira que diferentes demandas – tanto paladar quanto preço – possam ser atendidas. Esta diversidade permite o desenvolvimento dos mais variados *blends* (MAPA, 2014).

Pertencente à família *Rubiaceae*, o café possui cerca de 500 gêneros e mais de 6000 espécies. Porém, as duas espécies economicamente mais importantes são *Coffea arabica* (café arábica), o qual é responsável por 60% do mercado mundial de

café; e *Coffea canephora* (café robusta), responsável pelos 40% restantes (FARAH, 2009; ICO, 2015).

A composição básica do café depende principalmente de aspectos genéticos (espécie e variedade) e aspectos fisiológicos (grau de maturação). Em menor proporção, também depende de fatores extrínsecos, como composição do solo e clima, práticas agrícolas, condições de armazenamento. Os cafés arábica e robusta possuem composições químicas bastante distintas, sendo o café arábica responsável por uma bebida com qualidade superior. Já o café robusta, apesar de proporcionar uma bebida de qualidade inferior, quando comparada ao arábica, apresenta menos custo de plantio, devido as suas árvores resistentes. Por décadas, pesquisadores tentaram combinar a resistência dos pés de café robusta com a qualidade da bebida proporcionada pelo café arábica, porém, descobriu-se que os fatores químicos responsáveis pela alta resistência, também são responsáveis pela bebida de inferior qualidade. Um exemplo é a concentração de ácidos clorogênicos - presente em alta quantidade nos pés de robusta-, responsáveis pela proteção contra microrganismos, insetos e radiação UV. Porém, baixas concentrações de ácidos clorogênicos são importantes para a formação do aroma do café, pois altos níveis provocam a elevação da oxidação antes da torrefação (FARAH, 2009).

As misturas de café são utilizadas quando se deseja manter uma uniformidade do sabor e reduzir a acidez do produto. Nos blends podem ser adicionadas diferentes espécies ou até mesmo safras de café (FERNANDES et al., 2003). Assim, são necessárias metodologias confiáveis para a determinação de cada tipo de café nos *blends*.

A espectroscopia de infravermelho médio vem sendo cada vez mais empregada para análise quantitativa e qualitativa de alimentos. O FTIR é rápido e tem ampla aplicação analítica em análises químicas e de controle de qualidade (HELPER et al., 2006; MORGANO et al., 2005). A porção de radiação infravermelha de maior utilidade em análises orgânicas está situada entre 4000 cm^{-1} e 400 cm^{-1} . Certos grupos de átomos dão origem a bandas que ocorrem mais ou menos na mesma frequência, independentemente da estrutura da molécula e, a presença destas bandas permite, através do exame do espectro, a obtenção de informações estruturais úteis (SILVERSTEIN; WEBSTER; KIEMLE, 2007).

Técnicas como a espectroscopia no infravermelho próximo (NIRS) associada aos avanços computacionais e à utilização de métodos quimiométricos apresentam bons resultados na avaliação de qualidade de alimentos (MORGANO et al., 2005; SANTOS et al., 2006). A espectroscopia de infravermelho próximo atua no intervalo de número de onda de 12800 cm^{-1} até 4000 cm^{-1} . Apresenta uma série de vantagens, como a rapidez na obtenção dos espectros, cálculos e apresentação dos resultados; o fato de ser um método não destrutivo, permitindo o uso posterior da amostra; o fato de ser uma técnica não evasiva, porém com alta penetração de radiação (BUENO, 2004).

Apesar das vantagens, tanto o FTIR quanto o NIRS apresentam espectros complexos, os quais exigem a utilização de métodos de análise multivariada, de maneira que as amostras possam ser diferenciadas de maneira criteriosa.

A Análise de Componentes Principais (PCA – *Principal Component Analysis*) é uma ferramenta padrão na análise de dados moderna. É utilizada em diversos campos, desde a neurociência até a computação gráfica, pois se trata de um método capaz de extrair informações relevantes de um complexo conjunto de dados (SHLENS, 2014). Outra maneira de reduzir a dimensionalidade dos dados é a regressão por mínimos quadrados parciais.

Os mínimos quadrados parciais (PLS - *Partial Least Squares*) vem ganhando importância devido principalmente à simplicidade do algoritmo e excelente poder de previsão (PERALTA et al., 1997). Durante a etapa de calibração, a modelagem PLS utiliza tanto a informação da matriz de dados X como da matriz de concentração Y , obtendo-se assim variáveis latentes. Construído o modelo de calibração, este é utilizado para se fazer a estimativa (validação) em novas amostras.

O trabalho com redes neurais artificiais (RNAs) têm sido motivado desde o começo pelo reconhecimento de que o cérebro humano processa informações de uma forma inteiramente diferente de um computador convencional. O cérebro é um computador altamente complexo, não-linear e paralelo, possuindo a capacidade de organização de seus componentes estruturais – os neurônios – de forma a realizar certos processamentos (reconhecimento de padrões, percepção, controle motor, etc.) muito mais rapidamente que o mais rápido computador digital atual (HAYKIN, 2001).

As redes neurais artificiais são um conjunto de métodos matemáticos e algoritmos computacionais especialmente projetados para simular o processamento de informações e aquisição de conhecimento do cérebro humano (BUENO, 2004). Nas RNA o processamento das informações ocorre nos neurônios e os estímulos são transmitidos por meios de conexões, às quais são ajustados pesos, determinando-se assim o estímulo de saída (MORAES, 2012). São extremamente versáteis para realizar o mapeamento de relações complexas e não lineares entre múltiplas variáveis de entrada e saída (BHISOP, 2006).

Com o intuito de criar um modelo que seja capaz de prever a porcentagem de café arábica em misturas com café robusta, foram testadas as redes neurais Perceptron de Múltiplas Camadas (MLP – *Multilayer Perceptron*) e Rede de Função de Bases Radiais (RBF – *Radial Basis Function*), sendo utilizado como alimentação para as redes, os scores da PCA e do PLS (VARELLA, 2008)

2. OBJETIVOS

2.1. OBJETIVO GERAL

Tem-se como objetivo deste trabalho a determinação da porcentagem de café arábica, nos *blends* de arábica e robusta, utilizando espectroscopia e modelos de dois estágios.

2.2. OBJETIVOS ESPECÍFICOS

- Preparar os *blends* de arábica e robusta, com diferentes porcentagens de arábica, numa faixa de 0 % a 100%;
- Obter os espectros FTIR e NIRS dos *blends*;
- Realizar o pré-tratamento dos espectros: correção do espalhamento multiplicativo (MSC) e 2ª derivada do espectro usando o algoritmo de *Savitzky-Golay*;
- Construção do primeiro estágio usando técnicas lineares: análise de componentes principais (ACP) e mínimos quadrados parciais (PLS);
- Construção do segundo estágio usando técnicas não-lineares: perceptrons de múltiplas camadas (MLP) e redes de função de bases radiais (RBF);
- Definição da combinação entre a melhor forma de apresentação do espectro (puro ou segunda derivada), melhor método de redução de dimensionalidade de dados e melhor combinação de parâmetros das redes.

3. METODOLOGIA

O fluxograma abaixo mostra a sequência realizada nesse trabalho.

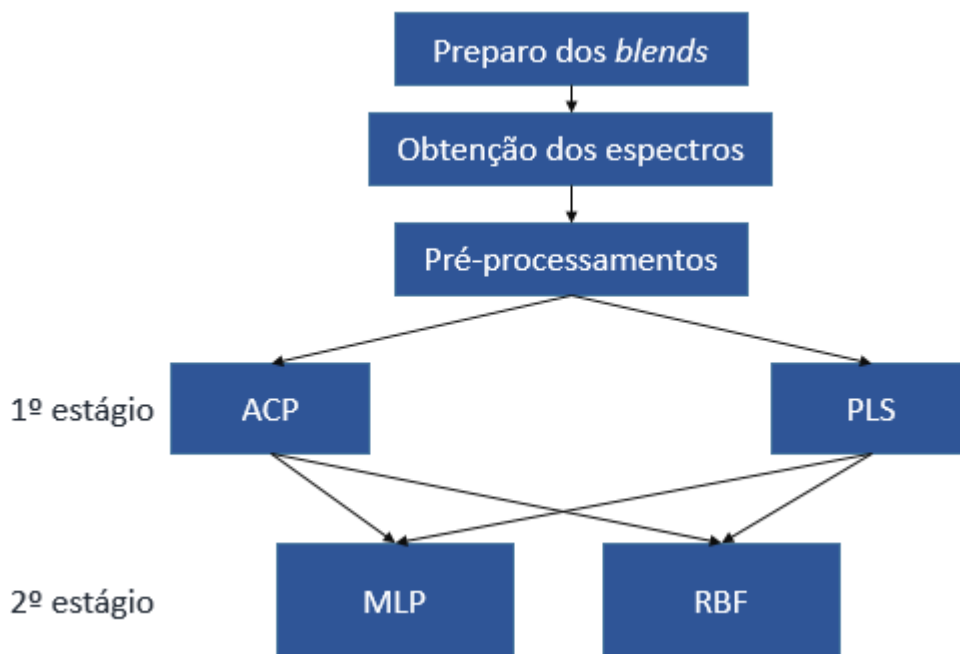


Figura 1. Fluxograma da metodologia do trabalho.

3.1 AMOSTRAS

As amostras de café arábica e robusta foram doadas por uma empresa localizada na região Centro-Oeste do Paraná. As mesmas foram torradas em uma temperatura de 197°C por 15 minutos e posteriormente moídas. Foram realizadas então misturas entre os dois tipos de café, em uma quantidade total de 10g, de maneira que se variou a porcentagem de arábica no *blend*, desde 0% até 100%, com um passo de 2% entre uma amostra e outra. Os *blends* foram feitos em duplicata para cada porcentagem, totalizando 306 amostras.

3.2. ESPECTROSCOPIA DE INFRAVERMELHO MÉDIO COM TRANSFORMADA DE FOURIER (FTIR)

Para preparar as pastilhas foram adicionados 100 mg de KBr seco (padrão cromatográfico, Sigma- Aldrich) e aproximadamente 1 mg de amostra em molde padrão (ICL, ICL's Macro/Micro KBr die) sob 7 toneladas de pressão. Primeiramente foi realizada uma leitura sem amostra (*background*) com o intuito de descontar a presença do ar nos demais espectros. O FTIR (Shimadzu, IR Affinity-1) varreu a região de 4000 a 400 cm^{-1} em triplicata para cada amostra, sendo feitas 32 varreduras acumuladas para a formação do espectro final.

3.2.1 PRÉ-PROCESSAMENTOS

Após obtenção dos espectros foram aplicadas algumas transformações nos mesmos. Primeiramente, foi realizada a normalização do espectro (a maior banda obteve absorbância 1 e a menor 0), correção da linha de base, a suavização do espectro. Foi considerado apenas a região de 1900 e 800 cm^{-1} por tratar-se da faixa mais importante para a análise de café (BRIANDET; KEMSLEY; WILSON, 1996; LYMAN et al., 2003; WANG et al., 2009).

3.3 ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS)

Para a espectroscopia de infravermelho, os espectros foram obtidos em um aparelho JDSU MicroNIR 2200. As leituras foram realizadas à temperatura ambiente, na região de 900-1600 nm com intervalo de 6 nm. Assim como no FTIR, as leituras foram feitas em triplicata para cada amostra.

3.3.1 PRÉ-PROCESSAMENTOS

Após obtidos, os espectros passaram por tratamentos preliminares, os quais tem como objetivo reduzir fontes de variação que não carregam informações

relevantes. Foram feitas a correção de linha de base, correção do espalhamento multiplicativo (MSC – *Multiplicative Scatter Correction*) e a segunda derivada. O MSC corrige simultaneamente os efeitos aditivos do espalhamento de luz gerados por diferença na granulometria e morfologia das partículas. Para realizar a correção é utilizada uma regressão linear das variáveis espectrais versus o espectro médio. A segunda derivada, por meio do algoritmo de *Savitzky-Golay* remove os problemas devido a mudanças de inclinação entre as amostras (MARQUETTI, 2014).

3.4 PRIMEIRO ESTÁGIO

3.4.1 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

A análise de componentes principais (PCA – *Principal Component Analysis*) é um método multivariado estatístico comumente utilizado para reduzir a variância e resolver problemas de colinearidade. O método faz combinações lineares as quais buscam sintetizar os dados sem que haja perda de informação no processo (YAN, 2008). Além disso, a análise separa variações importantes dos dados do ruído experimental, auxilia no reconhecimento de padrões de comportamento, detecta amostras anômalas que não se encaixam no modelo, facilitando a visualização e a interpretação dos resultados (FARO et al., 2010).

O conjunto de variáveis originais é transformado em outro, de mesma dimensão sem que haja perda de informação. Os componentes principais apresentam propriedades importantes: cada componente principal é uma combinação linear de todas as variáveis originais, e possuem, em ordem de estimação, o máximo de informação, em termos de variação total contida nos dados (VARELLA, 2008). Assim, os dados podem ser visualizados e analisados por meio dos *scores* das novas variáveis associados aos vetores de peso (*loading*), proporcionando informações referentes à correlação que existe entre as variáveis de entrada (MELSSSEN; WEHRENS; BUYDENS, 2006). A PCA decompõe a matriz de variáveis originais \mathbf{X} ($m \times n$), com m amostras (linhas) e com n variáveis (colunas) referentes à porcentagem de arábica, em três matrizes:

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} \quad (1)$$

onde, \mathbf{T} ($m \times k$) e \mathbf{P}^T ($m \times k$) são as matrizes de escores e pesos, respectivamente, com k componentes principais (CP). A matriz \mathbf{E} ($m \times n$) contém os resíduos, que é a parte da variância dos dados que não é explicada pelo modelo de PCA (JAVIDNIA et al., 2013).

3.4.2 MÍNIMOS QUADRADOS PARCIAIS

Foi aplicado o método de regressão multivariada para dados de primeira ordem conhecido por PLS (do inglês *Partial Least Squares*). O objetivo do modelo PLS foi encontrar uma relação entre a matriz \mathbf{X} ($m \times n$), que contém os espectros infravermelhos corrigidos dos *blends* de arábica e robusta, e o vetor \mathbf{Y} que armazena as respectivas porcentagens de arábica presente em cada *blend*. O modelo PLS utiliza a resposta desejada para cada amostra de treinamento na decomposição dos dados em *scores* e *loadings*, sendo estabelecida uma relação linear entre a variável dependente \mathbf{Y} e a variável independente \mathbf{X} . A definição do número de variáveis latentes do modelo foi realizada pela observação do erro quadrático médio (MSE, do inglês *Mean Square Error*). No PLS, a relação entre a matriz \mathbf{X} e \mathbf{Y} é representado como uma relação algébrica linear entre seus *scores*. Por sua vez, estes são obtidos pela decomposição das matrizes de dados, expressados como:

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} = \sum_{i=1}^A \mathbf{t}_i \cdot \mathbf{p}_i^T + \mathbf{E} \quad (2)$$

$$\mathbf{Y} = \mathbf{T} \cdot \mathbf{q}^T + \mathbf{f} = \sum_{i=1}^A \mathbf{t}_i \cdot \mathbf{q}_i^T + \mathbf{f} \quad (3)$$

onde, a matriz \mathbf{E} e \mathbf{f} contém os resíduos de \mathbf{X} e \mathbf{Y} , respectivamente - partes que não foram explicadas pelo modelo. O vetor \mathbf{t}_i , que forma as colunas da matriz \mathbf{T} (matriz de escores) é chamado de vetor de *scores*, \mathbf{p}_i e \mathbf{q}_i são chamados de *loadings* (pesos), e \mathbf{A} é o número de variáveis latentes usados para a construção do modelo (VALDERRAMA; BRAGA; POPPI, 2007).

3.5 SEGUNDO ESTÁGIO

Para o segundo estágio foram testadas duas redes neurais artificiais, o perceptron de múltiplas camadas e a rede de função de bases radiais.

As redes neurais extraem seu poder computacional através da sua habilidade de aprender e, portanto, de generalizar. A generalização refere-se ao fato de a rede produzir saídas adequadas para entradas que não estavam presentes durante seu treinamento. Esta capacidade faz que a rede possa resolver problemas altamente complexos e de grandes escalas. O uso das redes neurais oferece as seguintes vantagens:

- Não linearidade: Uma rede é constituída por estruturas não-lineares, sendo ela própria não-linear. É uma propriedade importante, particularmente se o sinal de entrada for não linear.
- Mapeamento de entrada e saída: Refere-se à capacidade de treinamento (modificação dos pesos sinápticos pela aplicação de um conjunto de amostras de treinamento) através de um método supervisionado.
- Adaptabilidade: Possuem capacidade inata de adaptar seus pesos sinápticos a modificações do ambiente. Uma rede treinada para operar em um ambiente específico pode ser facilmente retreinada para lidar com pequenas modificações no ambiente.

3.5.1 PERCEPTRON DE MULTIPLAS CAMADAS (MLP)

As redes neurais artificiais do tipo MLP são constituídas por uma camada de entrada, uma ou mais camadas ocultas, responsáveis pela separação dos padrões de fronteiras de decisão, contendo uma quantidade de neurônios a ser definida para cada problema, e uma camada de saída, a qual constrói combinações lineares das fronteiras de decisão formadas pelos neurônios ocultos, com um único neurônio responsável pela previsão do teor de café arábica na mistura. Na camada de entrada há um neurônio para cada componente principal ou variável independente utilizada. Em cada neurônio da rede, é realizado um somatório ponderado pelos

pesos sinápticos dos sinais provenientes dos neurônios da camada anterior (HAYKIN, 2001; MARINI, 2009).

A figura abaixo mostra a arquitetura de um perceptron de múltiplas camadas com duas camadas ocultas e uma de saída. O fluxo do sinal progride através da rede para a frente, da esquerda para a direita e de camada em camada.

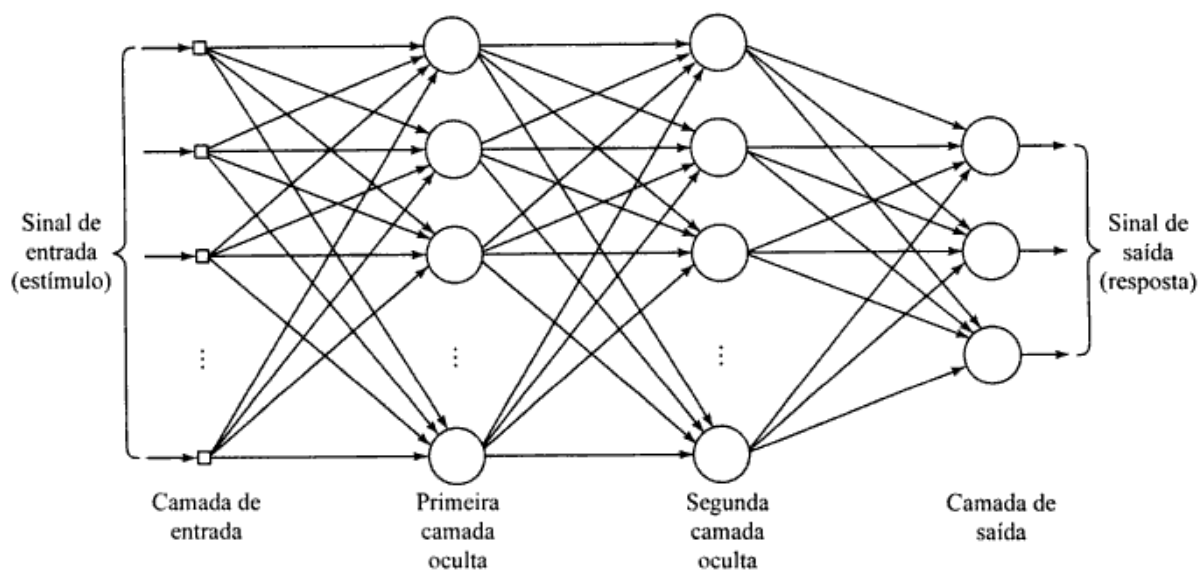


Figura 2. Representação de um perceptron de múltiplas camadas (MLP) (HAYKIN, 2001).

Os neurônios de saída constituem a camada de saída da rede. Os restantes formam as camadas ocultas. A primeira camada oculta é alimentada pela camada de entrada, constituída de unidades sensoriais. Por sua vez, as saídas resultantes da primeira camada oculta são aplicadas à próxima camada oculta, e assim sucessivamente. Cada neurônio oculto ou de saída da rede é programado para realizar dois cálculos:

- O cálculo do sinal funcional que aparece na saída de um neurônio, que é expresso como uma função não linear do sinal de entrada e dos pesos sinápticos associados com aquele neurônio;
- O cálculo do sinal de erro.

Para cada conexão entre os neurônios existe um peso w_{ij}^l associado, sendo i o índice do neurônio de entrada do sinal, j o neurônio de saída do sinal e l a camada

onde está localizado o neurônio de entrada. Em cada um dos N^l neurônios das L camadas da rede MLP é realizado um somatório ponderado pelos pesos sinápticos dos sinais provenientes dos neurônios da camada anterior. Esta soma, chamada de campo local induzido (4), é aplicada a uma função de ativação não linear (5) que irá produzir a saída do neurônio (HAYKIN, 2001).

$$v_i^l = \sum_{j=0}^{m_i} w_{ij}^l y_j^{l-1} \quad i = 1, 2, \dots, N^l; l = 1, 2, \dots, L. \quad (4)$$

$$y_i^l = \varphi(v_i^l) \quad i = 1, 2, \dots, N^l; l = 1, 2, \dots, L. \quad (5)$$

Nas equações (1) e (2) v_i^l é o campo local induzido do neurônio i na camada l , w_{ij}^l é o peso sináptico associado ao sinal de entrada y_j^{l-1} do neurônio j da camada anterior ($l - 1$). m_i é o número de entradas associadas ao neurônio i (HUANG; KANGAS; RASCO, 2007).

Para a MLP desenvolvida, será testada as funções de ativação do tipo logística (6), tangente hiperbólica sigmoide (7) e linear (8).

$$\varphi = \frac{1}{1+e^{-u}} \quad (6)$$

$$\varphi = \frac{2}{1+e^{-2u}} \quad (7)$$

$$\varphi = u \quad (8)$$

O algoritmo de *Levenberg-Marquardt* será utilizado para realizar o processo de aprendizagem das redes MLP construídas, tendo a função de modificar ordenadamente os pesos sinápticos da rede. Nesse algoritmo há um parâmetro que regula o tamanho do passo das correlações do peso, propondo uma solução de compromisso entre o algoritmo do gradiente descendente (retropropagação) e o método de Gauss-Newton (BHISOP, 2006).

3.5.2 REDES DE FUNÇÃO DE BASES RADIAIS (RBF)

A rede neural do tipo função de base radial consiste em uma camada de entrada, uma de saída e uma única camada oculta, as quais apresentam funções totalmente diferentes.

A rede RBF possui arquitetura mais simples, quando comparada à MLP, consistindo de duas camadas de pesos, sendo que a primeira possui os parâmetros das funções de base radial e a segunda cria combinações lineares das funções a fim de gerar uma saída. A camada de entrada é formada por nós de fonte, que são unidades sensoriais com a função de conectar a rede ao seu ambiente. A camada oculta realiza uma transformação não linear do espaço de entrada para o espaço oculto. Essas unidades ocultas fornecem um conjunto de funções radiais que constituem uma base arbitrária para os padrões de entrada. A camada de saída faz uma combinação linear de bases radiais, fornecendo a resposta da rede ao padrão de ativação aplicado à camada de entrada (HAYKIN, 2001). A imagem abaixo representa a estrutura de uma rede neural do tipo função de base radial.

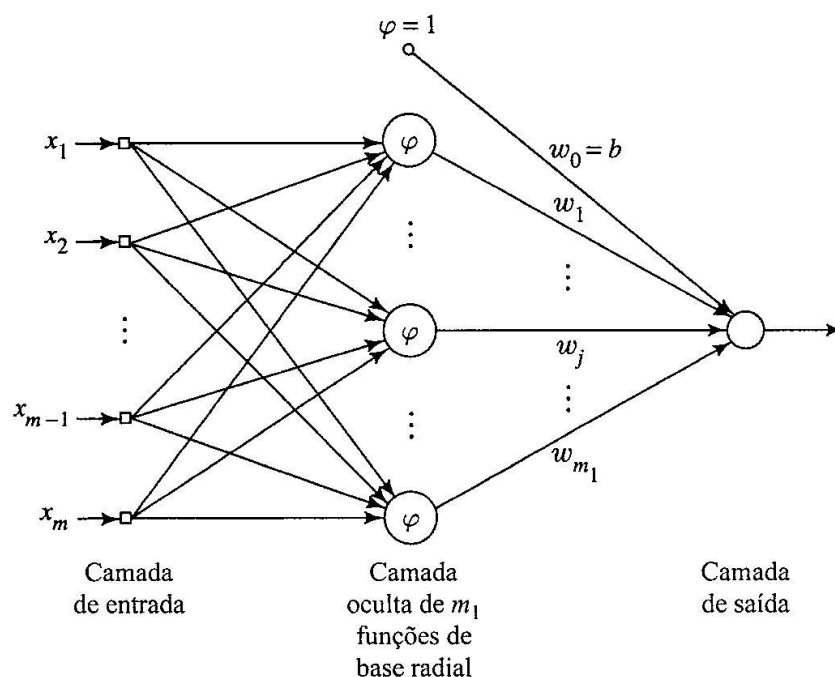


Figura 3. Representação de uma rede de função de bases radiais (RBF) (HAYKIN, 2001)

Foram utilizadas neste estudo, três tipos de funções de base radial: as funções do tipo gaussiana (9), multiquadrática (10) ou multiquadrática inversa (11).

$$\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (9)$$

$$\varphi(r) = (r^2 + \sigma^2)^{1/2} \quad (10)$$

$$\varphi(r) = \frac{1}{(r^2 + \sigma^2)^{1/2}} \quad (11)$$

Onde r é a distância euclidiana entre o centro da base e o vetor da amostra e σ é a largura da base, um parâmetro que controla a sua suavidade da função de interpolação. A função mais adequada foi determinada pelo método simplex sequencial.

O método de aprendizagem realizado foi híbrido, ou seja, em dois estágios. Primeiramente, a aprendizagem é auto organizada, estimando as localizações adequadas para os centros das funções de base radial na camada oculta. Por fim, faz-se uma aprendizagem supervisionada de rápida convergência, estimando a largura das bases (dispersão) e os pesos lineares da camada de saída (BHISOP, 2006; HAYKIN, 2001).

O processo de aprendizagem auto organizada foi realizado neste estudo utilizando um método de agrupamento que separa o conjunto de dados em subgrupos, o mais homogêneo possível. Para isso, foi utilizado o algoritmo de *K-means*, que posiciona os centros das funções de base radial apenas nas regiões densamente povoadas do espaço multidimensional da entrada (BHISOP, 2006).

A largura inicial das bases foi determinada de acordo com a equação (12) proposta por Haykin (2001).

$$\sigma = \frac{d_{\max}}{\sqrt{2m_1}} \quad (12)$$

Onde m_1 é o número de centros e d_{\max} é a distância máxima entre os centros escolhidos. Utilizando a equação (12), evita-se que as funções de base radial sejam pontiagudas ou planas demais. Depois, este valor foi otimizado com o método de busca direta Quase-Newton, a fim de minimizar o erro médio quadrático para as amostras de treinamento (BEVERIDGE; SCHECHTER, 1987).

Para a determinação dos pesos lineares da camada de saída da rede foi utilizado um procedimento direto que calcula a pseudo-inversa da matriz ϕ regularizada, conforme a equação (13) (BHISOP, 2006).

$$W = (\lambda I + \phi^T \phi)^{-1} \phi^T t \quad (13)$$

Sendo ϕ^T a matriz transposta de ϕ , que é definida por

$$\phi = \varphi_k \left(\|x_j - t_k\|^2, \sigma \right) \quad j = 1, 2, \dots, N; k = 1, 2, \dots, m_1 \quad (14)$$

onde x_j representa o j -ésimo vetor de entrada da amostra; k a k -ésima função de base radial e t_k é o centro da base k (TUDU et al., 2009).

3.6 OTIMIZAÇÃO SIMPLEX

Na etapa de construção das redes neurais, uma série de parâmetros devem ser escolhidos para que o modelo criado seja o melhor possível para a previsão da porcentagem de arábica.

O simplex é uma figura regular que se desloca sobre uma superfície, de modo a evitar regiões de resposta não satisfatória. No espaço n -dimensional o simplex é um poliedro com faces planas contendo $n+1$ vértices, onde n é o número de variáveis (contínuas ou discretas) independentes. O método é um procedimento recorrente, que tende a levar o simplex a um valor ótimo através da reflexão de pontos específicos. Uma vez nas vizinhanças do ótimo, o simplex pode sofrer

contração com o objetivo de determinar uma posição mais precisa (BONA et al., 2000).

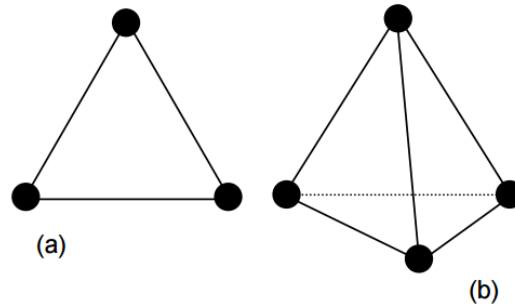


Figura 4. Interpretação geométrica do simplex para duas variáveis (a) e para três variáveis (b). Fonte: Bona (2000).

O método simplex sequencial utilizado consiste em uma figura de $n+1$ vértices, que se altera em tamanho e forma adaptando-se melhor ao espaço de resposta, onde n é o número de variáveis. A otimização é iniciada atribuindo-se limites inferiores (L_i) e superiores (U_i) para cada fator que será controlado. As coordenadas do simplex inicial são calculadas utilizando as equações 15, 16, 17 e 18; onde t é a distância entre dois vértices (geralmente tomada como 1), e são formadas conforme a matriz \mathbf{M} , onde as colunas representam os componentes dos vértices, numerados de 1 até $n+1$ e as linhas representam as coordenadas, $i = 1$ até n (LINK, 2013).

$$p = \frac{t}{n\sqrt{2}}(\sqrt{n+1}+n-1) \quad (15)$$

$$q = \frac{t}{n\sqrt{2}}(\sqrt{n+1}-1) \quad (16)$$

$$m_1 = L_i + p(U_i - L_i) \quad (17)$$

$$m_2 = L_i + q(U_i - L_i) \quad (18)$$

$$M = \begin{bmatrix} L_1 & m_1 & m_2 & m_2 \\ L_1 & m_2 & m_1 & m_2 \\ \dots & \dots & \dots & \dots \\ L_1 & m_2 & m_2 & m_1 \end{bmatrix} \text{ Matriz } n \times n+1$$

Com as respostas obtidas em cada iteração, os vértices do simplex são ordenados de acordo com seus valores em **B** (melhor), **N** (intermediários) e **W** (pior). O novo simplex é determinado rejeitando-se o vértice correspondente à pior resposta e substituindo-se esse vértice pela sua reflexão na hiperface definida pelos vértices restantes (BONA et al., 2000).

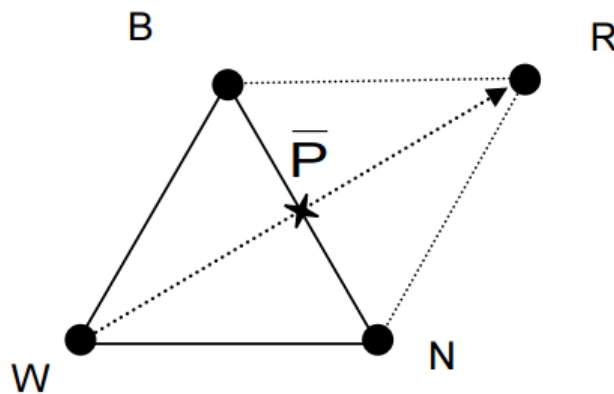


Figura 5. Movimento de reflexão em um simplex para duas variáveis. Fonte: Bona (2000).

O algoritmo utiliza as operações: reflexão, expansão, contração externa e interna e encolhimento (GAO; HAN, 2012). A otimização segue através de uma sequência das operações citadas até que o valor da resposta varie apenas dentro da tolerância estabelecida ou pela visualização gráfica que também pode ser utilizada como um critério de parada da otimização.

A tabelas abaixo mostram os parâmetros a serem otimizados, bem como os intervalos de variação, para as redes do tipo MLP e RBF, respectivamente.

Tabela 1. Parâmetros a serem otimizados para a rede do tipo MLP.

Parâmetros	Varição
Neurônios da primeira camada	6 a 13
Neurônios da segunda camada	0 a 8
Função de Ativação da Camada oculta	Logística ou tangente hiperbólica sigmoide
Função de Ativação da Camada de Saída	Logística tangente hiperbólica sigmoide ou linear
Função de Normalização da Camada de Entrada	Autoescalamento, mínimo e máximo [-1,1], identidade, vetor unitário
Função de Normalização da Camada de Saída	Autoescalamento, mínimo e máximo [-1,1]
Componentes Principais	9 a 13
Parâmetro de Regularização	0 a 0,4

Tabela 2. Parâmetros a serem otimizados para a rede do tipo RBF.

Parâmetros	Varição
Quantidade de bases radiais	5 a 80
Função de base radial	Gaussiana, multiquadrática ou multiquadrática inversa
Função de distância para o algoritmo <i>k-means</i>	Distância euclidiana, <i>cityblock</i> , cosseno ou correlação
Função de normalização dos dados de entrada	Autoescalamento, mínimo e máximo [-1,1], identidade, vetor unitário
Componentes principais	8 a 13
Parâmetro de regularização	0 a 1
Quantidade de bases radiais	5 a 80
Função de base radial	Gaussiana, multiquadrática ou multiquadrática inversa

3.7 IMPLEMENTAÇÃO COMPUTACIONAL

Todas as análises matemáticas e/ou estatísticas, otimização e redes neurais artificiais foram realizadas no software MATLAB R2008b (The MathWorks Inc., Natick, USA).

4. RESULTADOS E DISCUSSÕES

4.1 FTIR

Para cada tipo de tratamento do espectro, foram criadas 100 redes. A melhor rede de cada tipo foi selecionada pelo menor erro quadrático médio e maior capacidade de previsão.

4.1.2 PCA COMO PRIMEIRO ESTÁGIO

A tabela abaixo apresenta o resultado da otimização simplex para a rede do tipo MLP com a PCA como método de primeiro estágio.

Tabela 3. Otimização simplex e desempenho das redes MLP (PCA).

Espectros	Puro	1ª Derivada
Neurônios na primeira camada	5	5
Neurônios na segunda camada	0	1
Função de ativação da camada oculta	Logística	Tangente hiperbólica sigmoide
Função de ativação da camada se saída	Logística	Logística
Função de normalização da camada de entrada	Máximo e mínimo	Máximo e mínimo
Componentes principais	14	8
Parâmetro de regularização	$5,004 \times 10^{-4}$	$6,5144 \times 10^{-5}$
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0051	0,007
Erro absoluto médio (MAE)	7,16%	6,29%
Coeficiente de correlação (r)	0,9736	0,9631
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,005	0,0101
Erro absoluto médio (MAE)	5,3%	8,04%
Coeficiente de correlação (r)	0,9479	0,9091

Pode-se observar que para os valores de teste, o espectro puro apresentou menores valores de erro absoluto médio, assim como maior valor de coeficiente de correlação.

Os gráficos abaixo apresentam a relação entre os valores previstos e observados durante o teste das redes do tipo MLP.

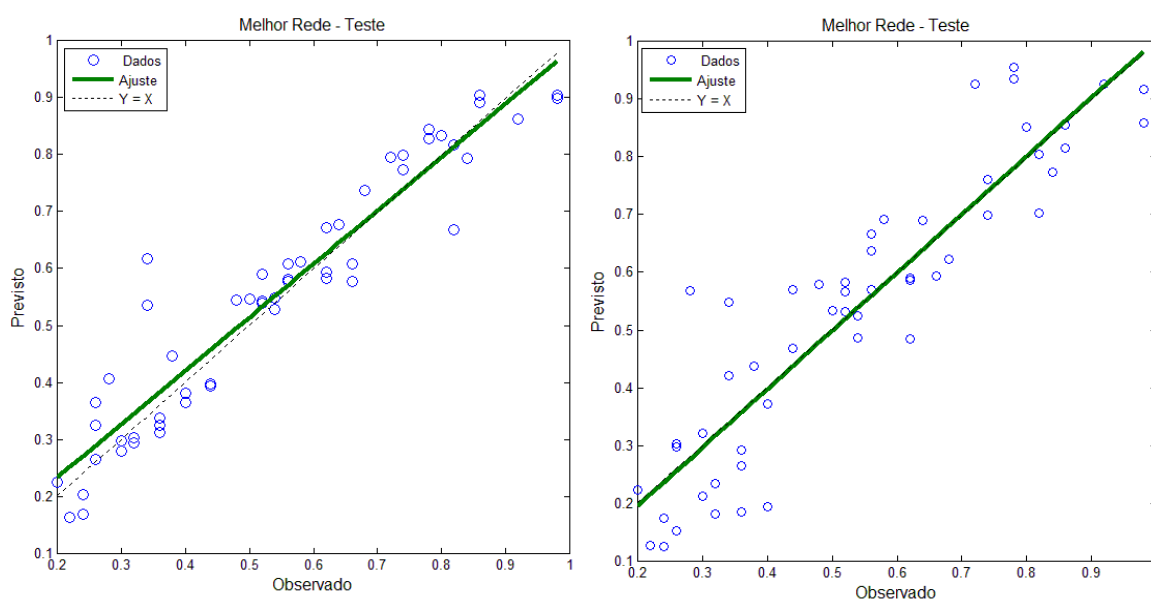


Figura 6. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede MLP com PCA como primeiro estágio (FTIR).

A tabela abaixo apresenta o resultado da otimização simplex para a rede do tipo RBF com a PCA como método de primeiro estágio.

Tabela 4. Otimização simplex e desempenho das redes RBF (PCA).

Espectros	Puro	1ª Derivada
Quantidade de bases radiais	18	45
Tipo de base radial	Multiquadrática	Gaussiana
Função de distancia k-means	<i>Cityblock</i>	<i>Cityblock</i>
Função de normalização dos dados de entrada	Máximo e mínimo	Máximo e mínimo
Componentes principais	11	8
Parâmetro de regularização	6.2891×10^{-05}	5.5175×10^{-05}
Desempenho da melhor rede para os dados de treinamento		

Erro quadrado médio (MSE)	0,01018	0,0103
Erro absoluto médio (MAE)	8,02%	8,34%
Coefficiente de correlação (r)	0,9419	0,9448
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,0072	0,0088
Erro absoluto médio (MAE)	6,37%	8,1%
Coefficiente de correlação (r)	0,9228	0,9020

Os gráficos abaixo apresentam a relação entre os valores previstos e observados durante o teste das redes do tipo RBF.

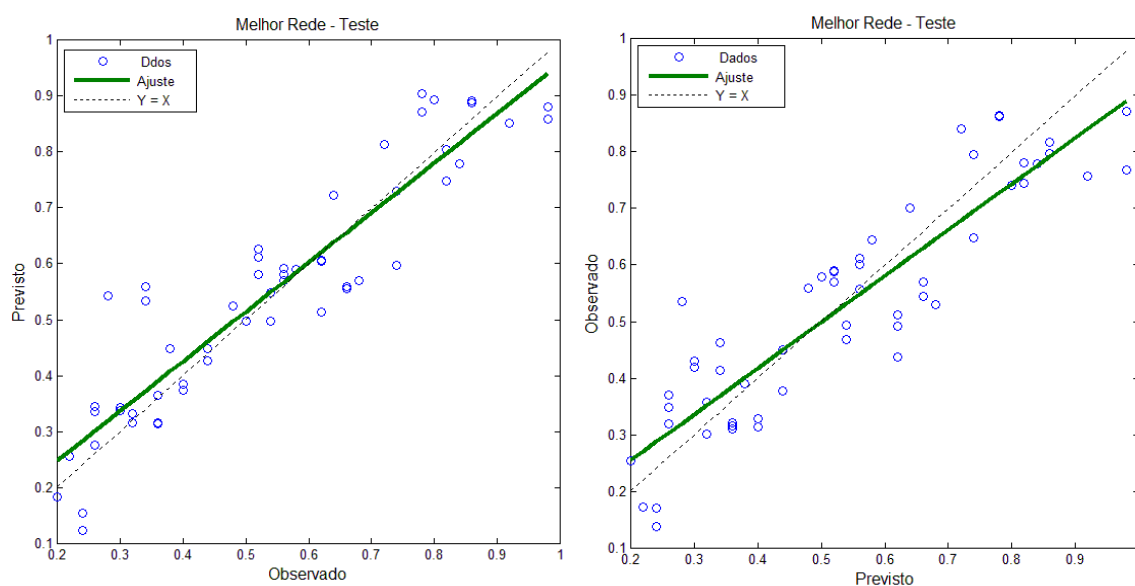


Figura 7. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede RBF com PCA como primeiro estágio (FTIR).

4.1.3 PLS COMO PRIMEIRO ESTÁGIO

A tabela a seguir apresenta o resultado da otimização simplex para a rede do tipo MLP com o PLS como método de primeiro estágio.

Tabela 5. Otimização simplex e desempenho das redes MLP (PLS).

Espectros	Puro	1ª Derivada
Neurônios na primeira camada	5	5
Neurônios na segunda camada	1	1
Função de ativação da camada oculta	Tangente hiperbólica sigmoide	Tangente hiperbólica sigmoide
Função de ativação da camada de entrada	Tangente hiperbólica sigmoide	Tangente hiperbólica sigmoide
Função de normalização da camada de entrada	Autoescalamento	Autoescalamento
Variáveis Latentes	12	9
Parâmetro de regularização	$6,6915 \times 10^{-05}$	$2,0083 \times 10^{-04}$
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0037	0,0041
Erro absoluto médio (MAE)	4,46%	5,15%
Coefficiente de correlação (r)	0,9808	0,9783
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,0034	0,0059
Erro absoluto médio (MAE)	4,76%	6,03%
Coefficiente de correlação (r)	0,9662	0,9382

A rede do tipo MLP utilizando os scores do PLS como alimentação, apresentou excelentes resultados de MSE e MAE para treinamento, tanto com o espectro puro quando com a primeira derivada do espectro, com erro absoluto médio de aproximadamente 5% e coeficiente de correlação próximo a 0,98. Os gráficos abaixo apresentam a relação entre os valores reais de porcentagem de arábica e os valores previstos pela rede.

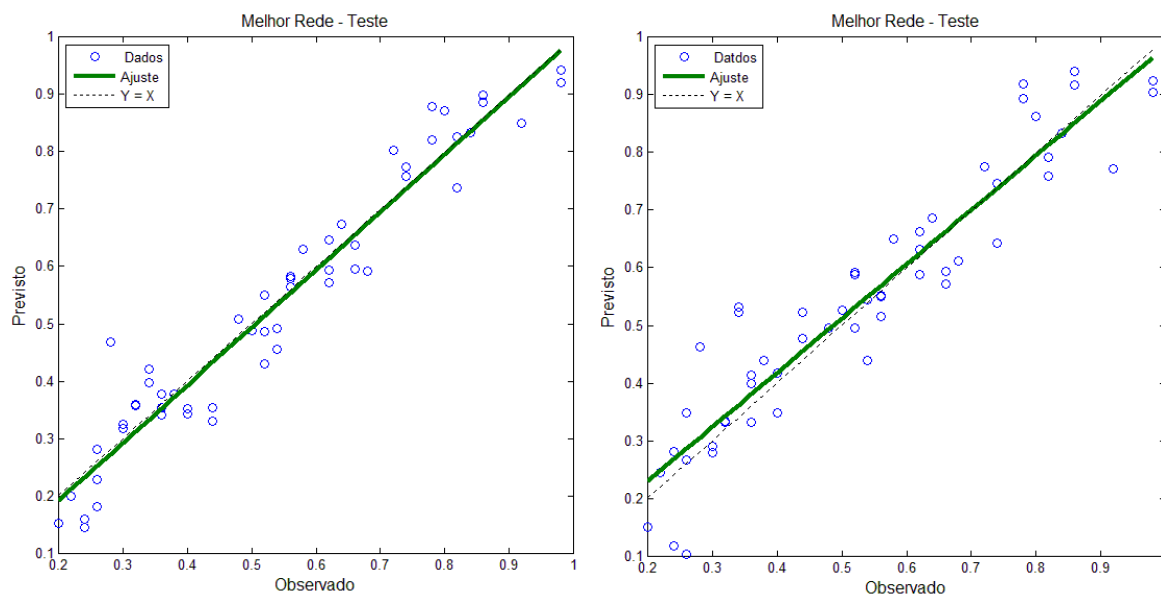


Figura 8. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede MLP com PLS como primeiro estágio (FTIR).

A tabela a seguir apresenta os parâmetros otimizados pelo método simplex, na rede do tipo RBF com o PLS como método de redução de dimensionalidade.

Tabela 6. Otimização simplex e desempenho das redes RBF (PLS).

Espectros	Puro	1ª Derivada
Quantidade de bases radiais	10	8
Tipo de base radial	Multiquadrática	Gaussiana
Função de distancia k-means	<i>Cityblock</i>	<i>Cityblock</i>
Função de normalização dos dados de entrada	Autoescalamento	Máximo e mínimo
Variáveis Latentes	13	13
Parâmetro de regularização	0,0013	0,0015
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0063	0,0046
Erro absoluto médio (MAE)	6,17%	5,46%
Coeficiente de correlação (r)	0,9668	0,9757
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,0045	0,0072
Erro absoluto médio (MAE)	5,28%	6,62%
Coeficiente de correlação (r)	0,9537	0,9213

Para a rede do tipo RBF, o espectro puro apresentou melhores valores, quando comparados à primeira derivada. Para o treinamento, a rede com espectro puro apresentou excelente valor de previsão, com erro próximo a 5%. As imagens abaixo comparam os valores reais de porcentagem e os valores previstos pela rede para o espectro puro e primeira derivada.

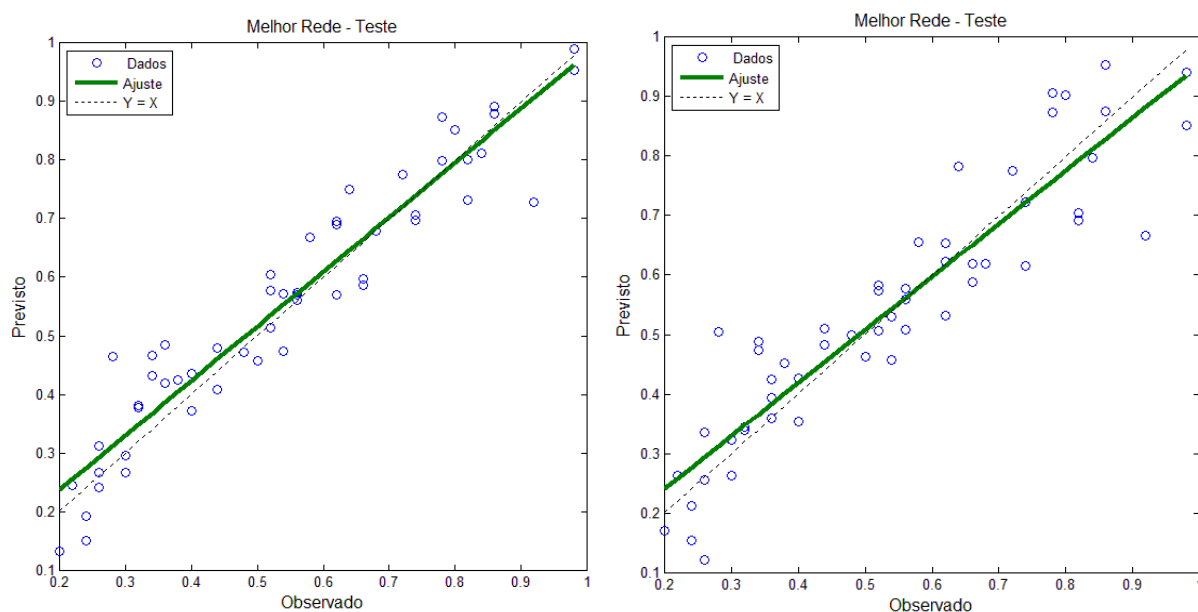


Figura 9. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede RBF com PLS como primeiro estágio (FTIR).

De maneira geral, pôde-se observar que o espectro puro conduz a melhores resultados que a primeira derivada. Já em relação ao método de redução de dimensionalidade de dados, as redes que foram alimentadas com os *scores* dos mínimos quadrados parciais apresentaram menor erro absoluto médio e maior coeficiente de correlação, quando comparado com as redes alimentadas com os *scores* da análise de componentes principais.

4.2 NIR

Para cada tipo de tratamento do espectro, foram criadas 100 redes. A melhor rede de cada tipo foi selecionada pelo menor erro quadrático médio e maior capacidade de previsão.

4.2.1 PCA COMO PRIMEIRO ESTÁGIO

A tabela abaixo apresenta os parâmetros otimizados da rede do tipo MLP com a PCA como alimentação.

Tabela 7. Otimização simplex e desempenho das redes MLP (PCA).

Espectros	Puro	2ª Derivada
Neurônios na primeira camada	5	6
Neurônios na segunda camada	1	1
Função de ativação da camada oculta	Tangente hiperbólica sigmoide	Tangente hiperbólica sigmoide
Função de ativação da camada se saída	Tangente hiperbólica sigmoide	Tangente hiperbólica sigmoide
Função de normalização da camada de entrada	Autoescalamento	Máximo e mínimo
Componentes principais	5	9
Parâmetro de regularização	$4,6026 \times 10^{-05}$	1×10^{-04}
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0085	0,0079
Erro absoluto médio (MAE)	6,9%	6,49%
Coefficiente de correlação (r)	0,9475	0,9509
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,005	0,0062
Erro absoluto médio (MAE)	5,5%	5,87%
Coefficiente de correlação (r)	0,9681	0,9598

Pode-se observar pela tabela que a rede neural com a segunda derivada, para o treinamento, apresentou menores valores de erro quadrático médio e erro absoluto médio, assim como maior coeficiente de correlação que a rede com os espectros puros. Já para os dados de teste, a rede com espectros puros como entrada apresentou menores valores de erro assim como maior coeficiente de correlação.

As figuras a seguir apresentam um comparativo entre os valores previstos pelas melhores redes MLP, com o espectro puro e sua segunda derivada como entrada.

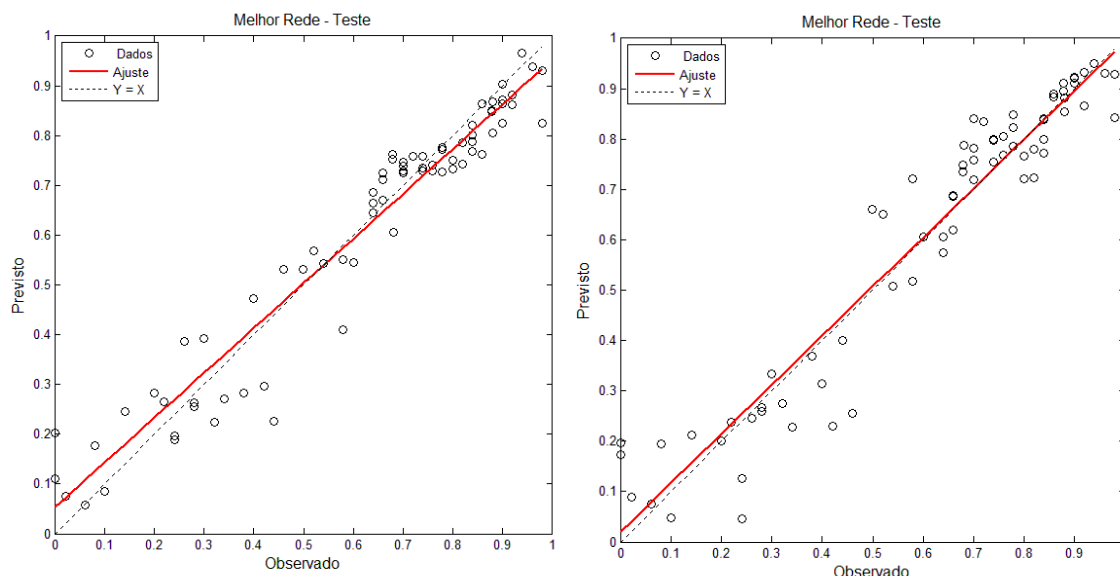


Figura 10. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede MLP com PCA como primeiro estágio (NIR).

Os resultados da otimização simplex para a RBF com a PCA como entrada são mostrados abaixo.

Tabela 8. Otimização simplex e desempenho das redes RBF (PCA).

Espectros	Puro	2ª Derivada
Quantidade de bases radiais	25	20
Tipo de base radial	Gaussiana	Gaussiana
Função de distancia k-means	<i>Cityblock</i>	<i>Cityblock</i>
Função de normalização dos dados de entrada	Linear	Linear
Função de normalização dos dados de entrada	Máximo e mínimo	Máximo e mínimo
Variáveis Latentes	11	10
Parâmetro de regularização	$9,1343 \times 10^{-4}$	$3,0627 \times 10^{-4}$
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0133	0,0191

Erro absoluto médio (MAE)	8,85%	10,88%
Coeficiente de correlação (r)	0,9156	0,8927
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,0057	0,0234
Erro absoluto médio (MAE)	5,8%	11,69%
Coeficiente de correlação (r)	0,9644	0,7743

A rede com os espectros puros como entrada apresentaram aproximadamente a metade do erro absoluto médio, quando comparada com a rede com a segunda derivada dos espectros como entrada, assim como um coeficiente de correlação muito maior.

As figuras a seguir apresentam um comparativo entre os valores previstos e observados para a RBF para o espectro puro e segunda derivada com a PCA como entrada.

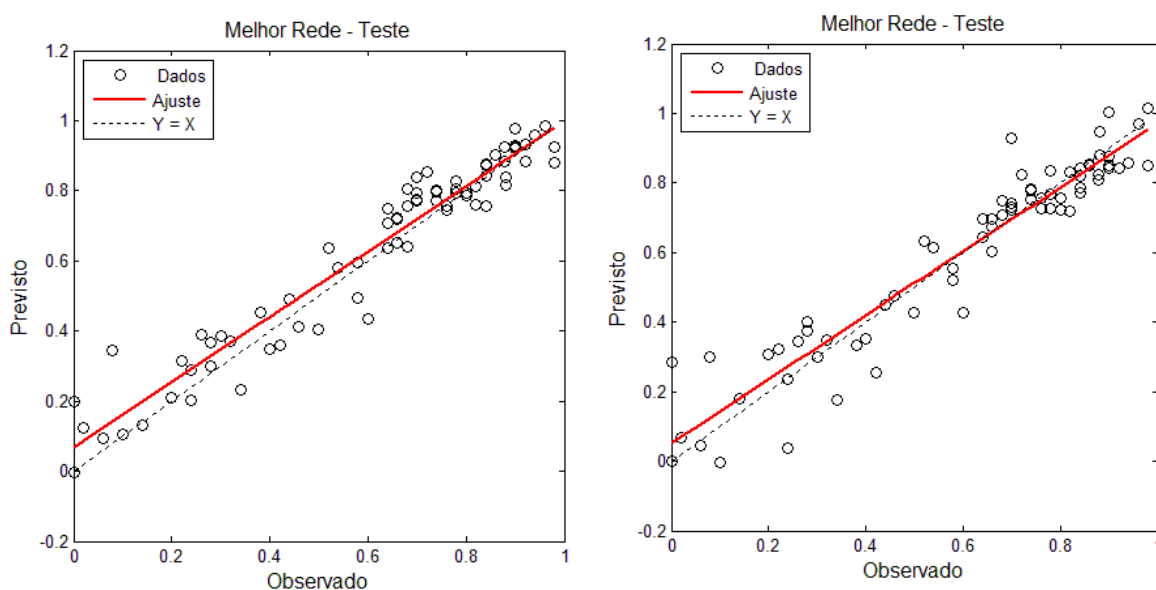


Figura 11. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede RBF com PCA como primeiro estágio (NIR).

4.2.2 PLS COMO PRIMEIRO ESTÁGIO

A otimização simplex para a rede MLP com o PLS como método de redução de dimensionalidade de dados é mostrada na tabela abaixo.

Tabela 9. Otimização simplex e desempenho das redes MLP (PLS).

Espectros	Puro	2ª Derivada
Neurônios na primeira camada	5	5
Neurônios na segunda camada	1	2
Função de ativação da camada oculta	Tangente hiperbólica sigmoide	Tangente hiperbólica sigmoide
Função de ativação da camada se saída	Tangente hiperbólica sigmoide	Linear
Função de normalização da camada de entrada	Máximo e mínimo	Máximo e mínimo
Componentes principais	9	13
Parâmetro de regularização	$5,9144 \times 10^{-05}$	$7,3243 \times 10^{-05}$
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0076	0,005
Erro absoluto médio (MAE)	6,88%	5,21%
Coefficiente de correlação (r)	0,9527	0,9690
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,0103	0,0101
Erro absoluto médio (MAE)	7,98%	7,84%
Coefficiente de correlação (r)	0,9439	0,9426

Tanto os espectros puros quanto a segunda derivada, para a rede do tipo MLP com o PLS como método de redução de dimensionalidade de dados, apresentaram valores semelhantes de erro e coeficiente de correlação, sendo os valores para a segunda derivada ligeiramente menores que os do espectro puro.

As imagens abaixo mostram a relação entre os valores previstos e observados para a rede MLP com PLS como entrada.

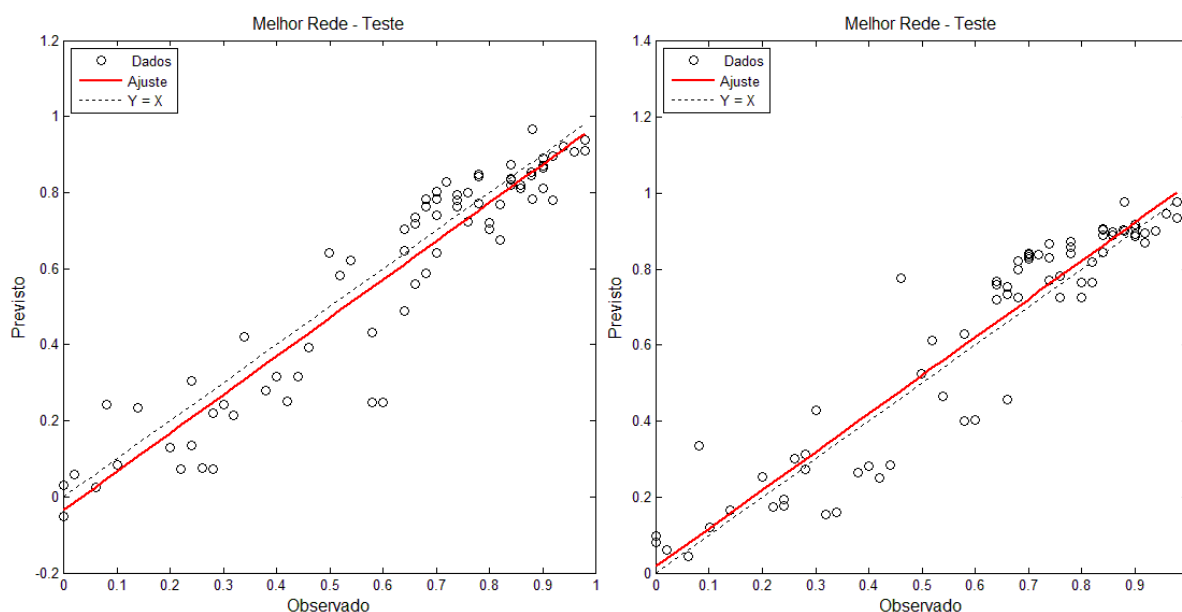


Figura 12. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede MLP com PLS como primeiro estágio (NIR).

A otimização simplex para a rede do tipo RBF com o PLS como primeiro estágio é mostrada na tabela a seguir:

Tabela 10. Otimização simplex e desempenho das redes RBF (PLS).

Espectros	Puro	2ª Derivada
Quantidade de bases radiais	6	5
Tipo de base radial	Gaussiana	Gaussiana
Função de distancia k-means	City block	Euclidiana
Função de normalização dos dados de entrada	Linear	Linear
Função de normalização dos dados de entrada	Máximo e mínimo	Vetor unitário
Variáveis Latentes	12	11
Parâmetro de regularização	0,1107	0,0734
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0133	0,016
Erro absoluto médio (MAE)	9,15%	10,03%
Coeficiente de correlação (r)	0,9173	0,8987
Desempenho da melhor rede para os dados de teste		

Erro quadrado médio (MSE)	0,0166	0,0216
Erro absoluto médio (MAE)	10,66%	12,49%
Coefficiente de correlação (r)	0,9392	0,9255

Foi observado que tanto para o espectro puro quanto para a segunda derivada, os modelos não apresentaram resultados satisfatórios, tendo erros absolutos próximos a 10%, assim como coeficiente de correlação relativamente baixos, quando comparado com os outros modelos construídos.

As imagens abaixo apresentam os valores previstos e observados para a RBF com o PLS como estágio linear.

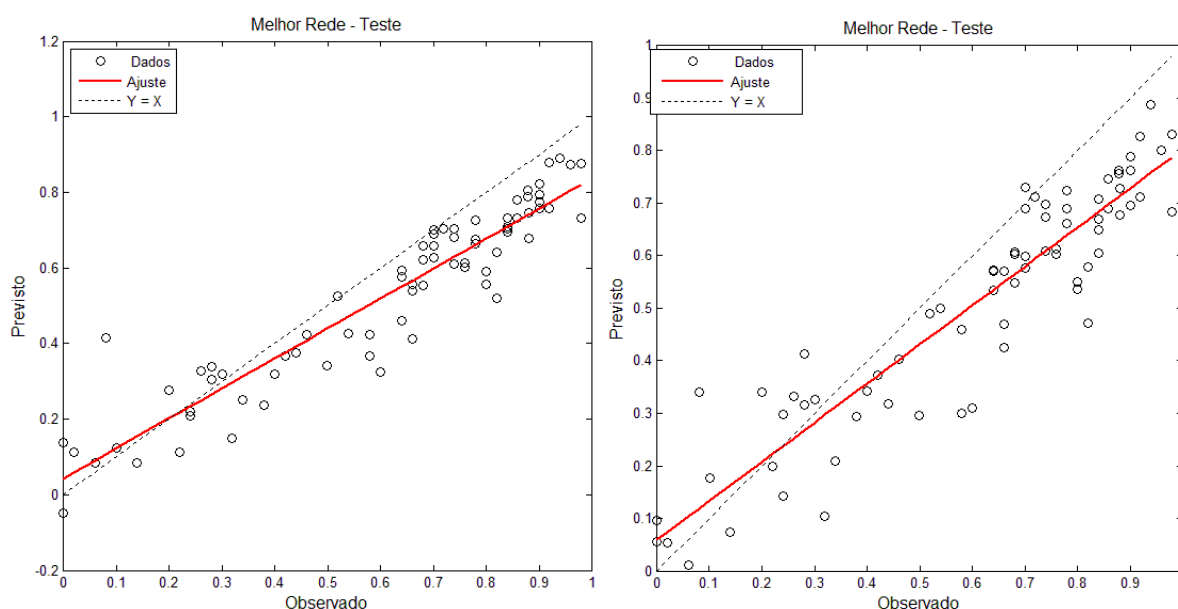


Figura 13. Comparação entre os valores previstos e observados para o espectro puro e primeira derivada, respectivamente, na rede RBF com PLS como primeiro estágio (NIR).

4.3. COMPARAÇÃO ENTRE NIR E FTIR

De maneira geral, pôde-se observar que os espectros obtidos no FTIR apresentaram-se melhores que os espectros NIR, conduzindo a menores erros e maiores coeficientes de correlação. A Tabela 11 abaixo compara a melhores redes alimentadas com espectros FTIR e NIR.

Tabela 11. Comparação entre a melhor rede alimentada com espectros FTIR e a melhor alimentada com espectros NIR.

Rede	FTIR – MLP (PLS)	NIR – RBF (PCA)
Desempenho da melhor rede para os dados de treinamento		
Erro quadrado médio (MSE)	0,0037	0,0133
Erro absoluto médio (MAE)	4,46%	8,85%
Coeficiente de correlação (r)	0,9808	0,9156
Desempenho da melhor rede para os dados de teste		
Erro quadrado médio (MSE)	0,0034	0,0057
Erro absoluto médio (MAE)	4,76%	5,80%
Coeficiente de correlação (r)	0,9662	0,9644

5. CONCLUSÃO

Foi possível construir um modelo capaz de prever a porcentagem de café arábica em misturas com café robusta de maneira satisfatória (com erro menor que 5%). A partir dos resultados, pôde-se concluir que os espectros de FTIR conduziram a melhores resultados que os espectros NIR. O método de redução de dimensionalidade de dados mais satisfatório foi o PLS. Já a rede que apresentou melhor desempenho na criação do modelo de previsão foi a rede do tipo MLP, a qual, de maneira geral, obteve menores resultados de erro absoluto médio e maiores valores de coeficiente de correlação.

6. REFERÊNCIAS

- ABIC. **Associação Brasileira da Indústria de Café - História**. Disponível em: <<http://www.abic.com.br/publique/cgi/cgilua.exe/sys/start.htm?sid=38>>. Acesso em: 9 abr. 2015.
- BEVERIDGE, G. S. G.; SCHECHTER, R. S. **Optimization theory and practice**. Tokyo: Mac Graw-Hill & Sons, 1987.
- BHISOP, C. M. **Pattern Recognition and Machine Learning**. New York: Springe, 2006.
- BONA, E. et al. Aplicativo para otimização empregando o método simplex seqüencial. **Acta Scientiarum. Technology**, v. 22, n. 5, p. 1201–1206, 13 maio 2000.
- BRIANDET, R.; KEMSLEY, K.; WILSON, H. Approaches to Adulteration Detection in Instant Coffees using Infrared Spectroscopy and Chemometrics. **Journal of the Science of Food and Agriculture**, v. 71, p. 359–366, 1996.
- BUENO, A. **Instituto de Química Caracterização de Petróleo por Espectroscopia no Infravermelho Próximo Dissertação de**. [s.l.] Universidade Estadual de Campinas, 2004.
- FARAH, A. Coffee as a speciality and functional beverage. In: **Functional and speciality beverage technology**. Cambridge: Woodhead Publishing Limited, 2009. p. 370–395.
- FARO, A. et al. Análise por componentes principais de espectros nexafs na especiação do molibdênio em catalisadores de hidrotratamento. **Química Nova**, v. 33, n. 6, p. 1342–1347, 2010.
- FERNANDES, S. M. et al. Constituintes químicos e teor de extrato aquoso de cafés arábica (*coffea arabica* L.) e conilon (*coffea canephora pierre*) torrados. **Ciência e Agrotecnologia**, v. 27, n. 5, p. 1076–1081, out. 2003.
- GAO, F.; HAN, L. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. **Comput Optim Appl**, v. 51, p. 259–277, 2012.
- HAYKIN, S. **Redes Neurais: princípios e prática**. 2ª edição ed.Porto Alegre: Bookman, 2001.
- HELFER, G. A. et al. Pulicação de métodos de análise multivariada no controle qualitativo de essências alimentícias empregando espectroscopia no infravermelho médio. **Ciência e Tecnologia de Alimentos**, v. 26, n. 4, p. 779–786, dez. 2006.

HUANG, Y.; KANGAS, L.; RASCO, B. A. Applications of artificial neural networks (ANNs) in food science. **Critical Reviews in Food Science and Nutrition**, v. 47, p. 13, 2007.

ICO. **International Coffee Organization**. Disponível em: <http://www.ico.org/botanical.asp?section=About_Coffee>. Acesso em: 9 jul. 2015.

JAVIDNIA, K. et al. Discrimination of edible oils and fats by combination of multivariate pattern recognition and FT-IR spectroscopy: a comparative study between different modeling methods. **Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy**, v. 104, p. 175–81, mar. 2013.

LINK, J. **Estudo dos Genótipos de Café Arábica Utilizando FTIR e Redes Neurais Artificiais**. [s.l.] Universidade Tecnológica Federal do Paraná, 2013.

LYMAN, J. et al. Analysis of Brewed Coffee: Effect of Roasting Conditions. **Journal of Agricultural and Food Chemistry**, v. 51, p. 3268–3273, 2003.

MAPA. **Ministério da agricultura, pecuária e abastecimento**. Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/cafe/saiba-mais>>. Acesso em: 9 abr. 2015.

MAPA. **Ministério da agricultura, pecuária e abastecimento: café - 2014**. Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/cafe>>. Acesso em: 9 abr. 2015.

MARINI, F. Artificial neural networks in foodstuff analyses: Trends and perspectives A review. **Analytica Chimica Acta**, p. 121–131, 2009.

MARQUETTI, I. **Classificação de genótipos de café arábica usando espectroscopia de infravermelho próximo**. [s.l.] Universidade Tecnológica Federal do Paraná, 2014.

MELSSSEN, W.; WEHRENS, R.; BUYDENS, L. Supervised Kohonen networks for classification problems. **Chemometrics and Intelligent Laboratory Systems**, v. 83, n. 2, p. 99–113, set. 2006.

MORAES, L. DE G. **Uma abordagem alternativa de behavioral scoring usando modelagem híbrida de dois estágios com regressão logística e redes neurais**. Porto Alegre: Univerisdade Federal do Rio Grande do Sul, 2012.

MORGANO, M. A. et al. Determinação de proteína em café cru por espectroscopia NIR e regressão PLS. **Ciência e Tecnologia de Alimentos**, v. 25, n. 1, p. 25–31, mar. 2005.

PERALTA, P. et al. Alternativas quimiométricas para a resolução de problemas analíticos clássicos: determinação espectrofotométrica de misturas de zircônio e háfnio. **Química Nova**, v. 20, n. 5, p. 469–474, out. 1997.

SANTOS, K. et al. Classificação de Café Torrado e Moído Empregando Espectroscopia NIR e Análise Multivariada. 2006.

SHLENS, J. A Tutorial on Principal Component Analysis. 3 abr. 2014.

SILVERSTEIN, R.; WEBSTER, F.; KIEMLE, D. **Identificação espectrométrica de compostos orgânicos**. 7. ed. Rio de Janeiro: LTC, 2007.

TUDU, B. et al. Electronic nose for black tea quality evaluation by an incremental RBF network. **Sensors and Actuators B: Chemical**, v. 138, n. 1, p. 90–95, abr. 2009.

VALDERRAMA, P.; BRAGA, J.; POPPI, R. Variable Selection, Outlier Detection, and Figures of Merit Estimation in a Partial Least-Squares Regression Multivariate Calibration Model. A Case Study for the Determination of Quality Parameters in the Alcohol Industry by Near-Infrared Spectroscopy. **Journal of Agricultural and Food Chemistry**, v. 55, n. 21, p. 8331–8338, 2007.

VARELLA, C. **Análise Multivariada Aplicada as Ciências Agrárias: Análise de Componentes Principais**. [s.l.] Programa de Pós-Graduação em Agronomia - Universidade Federal Rural do Rio de Janeiro, 2008.

WANG, J. et al. Fourier Transform Infrared Spectroscopy for Kona Coffee Authentication. **Journal of Food Science**, v. 5, n. 74, p. 1750–3841, 2009.

YAN, S. Principle Component Analysis and Partial Least Squares: Two Dimension Reduction Techniques for Regression. **Casualty Actuarial Society**, 2008.