

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
MESTRADO EM COMPUTAÇÃO APLICADA

CARLOS ALBERTO PROENÇA

REDES NEURAIS ARTIFICIAIS PARA PREDIÇÃO DOS TEORES DE MATÉRIA
ORGÂNICA E ARGILA DO SOLO NA REGIÃO DOS CAMPOS GERAIS UTILIZANDO
ESPECTROSCOPIA DE REFLECTÂNCIA DIFUSA

PONTA GROSSA
2012

CARLOS ALBERTO PROENÇA

REDES NEURAIS ARTIFICIAIS PARA PREDIÇÃO DOS TEORES DE MATÉRIA
ORGÂNICA E ARGILA DO SOLO NA REGIÃO DOS CAMPOS GERAIS UTILIZANDO
ESPECTROSCOPIA DE REFLECTÂNCIA DIFUSA

Dissertação apresentada no Programa de Pós-Graduação em Computação Aplicada, curso de Mestrado em Computação Aplicada da Universidade Estadual de Ponta Grossa, como requisito para obtenção do título de Mestre.

Orientação: Dra. Alaine Margarete Guimarães

PONTA GROSSA
2012

Ficha Catalográfica Elaborada pelo Setor Tratamento da Informação Bicen/UEPG

P964r Proença, Carlos Alberto
Redes neurais artificiais para predição dos teores de matéria orgânica e argila do solo na região dos Campos Gerais utilizando espectroscopia de reflectância difusa / Carlos Alberto Proença. Ponta Grossa, 2012.
78 f.

Dissertação (Mestrado em Computação Aplicada - área de concentração Computação Aplicada a Agricultura), Universidade Estadual de Ponta Grossa. Orientador: Profa. Dra. Alaine Margarete Guimãres.

1. Redes neurais artificiais. 2. Argila. 3. Análise de regressão. 4. Matéria orgânica I. Guimarães, Alaine Margarete. II. Universidade Estadual de Ponta Grossa. Mestrado em Computação Aplicada. III. T.

CDD: 006.3

TERMO DE APROVAÇÃO

CARLOS ALBERTO PROENÇA

REDES NEURAIS ARTIFICIAIS PARA PREDIÇÃO DOS TEORES DE MATÉRIA ORGÂNICA E ARGILA DO SOLO NA REGIÃO DOS CAMPOS GERAIS UTILIZANDO ESPECTROSCOPIA DE REFLECTÂNCIA DIFUSA

Dissertação aprovada como requisito para obtenção do grau de Mestre no Programa de Pós-Graduação em Computação Aplicada, curso de Mestrado em Computação Aplicada da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora.

Orientador:


Alaine Margarete Guimarães
UEPG


Eduardo Fávero Caires
UEPG


Fabricio Pinheiro Povh
Fundação ABC

Ponta Grossa, 1 de agosto de 2012.

AGRADECIMENTOS

A Deus por todas as graças concedidas durante toda a minha vida e agora neste momento, dando-me sabedoria necessária para desenvolver este estudo.

Agradeço a toda a minha família, especialmente à minha esposa Rosângela, as filhas Ana Claudia e Caroline e minha Mãe Zura pela paciência e apoio, fundamentais neste período de conciliação entre família, trabalho e estudos.

A minha orientadora Dra. Alaine Margarete Guimarães pela oportunidade de me ajudar a elaborar este trabalho, contando com toda a sua experiência e conhecimento, sempre procurando e incentivando a desenvolvê-lo da melhor forma possível.

A Fundação ABC, por todo o apoio durante este período, disponibilizando todos os recursos necessários para esta pesquisa, em especial, ao Laboratório de Solos pelas análises das amostras por diferentes métodos. Também agradeço a todos meus colegas de trabalho pelo suporte dado nos dias em que me ausentei para estar na UEPG e aos colegas Carlos André Schipanski e Fabricio Pinheiro Povh pelas dicas agronômicas.

A todos os amigos que fiz durante o Mestrado, principalmente Alison e Juscelino, que sempre me incentivaram nos trabalhos, e através do companheirismo me ajudaram a encarar e resolver todos os desafios.

A todos os professores dos cursos de Computação e Agronomia do mestrado que se dedicaram em transmitir seus conhecimentos, e que, com certeza me ajudaram muito na elaboração desta dissertação.

Muito Obrigado!

RESUMO

A determinação dos teores de matéria orgânica e argila consistem em importantes indicadores da qualidade do solo. Suas quantificações são fundamentais no auxílio do manejo agrônomo podendo fornecer subsídios para as recomendações de corretivos e fertilizantes. Para esta quantificação, a análise de solo se torna uma “ferramenta” indispensável, sendo utilizada cada vez mais, principalmente após a chegada da tecnologia da agricultura de precisão, em que o produtor realiza um número muitas vezes superior de análises visando a identificação da variabilidade do solo da sua propriedade. Porém, as análises laboratoriais trazem alguns inconvenientes, como o tempo necessário para as determinações e a geração de resíduos. Uma opção para realizar as análises de matéria orgânica e argila do solo, de forma rápida e sem geração de resíduos químicos é a espectroscopia de infravermelho visível e próximo (vis-NIRS – *visible and Near Infrared Spectroscopy*). O objetivo deste trabalho foi propor uma metodologia para predição dos teores de matéria orgânica e argila, envolvendo a combinação do uso de Análise de Regressão e Redes Neurais Artificiais com o desenvolvimento de modelos de estimativa destes atributos, utilizando uma base de dados com informações de análises de solo obtidas pelo método convencional e pelo método vis-NIRS. Primeiramente foram selecionadas as bandas espectrais que melhor correlacionavam com as variáveis de resposta (matéria orgânica e argila), por meio de um modelo de regressão multivariada. O grupo que obteve o maior coeficiente de determinação foi utilizado como entrada das Redes Neurais Artificiais visando melhorar a estimativa dos teores de matéria orgânica e argila. Para isto, 111 amostras de solo foram utilizadas para a calibração de modelos de análises de solos, sendo seus espectros obtidos em um espectrofotômetro de infravermelho próximo modelo FOSS XDS. Os resultados dos modelos foram avaliados por meio do coeficiente de determinação (R^2), e pelo grau de significância ao nível de 5%. Correlações de 0,89 e 0,94 foram obtidas na predição do teor de matéria orgânica e argila, respectivamente, com índices altamente significativos ($P < 0,001$), o que indica que a metodologia proposta pode ser utilizada para a predição dos atributos estudados.

Palavras-Chave: Matéria orgânica, argila, análise de regressão, Redes Neurais Artificiais, *near-infrared*.

ABSTRACT

Determining the soil organic matter and clay are important to obtain indicators of soil quality. Such measurements are help the agronomic management providing support for the recommendation of lime and fertilizer. For this quantification, analysis of soil becomes a "tool" indispensable, being increasingly used, especially when associated the the precision farming technology, in which the producer performs a higher number of analyzes aiming to identify soil variability of the property. However, laboratory tests bring some disadvantages, such as the time required for the analyses, and also the generation of waste. An option to perform the analyzes of organic matter and clay soil, quickly and without chemical residues, is by the use of visible to infrared spectroscopy and near (vis-NIRS - Visible and Near Infrared Spectroscopy). The aim of this work was to propose a methodology for predicting the soil organic matter and clay, by combining the use of Regression Analysis and Artificial Neural Networks in order to develop models to estimate these attributes. A database with information about soil analysis obtained by the conventional method and the method vis-NIR was used. The first step was to select the spectral bands that presented a better correlation with the response variables (clay and organic matter) by means of multivariate regression model. In order to improve the estimation of soil organic matter and clay, the group that presented the highest coefficient of determination was used as input of the Artificial Neural Networks. The quantity of 111 soil samples were used for calibration the models of soil analysis, and their spectra were obtained on a spectrophotometer FOSS NIR model XDS. The results were evaluated by the coefficient of determination (R^2), considering the significance level of 5%. Coeficients of 0,89 and 0,94 were obtained in the prediction of organic matter and clay respectively, with indices highly significant ($P < 0,001$), indicating the proposed methodology could be useful to predict the attributes studied.

Keywords: organic matter, clay, regression analysis, Artificial Neural Networks, near-infrared.

LISTA DE SIGLAS

AP	Agricultura de Precisão.
CTC	Capacidade de Troca Catiônica.
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária.
EQM	Erro Quadrático Médio.
MAE	Erro Médio Absoluto (do inglês Mean Absolute Error).
MID	Espectroscopia no Infravermelho Médio (do inglês, mid-infrared Spectroscopy).
MO	Matéria Orgânica.
MOS	Matéria Orgânica do Solo.
MLP	Multilayer Perceptron.
NM	Nanômetros.
NIRS	Espectroscopia no Infravermelho Próximo (do inglês, Near Infrared Spectroscopy).
PCR	Regressão por Componentes Principais (do inglês, Principal Component Regression).
PLS	Mínimos Quadrados Parciais (do inglês, Partial Least Squares).
PLSR	Regressão por Mínimos Quadrados Parciais (do inglês, Partial Least – Squares Regression).
RAE	Erro Relativo Absoluto (do inglês, Relative Absolute Error).
RLS	Regressão Linear Simples.
RLM	Regressão Linear Múltipla.
RNA	Rede Neural Artificial.
R ²	Coeficiente de Determinação.
RMSE	Erro Quadrático Médio.
SMLR	Regressão Linear Múltipla Stepwise (do inglês Stepwise Multiple Linear Regression).
RRSE	Erro Médio da Raiz Quadrada (do inglês, Root Relative Squared Error)
VIS-NIRS	Espectroscopia no Infravermelho Visível e Próximo (do inglês, Visible and Near Infrared Spectroscopy).

LISTA DE FIGURAS

FIGURA 1 - Estrutura de um agregado de solo	15
FIGURA 2 - Ilustração de Reflectância Difusa.....	22
FIGURA 3 - Gráfico dos espectros de reflectância na região do vis-nirs, de quatro amostras de argila utilizadas para a construção dos modelos	23
FIGURA 4 - Representação de escala de vários comprimentos de ondas, com destaque ao comprimento de onda do visível.....	25
FIGURA 5 - Representação de um espectro do solo mostrando as faixas do visível, infravermelho próximo e infravermelho médio.....	26
FIGURA 6 - Ilustração dos mapas krigados e as calibrações utilizadas para prever as amostras resultantes	28
FIGURA 7 - Ilustração básica de uma RNA	32
FIGURA 8 - Representação da sinapse artificial.....	33
FIGURA 9 - Ilustração de um exemplo de arquitetura de uma RNA utilizada para a determinação de MO	34
FIGURA 10 - Estrutura de uma RNA do tipo <i>multilayer feedforward</i>	35
FIGURA 11 – Exemplo de um relatório de saída do programa Weka.....	40
FIGURA 12 - Figura com a localização da fazenda e gleba estudada	42
FIGURA 13 - Mapa de classificação de solos da área	43
FIGURA 14 - Mapa com o grid dos pontos onde foram retiradas as amostras de solo	44
FIGURA 15 - Foto do espectrômetro vis-NIRS, marca FOSS, modelo XDS	45
FIGURA 16 - Foto do preparo da amostra de solo para ser analisada	46
FIGURA 17 - Foto de uma amostra de solo pronta para ser analisada pelo equipamento FOSS modelo XDS.....	46
FIGURA 18 – Exemplo de representação gráfica dos espectros de reflectância	47
FIGURA 19 - Gráfico de correlação com os valores de MO observados versus valores estimados pelo modelo vis-NIRS com dados não transformados.....	53
FIGURA 20 - Gráfico de correlação com os valores de argila observados versus valores estimados pelo modelo vis-NIRS com dados não transformados.....	54

FIGURA 21 - Gráfico de correlação com os valores de MO observados versus valores estimados pelo modelo vis-NIRS com dados transformados	55
FIGURA 22 - Gráfico de correlação com os valores de argila observados versus valores estimados pelo modelo vis-NIRS com dados transformados	56
FIGURA 23 - Representação gráfica da configuração da RNA para MO	57
FIGURA 24 - Representação gráfica da configuração da RNA para argila.....	57
FIGURA 25 - Gráfico de dispersão com os valores de MO observados nas análises laboratoriais versus valores estimados pelo modelo vis-NIRS.....	58
FIGURA 26 - Gráfico de dispersão com os valores de argila observados nas análises laboratoriais versus valores estimados pelo modelo vis-NIRS.....	58
FIGURA 27 – Mapa da área estudada dividida em cinco faixas de argila.....	61
FIGURA 28 - Gráfico de dispersão com os valores de MO observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS	64
FIGURA 29 - Gráfico de dispersão com os valores de argila observados nas análises laboratoriais versus valores estimados pelo modelo vis-NIRS.....	66

LISTA DE TABELAS

TABELA 1 - Escala internacional de classificação das partículas no solo	20
TABELA 2 - Exemplo da saída dos resultados estimados	51
TABELA 3 - Bandas espectrais selecionadas na primeira regressão	53
TABELA 4 - Bandas espectrais selecionadas na segunda regressão com os dados transformados	55
TABELA 5 - Valores com a melhor configuração da RNA para estimar os valores de MO e argila no modelo final.....	56
TABELA 6 - Bandas espectrais selecionadas na primeira regressão	62
TABELA 7 - Bandas espectrais selecionadas na segunda regressão com os dados transformados	63
TABELA 8 - Valores com a melhor configuração para a RNA para estimar os valores de MO e argila.....	63
TABELA 9 - Diferenças entre o valor estimado pelo modelo vis-NIRS para o teor de MO observada nas amostras de solo utilizadas com o erro absoluto (EA) para cada amostra	65
TABELA 10 - Diferenças entre o valor estimado pelo modelo vis-NIRS para o teor de argila observada nas amostras de solo utilizadas com o erro absoluto (EA) para cada amostra.....	67

LISTA DE QUADROS

QUADRO 1 – Primeiro script para geração dos relatórios do programa Weka.....	49
QUADRO 2 – Segundo script para transformar os relatórios do Weka em uma planilha	50
QUADRO 3 – Terceiro script para gerar o modelo final no programa Weka.....	51

SUMÁRIO

1	INTRODUÇÃO	12
2	REVISÃO BIBLIOGRÁFICA.....	14
2.1	SOLOS.....	14
2.2	MATÉRIA ORGÂNICA.....	16
2.3	ARGILA	18
2.4	AMOSTRAGEM E ANÁLISES DE SOLOS	18
2.5	AGRICULTURA DE PRECISÃO	20
2.6	ESPECTROSCOPIA NO INFRAVERMELHO VISÍVEL E PRÓXIMO.....	22
2.7	MO X REFLECTÂNCIA.....	29
2.8	ARGILA X REFLECTÂNCIA	30
2.9	REDES NEURAIS ARTIFICIAIS - RNAs.....	31
2.10	ARQUITETURA MULTILAYER PERCEPTRON (MLP)	36
2.11	SOFTWARE ESTATÍSTICO SAS.....	38
2.12	A FERRAMENTA WEKA	38
3	Materials e métodos	42
3.1	LOCAL DO ESTUDO	42
3.2	COLETA E ANÁLISES DE SOLO	44
3.3	PREPARO E ANÁLISE DOS ATRIBUTOS FÍSICOS E QUÍMICOS DO SOLO	45
3.4	PREPARO E ANÁLISE COM O EQUIPAMENTO NIRS.....	45
3.5	OBTENÇÃO DOS DADOS vis-NIRS.....	47
3.6	PRÉ-PROCESSAMENTO	47
3.7	ANÁLISE ESTATÍSTICA DOS DADOS	48
3.8	CONSTRUÇÃO DA RNA – TREINAMENTO E VALIDAÇÃO.....	49
4	Resultados e Discussões.....	53
4.1	Primeira Etapa – Base de dados Completa.....	53
4.1.1	REGRESSÃO COM DADOS NÃO TRANSFORMADOS.....	53
4.1.2	REGRESSÃO COM DADOS TRANSFORMADOS	55
4.1.3	CONFIGURAÇÃO DA RNA.....	56
4.2	SEGUNDA ETAPA - VALIDAÇÃO DA METODOLOGIA.....	61
4.2.1	REGRESSÃO COM DADOS NÃO TRANSFORMADOS.....	62
4.2.2	REGRESSÃO COM DADOS TRANSFORMADOS	62
4.2.3	CONFIGURAÇÃO DA RNA.....	63
	CONCLUSÕES	68
	TRABALHOS FUTUROS	69
	REFERÊNCIAS	70

1 INTRODUÇÃO

O Brasil é uma potência agropecuária que vem crescendo e se superando a cada novo desafio. Projeções recentes do relatório da Organização para Cooperação e Desenvolvimento Econômico (OCDE)-FAO (2010) indicaram para a agricultura na próxima década, que, no período de 2009 a 2019, a expansão do agronegócio brasileiro deverá ser de 38%, o dobro da média mundial e superior à expansão projetada para outros importantes países produtores de alimentos como Estados Unidos.

Diante deste fato, é fundamental o investimento em pesquisas para novas tecnologias no campo tanto em entidades públicas como privadas, transferindo conhecimento do pequeno ao grande produtor.

Neste contexto, a região dos Campos Gerais no centro sul do Paraná se caracteriza pela alta tecnologia aplicada na atividade agropecuária, com médias de produção comparáveis as melhores do mundo.

Sendo o estudo do solo essencial para o incremento das produtividades, e esta região possuir como características uma alta variabilidade de solo, a busca para interpretar, compreender e avaliar a fertilidade nestes diferentes tipos de solos, que esta, diretamente ligada às propriedades físico-químicas do mesmo, pesquisadores procuram aperfeiçoar e/ou criar novas técnicas e procedimentos.

Um dos segmentos do setor agrícola que visa gerar conhecimentos úteis para o entendimento da variabilidade do solo e que vem crescendo rapidamente é a Agricultura de Precisão (AP).

A AP é uma das técnicas que podem auxiliar o produtor a entender melhor a variabilidade de solos e consequentemente as diferenças de produtividade. A geração de mapas de fertilidade de solo é um dos produtos obtidos por meio da AP e para que isso seja realizado é necessária a coleta de um grande volume de amostras de solo em uma mesma área, gerando um aumento de custo para o produtor além do tempo de espera pelos resultados.

O primeiro passo para um bom resultado na lavoura é a realização das análises químicas, e a partir da interpretação dos números obtidos, é possível ter indicativos sobre o que e quanto aplicar de fertilizantes e corretivos, não desperdiçando dinheiro, tempo e melhorando a qualidade ambiental. Estas recomendações são a base para a correção do solo e sua fertilidade, fundamentais para o plantio das culturas e de extrema importância para o fracasso ou sucesso da lavoura.

Os métodos utilizados para as análises químicas e físicas de rotina em laboratórios são trabalhosas, demoradas, caras e necessitam de descarte apropriado de reagentes, caso contrário, serão poluentes ao meio ambiente (VISCARRA ROSSEL e MCBRATNEY, 1998).

Um dos grandes desafios da pesquisa é o desenvolvimento de métodos de análises que não necessitem de tratamento prévio das amostras, sejam rápidos, não destrutivos e com bons índices de respostas. Uma alternativa que tem apresentado bons resultados e em poucos segundos, se obtém resultados de múltiplos atributos do solo em amostras preparadas ou não, é a espectroscopia de infravermelho visível e próximo (vis-NIRS). Esta metodologia fundamenta-se na medida da intensidade de absorção de radiação eletromagnética na região do infravermelho próximo.

A análise utilizando vis-NIRS (visible and Near-infrared spectroscopy) ou espectroscopia de reflectância na região do visível e no infravermelho próximo é uma integração da espectroscopia, estatística e computação. O princípio mecânico é o de iluminar uma amostra com luz de comprimento de onda específico e conhecido da região do espectro eletromagnético. A absorção de luz é medida por diferenças entre a quantidade de luz emitida e a quantidade de luz refletida pela amostra, relação da qual pode-se prever a sua composição química, desde que as leituras obtidas possam ser instantâneas, efetivamente comparadas e ajustadas na matriz de um banco de dados armazenado que calibra o software de logística do equipamento (BORGES NETO, 2005). Além dos modelos estatísticos que utilizam métodos multivariados que estão sendo largamente aplicados para encontrar as características ou propriedades do solo por meio da análise via vis-NIRS, pode-se aplicar também técnicas de uma área chamada Mineração de Dados, a qual utiliza Inteligência Artificial como, por exemplo, Redes Neurais Artificiais (RNAs).

O objetivo deste trabalho foi propor uma metodologia híbrida, envolvendo a combinação do uso de Análise de Regressão e RNAs, para desenvolvimento de modelos de estimativa da matéria orgânica (MO) e argila do solo utilizando uma base de dados com informações obtidas pelo método convencional de análise e pelo método vis-NIRS.

2 REVISÃO BIBLIOGRÁFICA

2.1 SOLOS

Solo é a parte superficial intemperizada não consolidada da crosta terrestre, contendo matéria orgânica e seres vivos. Nele se desenvolvem os vegetais, que obtém do solo, por meio das raízes, a água e os nutrientes necessários para seu desenvolvimento (Raij, 1991). É composto por partículas sólidas e espaços porosos ocupados por água e ar.

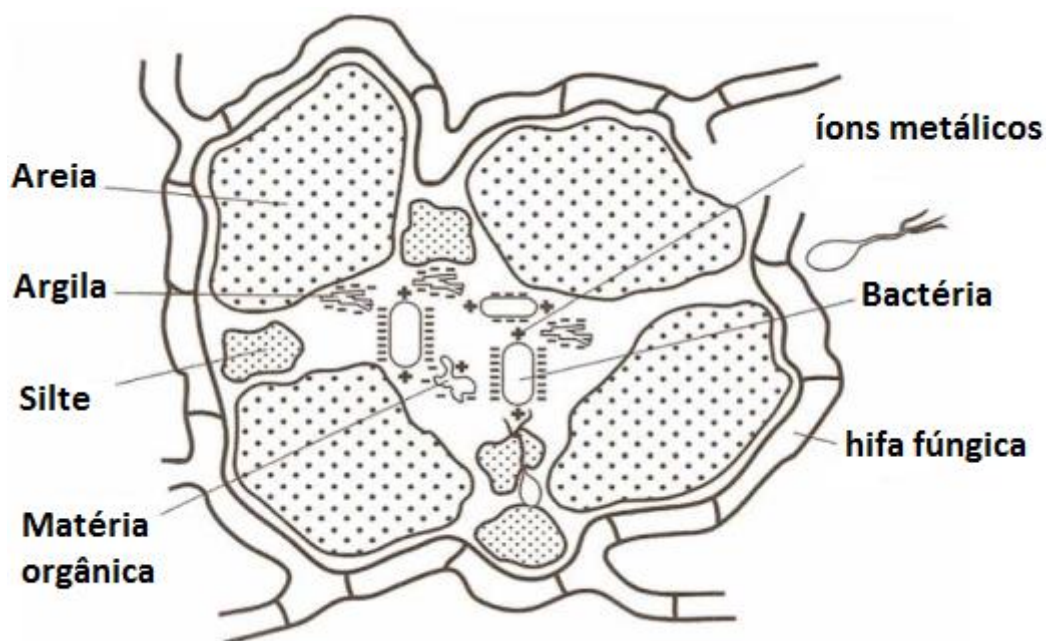
Segundo Essington (2004), o solo consiste no maior reservatório de elementos e substâncias que supre as necessidades nutricionais de plantas e animais.

Torstensson (1998) definiu solo como uma estrutura complexa formada pela influência de fatores geológicos, topográficos, climáticos e temporais, sendo a parte mais importante da geosfera e que provém a maior parte dos nossos alimentos. Além disso, ele funciona como uma espécie de filtro ambiental para a purificação da água e do ar.

O solo funciona como uma máquina recicladora, recebendo restos de seres mortos e resíduos de atividades urbanas e agrícolas. Os solos, de uma maneira geral são formados por 45% de substâncias minerais, 25% de água, 25% de ar e 5% de matéria orgânica (MO). Em situações naturais, a proporção de ar e água é sujeita à alterações, dependendo do tipo de solo, região e seus períodos de chuvas. A textura, ou granulometria do solo também varia dependendo diretamente do tipo e da degradação da rocha matriz, desde grandes partes de rochas até o pó fino.

No caso dos solos muito arenosos ou várzeas, a granulometria pode influir diretamente na produtividade, considerando que estes solos possuem baixos teores de MO e consequentemente baixa capacidade de retenção de nutrientes e água (MOREIRA e SIQUEIRA, 2006). Uma amostra de uma partícula de um agregado de solo esta representado na Figura 1.

Figura 1 - Estrutura de um agregado de solo com destaque ao papel agregador da MO e dos microorganismos em interação com partículas de areia, argila e silte.



Fonte: Adaptado de Moreira e Siqueira (2006).

O território brasileiro é formado por cerca de 70% de solos ácidos, com baixo conteúdo de bases trocáveis. Isto acontece devido, principalmente, à ação das chuvas que ocasionam lixiviação de cátions básicos. Uma das formas de correção ou neutralização da acidez do solo é feita por meio da aplicação de calcário. A calagem contribui de forma relevante para o aumento da produtividade e da renda agrícola, em vista dos benefícios que traz as propriedades químicas, físicas e biológicas dos solos (Quaggio, 2000).

No Brasil, os solos estão localizados, em sua maioria (73%), na região do bioma Cerrado, em uma altitude entre 300 e 900 m, sendo Latossolo o tipo predominante (46%). Esses solos são profundos, bem estruturados, com alta porosidade, considerados de grande potencial agrícola para produção de culturas anuais e perenes, e também para pastagens (FERREIRA, et al., 1999; RESCK et al., 2008; EBERHARDT et al., 2008). A textura dos Latossolos brasileiros é, principalmente, argilosa, predominando a caulinita, gibbsita, goethita e hematita, que são minerais com baixa capacidade de troca de cátions (CTC) (FERREIRA et al., 1999). Solos com argila de alta atividade apresentam CTC alta e podem reter grandes quantidades de cátions e solos arenosos apresentam baixa CTC e são mais suscetíveis às perdas de nutrientes por lixiviação.

Mais especificamente na região Centro-Sul do Paraná, na região dos Campos Gerais, os solos são originados principalmente de rochas eruptivas da formação Serra Geral, com predominância de basalto, sendo assim mais argilosos e com elevados teores de ferro (Embrapa, 1984).

De acordo com Silva et al. (2003), a textura e o teor de MO são características importantes de serem conhecidas no manejo dos solos, e também no planejamento para áreas que utilizam irrigação, porque têm influência direta na taxa de infiltração de água, na aeração, na capacidade de retenção de água, na nutrição, e também na aderência ou força de coesão das partículas do solo. A textura com os teores de areia, silte e argila no solo afetam diretamente o ponto de aderência aos implementos de preparo do solo e de semeadura, facilitando ou dificultando o trabalho das máquinas. A textura do solo tem uma estreita relação com a retenção de água e a CTC (RESENDE et al., 1997).

Diferentes usos da terra podem causar efeitos diferentes sobre as propriedades do solo, e algumas propriedades químicas ou físicas podem ser mais sensíveis do que outras (TRANGMAR et al., 1985). Além de diferentes usos pode-se dizer também que o solo é um sistema heterogêneo cujos processos e mecanismos são complexos e difíceis de serem totalmente compreendidos (VISCARRA ROSSEL, 2006). Sendo heterogêneo e possuindo grande variabilidade espacial, a sua amostragem é uma operação que visa quantificar a variação espacial, a qual depende diretamente do manejo adotado pelo produtor como: culturas instaladas, sistema de cultivo, aplicação de corretivos, fertilizantes e utilização de produtos químicos.

2.2 MATÉRIA ORGÂNICA

Dentre os constituintes do solo, um dos mais estudado é a MO, sendo um fator central na sustentabilidade dos sistemas agrícolas. Práticas que favoreçam a conservação da MO melhoram as propriedades do solo e ajudam a reduzir o risco de erosão. Nos chamados sistemas de manejo sustentável, microrganismos benéficos são incorporados à camada superficial do solo juntamente com restos de colheita e outros resíduos orgânicos, reduzindo a aplicação de pesticidas e fertilizantes, levando a um aumento na estabilidade e na conservação do solo (POPPI e SENA, 1999).

A MO, também denominada fração orgânica, ou húmus, tem seu carbono originado no processo da fotossíntese, realizados por vegetais clorofilados, que transformam gás carbônico, oxigênio, hidrogênio e nutrientes minerais em compostos orgânicos (RAIJ, 2001).

É um dos componentes do solo que afeta favoravelmente a produtividade e merece cuidados, porque sua perda pode ocorrer por meio do uso inadequado do solo como queimadas, monoculturas e revolvimentos, causando reflexos na CTC e na disponibilidade de nutrientes.

Anjos et al. (2008) relatam que o teor de MO do solo pode afetar: estabilidade dos agregados, a disponibilidade de nutrientes e a retenção de água. Solos com textura média a argilosa, moderadamente drenados e com alta atividade microbiana, geralmente apresentam grande quantidade de agregados estáveis responsáveis por proteger fisicamente a MO.

Sob vegetação natural, os solos da região dos Campos Gerais apresentam teores de MO estável, variando de médio ($25\text{-}50\text{ g/dm}^3$) a alto ($>50\text{ g/dm}^3$), decorrente de sua textura argilosa e de temperaturas amenas (Embrapa, 1984). Nessa região há predomínio de caulinita e óxidos de ferro e de alumínio na fração argila. Assim, a MO é certamente o componente responsável por uma grande fração da CTC do solo (FONTOURA et al., 2009). Sabendo-se destas características, a preservação da MO se tornou uma das principais preocupações dos pesquisadores. Se a fertilidade é naturalmente muito baixa ou diminui abaixo de um nível crítico após um longo período de utilização agrícola sem reposição de nutrientes ou como resultado da erosão, sua reconstituição pode ser uma condição prévia para a agricultura produtiva. Uma melhor compreensão das interações entre as culturas e os solos ajuda a manter as expectativas em um nível realista e reconhecer o que o ecossistema agrícola pode ou não alcançar (DU e ZHOU, 2009).

Para a determinação do teor de MO do solo o procedimento mais utilizado é o método proposto por Walkley-Black (WALINGA, 1992), o qual utiliza o dicromato ($\text{Cr}_2\text{O}_7^{2-}$), causador de contaminação ambiental por cromo. Seguindo esta afirmação, busca-se novas alternativas ao método tradicional, sendo a utilização da espectroscopia de infravermelho próximo uma opção interessante para estimar o teor de MO no solo, pois a necessidade de reagentes é descartada, reduzindo tempo de análise e eliminando a possibilidade de contaminação do ambiente por cromo.

2.3 ARGILA

Assim como a MO, outro importante constituinte do solo é o teor de argila. A argila possui cargas elétricas negativas responsáveis pela CTC, que é um dos requerimentos para recomendação de doses de fertilizantes e corretivos e também possui uma estreita relação com a retenção de água do solo.

Minerais de argila diferentes possuem propriedades diferentes, por exemplo, alguns são capazes de reter a água em suas malhas, dando-lhes o seu comportamento inchar-encolher, enquanto outros são importantes como fonte de potássio (ANDRIST-RANGEL et al., 2006). Partículas de argila são caracterizadas por superfícies carregadas negativamente e alguns minerais de argila têm superfícies mais carregadas negativamente do que outros.

Estas superfícies carregadas são importantes em termos de física e química do solo na regulação dos processos de agregação e na CTC do solo, que afeta a liberação e/ou retenção de nutrientes (HILLEL e HILLEL, 1998).

Uma das variáveis úteis na previsão de propriedades físicas dos solos é sem dúvida o teor de argila, pois se relaciona diretamente com a retenção de umidade, resistência do solo e muitos outros processos físicos e químicos (EMBRAPA, 2003).

O método da pipeta (EMBRAPA, 1997) é um dos mais utilizados nos laboratórios para a determinação do teor de argila.

2.4 AMOSTRAGEM E ANÁLISES DE SOLOS

A amostragem de solo representa a fase mais crítica do processo para recomendação de adubação e correção da acidez do solo por meio da aplicação de calcário para a maioria das culturas (PAULETTI, 2004). Watanabe et al. (2002) comentam que antes da semeadura, para o agricultor atingir uma boa produtividade, a análise de solo é uma “ferramenta” essencial, porque somente os dados obtidos a campo pela observação visual não são suficientes para se determinar possíveis problemas relacionados à nutrição das plantas.

A retirada de amostras de solo nada mais é que uma coleta de porção de terra, sendo o primeiro procedimento solicitado pelo assistente técnico ao produtor para que

se constate como está a qualidade química e física do solo, tendo assim, embasamento para as recomendações de calagem e fertilizantes apropriadas sem desperdício, racionalizando tempo e recursos financeiros.

Normalmente os agricultores têm feito a amostragem do solo de uma dada área e uniformizado as sub-amostras em uma única, que passa a representar as características de fertilidade daquela área, assumida como uniforme. Com base nessa interpretação fazem a aplicação de quantidades uniformes de insumos na propriedade. No entanto as propriedades do solo e as infestações de doenças e pragas variam de local para local dentro de uma mesma lavoura (WIEDA e BORGELT, 1993; SCHUELLER, 1992). Essa variabilidade gera a necessidade de uma melhor compreensão do solo levando ao uso de técnicas mais quantitativas para sua modelagem e mapeamento (WEBSTER e BUTLER, 1976).

A quantidade de amostras de solo varia não só de campo para campo, como também depende dos parâmetros da fertilidade do local (CAHN et al., 1994; CAMBARDELLA et al., 1999). Larson et al. (1994), propuseram um conjunto mínimo de variáveis químicas, físicas e biológicas, que, acompanhadas ao longo do tempo, são capazes de detectar as alterações da qualidade do solo em função do manejo. O carbono orgânico total (COT) e a MO do solo encontram-se entre essas variáveis.

A MO tem sido representada em porcentagem (%). Como nos laboratórios de análise de solo são feitas medidas volumétricas de solo, isso na realidade significa $\text{g}/100 \text{ cm}^3$ de terra. Alguns laboratórios ainda representam os resultados em carbono, mas é preferível usar MO, lembrando que $\text{MO} = 1,724 \times \text{C}$ (RAIJ, 1991).

Os parâmetros físicos e químicos, importantes para o crescimento de plantas, são controlados pela textura, que determina a extensão da superfície específica. A textura expressa a distribuição percentual das partículas primárias e é determinada pela análise granulométrica (CAMARGO et al., 2009). Esta análise visa classificar as partículas do solo pelos respectivos tamanhos e suas frações correspondentes, de acordo com a Tabela 1:

Tabela 1- Escala internacional de classificação das partículas no solo.

Fração Granulométrica	Limites dos diâmetros das partículas(mm)
Argila	< 0,002
Silte	0,002-0,02
Areia fina	0,02-0,2
Areia grossa	0,2-2,0
Cascalho	2 – 20
Pedra	> 20

Fonte: Raij et.al, (2001).

De acordo com as frações da Tabela 1 percebe-se que nestes intervalos existem diferentes granulometrias, e conseqüentemente, diferentes concentrações e tamanho de partículas. Estes dois itens formam a estrutura do solo que é composta por um arranjo físico e a agregação dessas partículas. A textura e a estrutura são responsáveis pela quantidade e tamanho dos espaços porosos no solo, que são ocupados pela água e pelo ar. No caso de um solo ideal para cultivo, sem problema com excesso de umidade, o ar ocupa os poros maiores que 5 mm e a água os poros menores que 5 mm. Dentre as frações, a argila é a mais ativa quimicamente, sendo constituída por minerais secundários, como a caulinita, montmorilonita, e óxidos de alumínio e ferro (RAIJ, 2001).

2.5 AGRICULTURA DE PRECISÃO

Odlare et al. (2005), afirmam que o conhecimento da variação espacial do solo é importante em várias disciplinas, como em pesquisa de testes de campo agrícola e agricultura de precisão. O objetivo da AP é gerenciar os insumos agrícolas com respeito a variação espacial e produzir mapas de solos com base em um grande número de análises químicas, físicas e biológicas, sendo um método demorado e caro.

A AP é uma tecnologia ainda nova para a maioria dos produtores no Brasil, mas que vem ganhando espaço pelo fato de estar com custo dos equipamentos mais acessíveis e por existirem mais técnicos com formação especialista para a interpretação dos resultados obtidos. É um método de administração cuidadosa e datalhada do solo e da cultura para adequar as diferentes condições encontradas em cada parte da lavoura, tendo em vista a desuniformidade intrínseca dos solos (SCHUELLER, 1992; WEIDA E BORGELT, 1993).

Antes do advento da AP, o agricultor tratava o solo como único e, assim, tinha como banco de dados poucas análises químicas que serviam para toda a área, como também o desconhecimento de qual parte daquela área estava produzindo mais ou menos. O primeiro produto gerado pela AP deveria ser o mapa de colheita e, por meio deste, os produtores poderiam visualizar que uma área tem diferentes produtividades, por apresentar diferentes texturas e níveis de nutrientes.

Com a chegada da AP e sua forma de trabalhar com a produção de mapas de fertilidade e produtividade, a geração de dados que era reduzida, agora se torna um grande banco de dados, com muitas informações e consequentemente muitas dificuldades na sua interpretação. O estudo de novas técnicas de análise de solo é fundamental para apoiar o desenvolvimento da AP, principalmente para agilizar e baratear este processo.

Pierce e Nowak (1999) utilizam a seguinte definição: “Agricultura de Precisão é a aplicação de princípios e tecnologias para manejar a variabilidade espacial e temporal, associada com todos os aspectos da produção agrícola, com o objetivo de aumentar a produtividade na agricultura e a qualidade ambiental”.

A AP é um processo de gerenciamento agrícola que busca informações mais precisas e se completa com decisões mais refinadas. Conforme Canzian et al. (1999), algumas áreas podem ser uniformes, mas outras apresentam variações no tipo de solo, fertilidade e outros fatores que afetam a produção agrícola. Se a variabilidade do campo puder ser medida e registrada, estas informações poderão ser usadas para otimizar as aplicações em cada ponto. A obtenção de dados por amostras para medir variabilidades em campo é geralmente bastante precisa, no entanto requer tempo e tem um custo significativo em termos de análises (SCHUELLER, 1992). De acordo com Robert (2002), os principais custos na AP envolvem: aquisição e operação de máquinas e equipamentos (GPS, sensores – monitor de colheita, computadores e programas) e amostragem de solos e análises.

A obtenção de um melhor manejo é um dos principais objetivos da AP, logo, tem que ser realizado adequadamente. Para isso o produtor necessariamente precisa fazer muitas análises de solo. Viscarra Rossel e McBratney (1998) observaram que como a AP esta relacionada diretamente a realização de uma maior quantidade de amostras de solo o volume de descartes dos reagentes químicos também aumenta, indo contra a uma tendência mundial de práticas menos poluentes. Os mesmos autores em 2001 comentam que os mapas gerados pela AP com atributos do solo dependem de amostragem adequada e análises laboratoriais para derivar modelos da variação espacial de atributos específicos do solo que, podem ser generalizados em toda a região nos locais não amostrados.

Como a demanda pela quantificação dos dados aumenta consideravelmente utilizando a tecnologia, de AP, Kitchen (2008), afirma que, o uso de computadores e sensores para decisões em tempo-real em sistemas agrícolas está crescendo rapidamente. Quando integrado com o conhecimento agrônômico, o valor desta tecnologia pode ser melhor percebido, resultando em um processo contínuo de avaliação, interpretação e operações direcionadas.

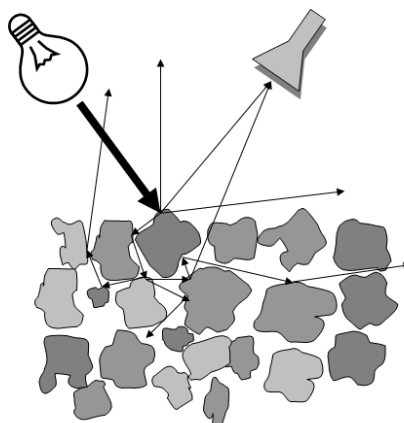
Uma das técnicas que pode auxiliar o setor agrícola é a utilização da espectroscopia no infravermelho vis-NIRS. Este método está potencialmente vinculado ao desenvolvimento progressivo dos sistemas de sensores próximos ou de varredura. Esses sistemas podem superar problemas atuais de alto custo, esforço, tempo e imprecisão de amostragem, aumentando eficazmente a precisão da representação da variabilidade espacial das propriedades medidas (VISCARRA ROSSEL e MCBRATNEY, 1998).

2.6 ESPECTROSCOPIA NO INFRAVERMELHO VISÍVEL E PRÓXIMO

O termo *espectroscopia* em uma definição simples consiste no estudo da radiação eletromagnética emitida ou absorvida por um corpo. Esta técnica é largamente empregada nas áreas de química, física, engenharias, astronomia, entre outras (SALIBA, 1999).

Viscarra Rossel et al. (2010), dizem que espectros são também altamente influenciados pelas propriedades estruturais da amostra, o que provoca efeitos não lineares de espalhamento de luz, daí o nome de *espectroscopia de reflectância difusa*. A Figura 2 exemplifica o espalhamento da luz.

Figura 2 - Ilustração de reflectância difusa. A luz recebida da esquerda foi espalhada. Do lado direito apenas uma fração da luz refletida atinge a sonda, o resto é absorvido ou perdido.



Fonte : Adaptado de Viscarra Rossel et. al. (2010).

O equipamento NIRS é uma integração da espectroscopia, estatística e computação. Seu princípio mecânico é o de iluminar uma amostra com luz de comprimento de onda específico e conhecido da região do espectro eletromagnético. A absorção de luz então é medida por diferenças entre a quantidade de luz emitida pelo NIRS e a quantidade de luz refletida pela amostra, relação pela qual pode-se prever a sua composição química, desde que as leituras obtidas possam ser instantâneas, efetivamente comparadas e ajustadas na matriz de um banco de dados armazenado que calibra o software do equipamento.

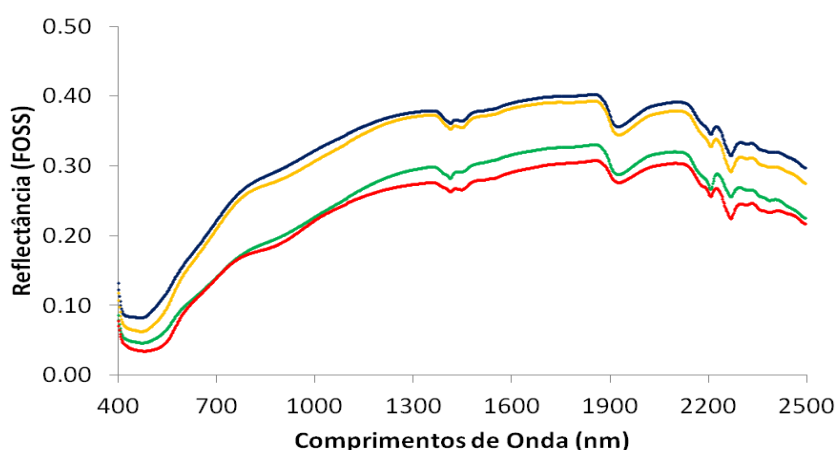
O equipamento utiliza energia luminosa de luz branca composta de todas as cores do espectro do arco-íris. Quando essa luz incide sobre um objeto, certas cores são absorvidas e outras refletidas ou transmitidas. A luz refletida pode ser captada pelos olhos. Por exemplo, quando a luz incide sobre uma folha verde são absorvidos o vermelho e o azul. As cores refletidas são interpretadas pelos nossos sentidos visuais como verde.

A energia em que qualquer pico de espectro de absorção que aparece corresponde à frequência de uma vibração de uma parte de uma molécula de amostra, o que torna espectroscopia no infravermelho um método alternativo na avaliação do solo (DEMATTE et al., 2004) e que segundo Almeida (2005), é um método promissor para a estimativa rápida de uma série de atributos do solo.

Como os solos apresentam diferentes constituintes, os mesmos podem ser identificados e, em certos casos, quantificados pela análise de sua resposta espectral. Os principais constituintes dos solos que influenciam seu comportamento espectral são a matéria orgânica, óxidos de ferro, argilominerais, além da distribuição granulométrica e umidade (DALMOLIN et al., 2005).

A medida de reflectância pode ser expressa em curvas espectrais como mostra a Figura 3, representando quatro amostras de argila.

Figura 3 - Exemplo dos espectros de reflectância na região do vis-NIRS, de quatro amostras de argila.



Segundo Dalmolin et al. (2005), a argila é um componente espectralmente caracterizado na região do infravermelho próximo e como cada solo apresenta uma diferente constituição, a respectiva curva espectral terá diferentes bandas de absorção. Os mesmos autores citam que a distribuição do tamanho de partícula no solo e a presença de diferentes argilominerais influem na resposta espectral dos solos. De uma maneira geral, solos de textura arenosa tendem a ter maior reflectância, devido à sua constituição mineralógica (rica em quartzo) e ao fato de geralmente apresentarem baixos teores de matéria orgânica, óxidos de ferro e menores teores de água.

O espectro de um determinado material obtido com radiação infravermelho é o resultado da absorção de energia (luz) por moléculas orgânicas, particularmente aquelas que possuem grande número de ligações do tipo carbono-nitrogênio, nitrogênio-hidrogênio e oxigênio- hidrogênio (AMORIM, 1996). Os avanços no conhecimento da relação entre a reflectância espectral e características do solo fornecem uma ferramenta para prever várias características físicas e químicas do solo de maneira rápida, segura, não invasiva e quando integrada com o conhecimento agrônômico, o valor desta tecnologia pode ser melhor percebido, resultando em um processo contínuo de avaliação, interpretação e operações precisas (KITCHEN, 2008).

Técnicas de espectroscopia do infravermelho próximo são altamente sensíveis a ambas as fases orgânica e inorgânica do solo, tornando a sua utilização nas ciências agrárias e ambientais particularmente relevante. Elas são rápidas, em alguns países tem menor custo, eficientes e não necessitam de tempo e custo no pré-processamento da amostra. Outra característica importante é a não utilização de extratores químicos, evitando assim agressão e toxicidade ao ambiente. Raji (1991) demonstrou sua preocupação com o sistema destrutivo dos recursos naturais para o avanço da agricultura e colocava que: uma agricultura moderna exige o uso de corretivos e fertilizantes em quantidades adequadas, de forma a atender a critérios racionais, que permitam conciliar o resultado econômico positivo com a preservação dos recursos naturais do solo e do meio ambiente.

A espectroscopia do visível e infravermelho pode, por vezes, ser mais simples do que a análise convencional do solo e também, em determinadas ocasiões, apresentam a vantagem do bom potencial e curto prazo de adaptabilidade destas técnicas para o uso em campo em tempo-real (VISCARRA ROSSEL et. al, 2006).

Obukhov e Orlov (1964) demonstraram as diferenças existentes entre grupos de solos em termos de reflectância difusa. Estas descobertas, porém, não foram exploradas até recentemente, quando Shepherd e Walsh (2002) demonstraram o uso da reflectância difusa na

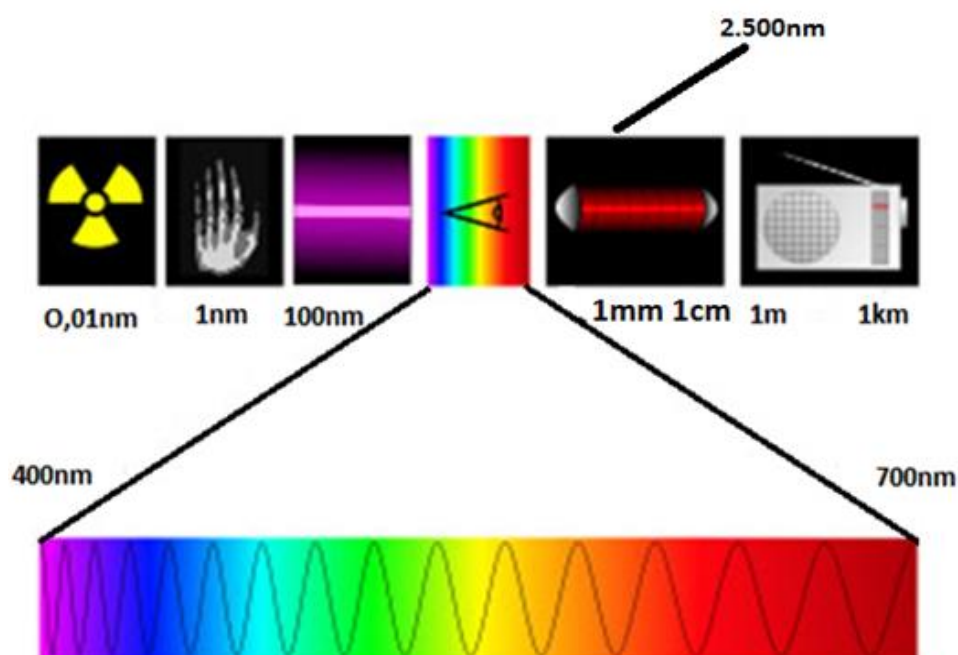
faixa do visível e do infravermelho próximo na estimativa rápida de alguns parâmetros de fertilidade na superfície do solo no leste da África. Fidêncio et al. (2002) demonstraram o uso da reflectância infravermelho próximo, na predição de carbono orgânico em solos brasileiros e também conduziram um estudo semelhante ao de Obukhov e Orlov (1964). Ainda, Stenberg et al. (2010) e Viscarra Rossel et al. (2006), consideram que a MO e o teor de argila, associados ao teor de nitrogênio total, são promissores no que se refere a avaliação de atributos do solo. Os autores afirmam que existe uma ação global no sentido do desenvolvimento de mais metodologias para análise de solo, pois há uma grande demanda para grandes quantidades de dados de boa qualidade a serem usados na agricultura. Comentam também que em 2009 propuseram um projeto para a criação de uma Biblioteca Espectral do Solo Global. Este projeto tem como objetivo desenvolver uma rede global de colaboração para espectroscopia de solo visando o aprofundamento da investigação e o incentivo de sua adoção em ciência do solo.

A necessidade de informação analítica aumenta em importância à medida em que se avança em direção a práticas mais respeitadoras do ambiente e sustentáveis no manejo intensivo de sistemas agrícolas (MORGAN e ESS, 1997). E de acordo com Mello (1999), a técnica vis-NIRS tem sido explorada como uma abordagem para análise rápida e “limpa” na área agrícola.

Os sensores laboratoriais para estudo do comportamento espectral de solo mais comuns utilizam as faixas espectrais obtidas em nanômetros as quais são: entre 400 - 700 nm (visível - VIS), 700 - 2.500 nm (infravermelho próximo - NIRS) e 2.500 - 25.000 nm (infravermelho médio - MIR). Existem outras faixas de comprimento de ondas exploradas por outras tecnologias. A Figura 4 representa de uma forma simples essas diferentes aplicações.

Absorção na região do visível ocorre principalmente devido a excitações eletrônicas que estão, sobretudo, relacionadas com os minerais que contêm ferro e com a MO, devido à cor escura dos ácidos húmicos (STENBERG et al., 2010). Com referência à água, esta possui fortes características de absorção no espectro, mais perceptível entre 1400 e 1900 nm e por fim, em toda a região do NIRS os minerais de argila podem ser absorvidos. (VISCARRA ROSSEL e BEHRENS 2010).

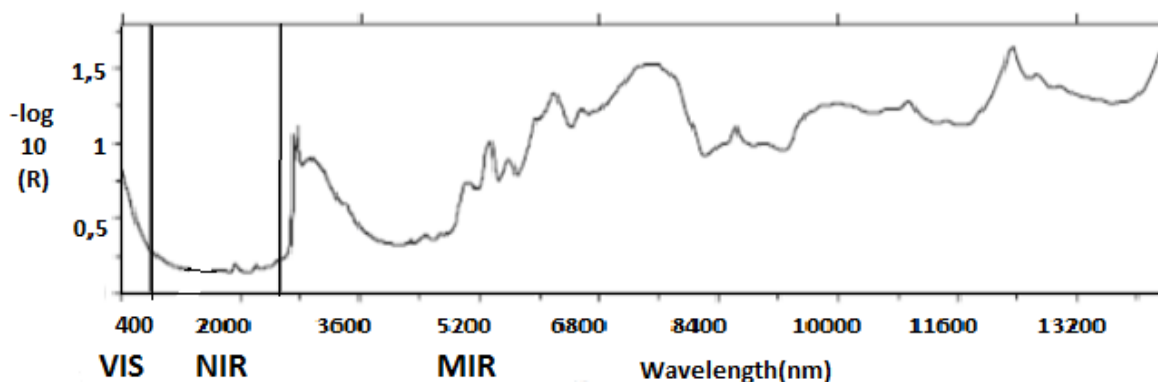
Figura 4 - Representação de escala de vários comprimentos de ondas, com destaque ao comprimento de onda do visível.



Fonte : <http://pt.wikipedia.org/wiki/Ficheiro:Spectre.svg>

A Figura 5 ilustra uma representação de um espectro de solo mostrando as faixas do visível (VIS), infravermelho próximo (NIRS) e infravermelho médio (MID).

Figura 5 - Representação de um espectro do solo mostrando as faixas do visível (VIS), infravermelho próximo (NIRS) e infravermelho médio (MID).



Fonte: Adaptado de Viscarra Rossel et al. (2006).

A análise realizada por espectroscopia de reflectância gera um elevado número de informações. Uma das opções para reduzir essa quantidade de dados é a aplicação de funções estatísticas de regressão, aumentando a eficiência nas avaliações qualitativas e quantitativas do comportamento espectral dos solos e obtendo modelos de calibração para a predição de atributos dos solos (TERRA, 2011).

As variáveis do solo analisadas por meio do vis-NIRS estão sendo calibradas utilizando métodos estatísticos como regressão linear múltipla, regressão polinomial, regressão de componentes principais, regressão linear múltipla stepwise (BEN-DOR e BANIN, 1995a; CHANG et al., 2001; DANIEL et al., 2004; ISLAM et al., 2003; UDELHOVEN et al., 2003) e técnicas de mineração de dados (BROWN et al., 2006; Shepherd e Walsh, 2002).

Extrair padrões matemáticos de espectros de absorção do solo e correlacioná-los com as propriedades do mesmo é uma tarefa complexa requerendo o uso de calibrações multivariadas (MARTENS e NAES, 1989).

Segundo Du e Zhou (2009), a aplicação da técnica vis-NIRS é em grande parte dependente de pré-tratamento e aplicação de métodos de calibração multivariada devido às fortes interferências nos espectros. Soluções matemáticas como Mínimos Quadrados Parciais (PLS) e RNAs são dois instrumentos que podem ser amplamente utilizados para executar as análises na predição de propriedades do solo.

Ingleby e Crowe (1999) utilizaram modelos de regressão linear para prever teor de carbono orgânico do solo em cinco campos experimentais em uma região do Canadá chamada Saskatchewan. Relataram que o melhor modelo para cada campo variava de tamanho e comprimentos de onda de reflexão, sugerindo que, o desenvolvimento de modelos específicos poderiam ser necessários. Eles também relataram que os métodos de modelos alternativos de desenvolvimento, como PLS, PCR (regressão por componentes principais) e RNA precisariam ser amplamente aplicados. Os mesmos autores em 2001 fizeram outro estudo utilizando RNA em comparação com modelos de regressão linear múltipla (MLR) para predição de MO. De acordo com a pesquisa o desempenho da RNA foi muitas vezes altamente dependente da quantidade de tempo de treinamento.

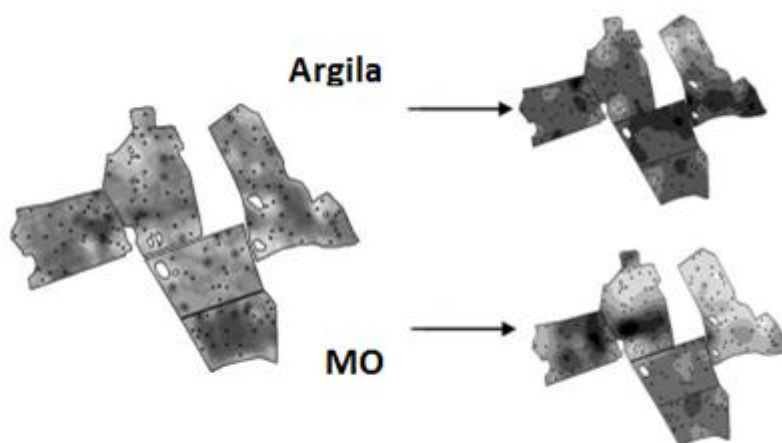
As RNAs têm sido utilizadas em uma ampla variedade de aplicações na agricultura, como na previsão de rendimento de milho (LIU e GOERING, 1999), estimativa do teor de açúcar no suco de frutas (CHEN et al., 1999) e modelagem de solo. Sudduth (1991) desenvolveu uma RNA para a previsão da carbono orgânico do solo usando dados do infravermelho próximo com reflectância entre 800 a 2600 nm como entrada. Os resultados da RNA foram decepcionantes quando comparados com técnicas de regressão. Entretanto o autor afirmou que a otimização da rede tinha sido limitada por restrições de tempo e que melhorias em software e hardware trariam maior precisão na utilização da RNA.

Morimoto et. al (2004) construíram um sensor para ser utilizado em tempo real e o testaram em sete campos diferentes no Japão, com plantações de arroz de sequeiro, fazendo

uso de um espectrofotômetro de infravermelho próximo, para prever MO, nitrogênio total (NT), pH e teor de umidade. O sensor coletou uma absorvância no espectro de solo (de 500-1650 nm com intervalo de 7 nm) e os modelos de predição foram baseados em um conjunto de 1300 pontos, utilizando uma arquitetura de RNA com aprendizado não-supervisionado e algoritmo Backpropagation. O desempenho da RNA, após treinamento e validação, para umidade, pH, MO e nitrogênio total indicaram que o modelo avaliado foi um sucesso, atingindo os coeficientes de determinação (R^2) de 0,96, 0,95, 0,92 e 0,75 respectivamente.

Outro estudo com dados publicados de testes em fazendas na Suécia sugere que a racionalidade pode ser adquirida com exploração em escala de calibrações construídas em apenas 25 amostras, em cerca de 100 ha, prevendo argila e MO com elevada precisão (Figura 6). Neste caso, Wetterlind et al. (2008b) geraram mapas para predição de MO e argila utilizando o método de interpolação por Krigagem (LANDIM, 2003), o que requer dados espacialmente dependentes e uma quantidade maior de amostras do que normalmente utilizado para pesquisas agrícolas. Os autores fizeram o experimento extrapolando as análises de solos feitas utilizando a técnica vis-NIRS que levou a criação de mapas, concluindo que os resultados foram satisfatórios, com pequenos erros e que esta técnica é promissora gerando um custo-benefício razoável e diminuindo o número de amostras para os fazendeiros que utilizarem o mapeamento de solo para o teor de MO e argila com calibrações vis-NIRS.

Figura 6 - Figura ilustrativa dos mapas krigados que estão mostrados no lado direito.



Fonte: Adaptado de Wetterlind et al. (2008b).

Outra aplicação atribuída ao espectro vem sendo os estudos sobre a caracterização direta da condição do solo. Como exemplo, pode-se citar Awiti et al. (2008), os quais

utilizaram espectroscopia de infravermelho próximo como ferramenta para diagnosticar o potencial de uso do solo para agricultura e manejo. Neste caso, a técnica era capaz de indicar três condições básicas de uso: boa, razoável e insuficiente. Viscarra Rossel et al. (2010), por sua vez, propuseram uma técnica versátil para identificação de níveis de fertilidade indicando se determinada área agricultável era: muito fértil, fértil ou pobre, utilizando o método estatístico de árvores de regressão e curvas espectrais sem a necessidade de amostras convencionais.

Em um trabalho desenvolvido por Santos et al.(2010), os teores de argila e de MO obtidos pelo vis-NIRS não diferiram dos resultados dos métodos convencionais nos solos estudados. Os autores afirmam que os coeficientes de correlação acima de 0,70 obtidos para teor de argila e MO indicam que o sistema de análises vis-NIRS pode, em alguns casos, auxiliar os métodos convencionais sem comprometer a precisão e exatidão dos resultados.

Em outro estudo conduzido por Islan et al. (2003) foi utilizada a espectroscopia nas faixas UV-VIS-NIRS para prever diversas propriedades do solo, ao autores avaliaram o coeficiente de determinação (R^2) para as propriedades do solo e depois compararam com vários comprimentos de onda. A maior correlação para ferro livre foi encontrada na região UV, para carbono orgânico na região do VIS e para argila na região do vis-NIRS. A região do vis-NIRS apresentou uma melhor resposta na predição do teor de água e CTC.

2.7 MO X REFLECTÂNCIA

A composição e o conteúdo de MO no solo são reconhecidamente fatores que apresentam muita influência sobre a reflectância dos solos. À medida que o teor de MO aumenta, a reflectância do solo decresce no intervalo de comprimento de onda de 400 a 2500 nm (HOFFER E JOHANNSEN, 1969).

Segundo Baumgardner et al. (1970), quando o teor de MO no solo excede a 2,0 %, ela assume um papel importante na determinação das propriedades espectrais do solo. Quando o teor é menor que 2,0%, outros constituintes do solo passam a ser mais influentes no seu comportamento espectral do que a MO. Os autores reforçam também que, mesmo as feições relacionadas a MO sendo muito sutis e muitas vezes não perceptíveis ao olho humano, a reflectância na faixa do visível é evidente.

Al-Abbas et al. (1972) comentam que os efeitos de diminuição da reflectância do solo pelo aumento do teor de MO podem estar associados também com o teor de argila, uma vez que deve existir uma boa correlação entre os teores de argila e de MO nos solos.

Para prever o teor de MO para solos australianos, baseando-se somente na informação espectral das amostras, Islam et al. (2003) conseguiram um coeficiente de determinação (R^2) de 0,72.

2.8 ARGILA X REFLECTÂNCIA

Quando se trata de textura do solo, o foco tem sido o teor de argila nos estudos de reflectância pela sua influência na estrutura, promovendo a formação de agregados do solo e propriedades de inchaço e diminuição, formando rachaduras. A dinâmica da água e a aeração do solo são altamente dependentes da textura e estrutura e, portanto, importantes para o crescimento das plantas, e também para regulação de microorganismos que fazem a decomposição e processos de ciclagem de nutrientes (STENBERG, 1999).

Apesar da argila ser definida com partículas menores que 2 mm, a influência da mineralogia na região vis-NIRS do espectro pode ser vista como um recurso valioso para as previsões de teor deste atributo. Ben-Dor e Banin (1995a) encontraram bandas importantes para calibrações de teor de argila.

Chang et al. (2001) afirmaram que o teor de argila é considerado o atributo que apresenta melhor desempenho nas previsões. Em outra pesquisa utilizando dados do comprimento de onda da região vis-NIRS do espectro, Moron e Cozzolino (2003) avaliaram o conteúdo de areia, argila e silte em solos do Uruguai. Os resultados de R^2 obtidos na calibração foram, superiores a 0,8 para as frações de areia, silte e argila.

Vários métodos estatísticos e matemáticos podem ser utilizados na calibração dos modelos envolvendo dados de reflectância e, por trabalhar bem com dados não lineares, o uso de técnicas de mineração de dados, tais como RNA é crescente (BROWN et al., 2006). Outras técnicas que podem trabalhar com muitas variáveis preditoras são regressões: utilizando componentes principais (RCP), mínimos quadrados parciais (PLS) e regressão linear simples (RLS) ou múltipla (RLM).

2.9 REDES NEURAIS ARTIFICIAIS - RNAS

Com a evolução dos computadores e equipamentos eletrônicos de um modo geral, aliados com a grande capacidade de armazenamento, a aquisição de dados se tornou mais eficiente. O desafio agora é como extrair destes dados padrões de comportamento e novos conhecimentos.

Uma das técnicas mais utilizadas no reconhecimento de padrões atualmente são as Redes Neurais Artificiais. A função de uma RNA nada mais é que estabelecer uma relação linear ou não, entre um conjunto de dados de entrada e uma saída (resposta correspondente).

Para chegar neste objetivo podem ser utilizadas transformações matemáticas nos dados de entrada, de modo a produzir uma saída desejada. A rede realiza este trabalho “aprendendo” como o sistema em estudo se comporta e, então, aplica este conhecimento a novos dados de entrada para prever a saída apropriada (POPPI e SENA, 1999).

A maioria dos sistemas naturais não é linear (STARK e HARDY, 2003). Smith (2003) comenta que um sistema é não linear quando não se obedecem, simultaneamente, os princípios da proporcionalidade e da aditividade. O tipo de arquitetura multicamada é um dos modelos de rede mais utilizado nos problemas envolvendo dados não lineares (HAYKIN, 1999).

Técnicas que utilizam algoritmos que trabalham com dados não lineares como as RNAs são uma boa opção e, entre as várias arquiteturas existentes, uma que se destaca é a multicamada perceptron utilizando a validação cruzada que, segundo Witten e Frank (2005), é um bom meio para determinar a potencialidade de métodos preditivos. A validação cruzada (do inglês cross-validation) é uma metodologia com boa aceitação, pois cria um modelo mais genérico, ou seja, ela utiliza toda a base e constrói o modelo sobre todos os dados.

Os métodos de RNAs têm a capacidade de generalização do aprendizado, de maneira a reconhecer instâncias similares que nunca haviam sido treinadas e que apresentam bom desempenho em tarefas com falta de conhecimento explícito sobre como encontrar uma solução (BITTENCOURT, 1998).

Com o objetivo de representar um evento do sistema nervoso, as RNAs foram desenvolvidas a partir do modelo neuronal proposto por McCulloch e Pitts em 1943. Os neurônios do cérebro humano processam informações e se comunicam com milhares de

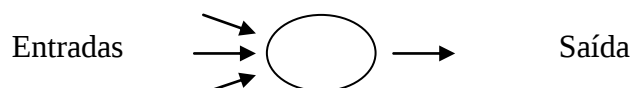
outros neurônios, continuamente e em paralelo. O cérebro humano é responsável pelo que se chama de emoção, pensamento, percepção e cognição, bem como pela execução de funções sensoriais, motoras e autônomas. Além disso, sua rede de neurônios tem a capacidade de reconhecer padrões e relacioná-los, usar e armazenar conhecimento por experiência, além de interpretar observações. Não se sabe ao certo a forma como as funções cerebrais são realizadas, mas é nessa estrutura fisiológica que se baseia as RNAs (HEWITSON e CRANE, 1996). Assim, uma RNA é composta por vários neurônios distribuídos em camadas. A primeira camada é chamada de entrada e tem a função de armazenar os dados que chegam a rede. As camadas intermediárias, denominadas camadas escondidas, realizam os processamentos da rede, e a última camada, definida como de saída, recebe os valores já processados pela rede. A quantidade de neurônios por camadas, o número de camadas por rede e os pesos ajustados após o aprendizado é chamado de arquitetura da rede neural (ITO et al., 2000).

As RNAs são compostas por cinco estruturas básicas:

1- Neurônios Artificiais

São as unidades básicas de processamento da informação, projetadas para simular o comportamento de neurônios biológicos. Assim, os neurônios artificiais possuem um corpo de processamento de informação com duas ramificações: entradas (dendritos) e saída (axônio), conforme representado na Figura 7.

Figura 7 - Representação básica de um neurônio artificial.



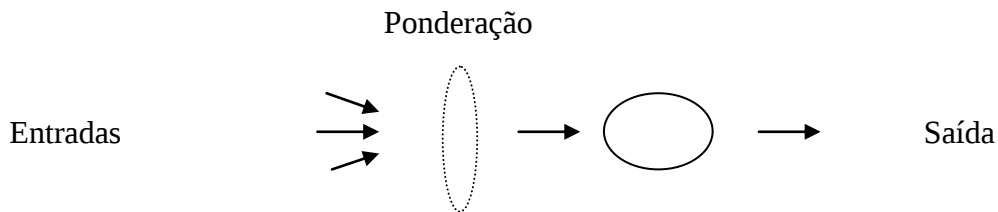
O neurônio artificial, elemento de processamento ou ainda unidade de processamento, denota um dispositivo simples capaz de calcular o somatório de N entradas ponderadas, cujo resultado passa por uma função não linear. Basicamente este dispositivo é caracterizado por um *offset* ou limiar interno e por algum tipo de função não linear. A saída Y do dispositivo é função somatório das entradas $X_0, X_1,$

..., X_{n-1} , ponderadas pelos pesos $W_0, W_1, \dots, W_{n-1}$ menos o *offset* θ (MEDEIROS, 1999).

2- Sinapses e Pesos

Em um neurônio artificial a intensidade da sinapse é simulada por um fator de ponderação chamado *peso da sinapse* ou simplesmente peso (Figura 8). O sinal total que entra no corpo de processamento de um neurônio artificial é calculado através da simples multiplicação do sinal que chega ao neurônio pelo peso da sinapse em questão (CERQUEIRA et al., 2001).

Figura 8 - Representação da sinapse artificial.



Neurônios possuem um grande número de entradas e, portanto, podem estabelecer várias sinapses com outros neurônios da RNA, simultaneamente.

3- Funções de transferências

A saída de um neurônio (output) será obtida em função do sinal de entrada (input). Esta função deve possuir as propriedades:

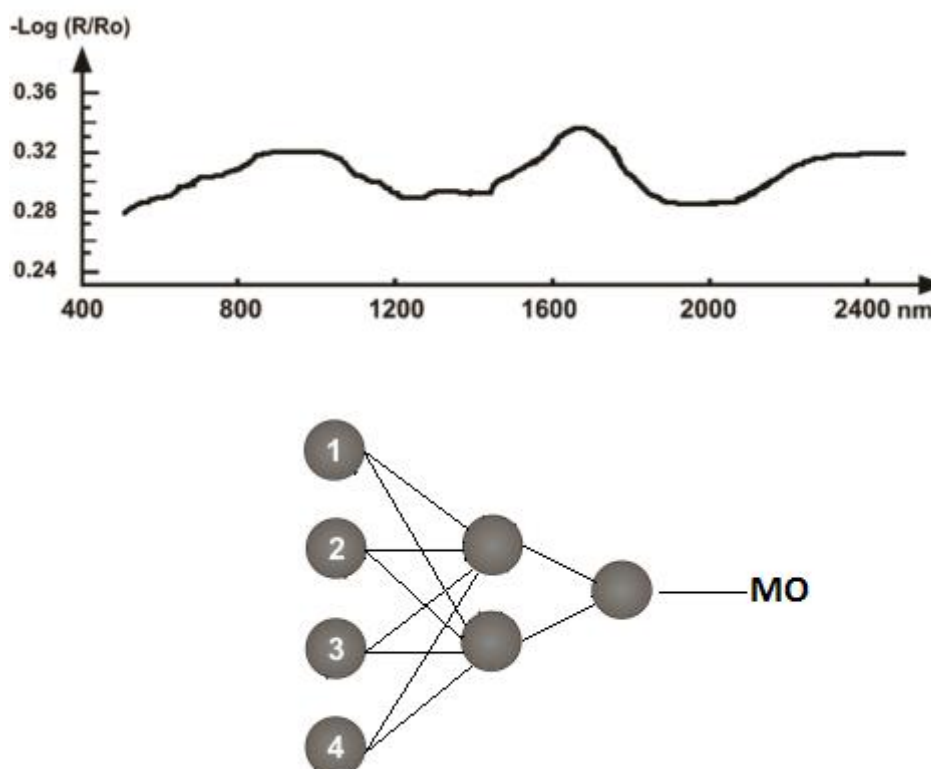
- não ser negativa, pois ou o neurônio troca sinal com outro neurônio ou não troca.
- deve ser contínua, pois um neurônio não pode trocar sinal com outro neurônio, indefinidamente.

4- Arquitetura de RNA

Uma RNA é composta por uma camada de entrada e uma camada de saída. Entre estas camadas, existe um número variável de camadas intermediárias (ocultas). A esta disposição das camadas e número de neurônios por camada, dá-se o nome de arquitetura da RNA.

Um exemplo simples do emprego de RNAs para predição da MO por meio de informações obtidas pela técnica vis-NIRS é apresentado na Figura 9. A arquitetura da RNA utilizada é formada por uma camada de entrada com quatro neurônios, uma camada intermediária com dois neurônios e uma camada de saída com um neurônio. As funções de transferência utilizadas estão indicadas na própria Figura 9. As entradas (inputs) foram as intensidades, isto é, valores lidos em diferentes comprimentos de onda em espectros de reflectância difusa na região do infravermelho visível e próximo, e a saída (output) foi o valor da MO obtido nas análises laboratoriais.

Figura 9 - Ilustração da arquitetura de uma RNA utilizada para a determinação de MO (camada de saída) a partir de dados resultantes da análise de espectroscopia de reflectância difusa no infravermelho visível e próximo da MO na camada de entrada.



5 - Treinamento e validação de RNAs

Existem muitos tipos de RNAs e várias maneiras de classificá-las. Talvez a mais importante seja quanto à forma de aprendizado que pode ser: supervisionado e não supervisionado.

- **Supervisionado:** são sucessivamente apresentados à rede conjuntos de padrões de

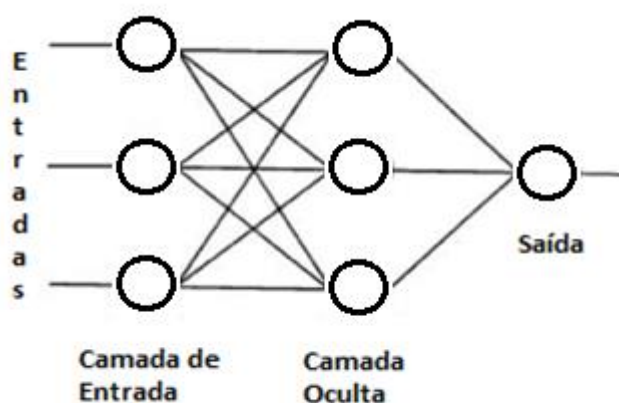
entrada e seus correspondentes padrões de saída. Durante este processo, a rede realiza um ajustamento dos pesos das conexões entre os elementos de processamento, segundo uma determinada lei de aprendizagem, até que o erro entre nos padrões de saída gerados pela rede e alcancem um valor mínimo desejado. Exemplos deste tipo de aplicação são os métodos *perceptron* e *backpropagation* (NIELSEN, 1989).

- **Não-supervisionado:** o treinamento consiste da apresentação apenas dos vetores de entrada, a partir dos quais são extraídas as características desse conjunto de padrões, agrupando-os em classes. O treinamento não supervisionado pode ser observado como um processo autônomo ou auto-organizável (WASSERMAN, 1989).

As redes também podem ser classificadas quanto as suas estruturas como as redes de múltiplas camadas - *multilayer feedforward network* -, cujo fluxo de dados segue uma única direção e redes recursivas - *recurrent network* ou *Backpropagation*.

O uso de RNA para solução de um dado problema consiste na repetição de um de conjunto de treinamentos. Treinar uma rede significa ajustar a sua matriz de pesos de entrada de forma que o vetor de saída coincida com certo valor desejado para cada vetor de entrada (Figura 10). O processo de treinamento é repetido até que o erro para o conjunto de treinamento alcance valor mínimo ou limite previamente determinado (HAYKIN, 1999).

Figura 10 - Estrutura de uma RNA do tipo *multilayer feedforward*.



Fonte: Adaptado de Hewitson e Crane (1994).

2.10 ARQUITETURA MULTILAYER PERCEPTRON (MLP)

Haykin (1999), define a MLP como uma rede neural artificial *feedforward* composta de uma camada de entrada, uma ou mais camadas escondidas, e uma camada de saída. Cada uma das camadas é composta de uma ou mais unidades simples de processamento de informação denominados neurônios. No caso da MLP, os neurônios existentes na camada de entrada são responsáveis por distribuir os sinais de entrada aos neurônios da(s) camada(s) ocultas(s) subsequentes. Os neurônios contidos nas camadas escondidas são elementos de processamento não linear, também conhecidos como perceptrons. Um perceptron recebe uma soma ponderada dos seus sinais de entrada e passa o resultado dessa soma por uma função não linear.

A arquitetura das redes MLP são acíclicas, ou *feedforward*, com uma ou mais camadas intermediárias. A diferença entre redes acíclicas e cíclicas, está no fato de uma rede cíclica possuir pelo menos um link de realimentação (feedback), pelo qual a saída de algum neurônio de uma camada i é utilizada como entrada para um neurônio de camada de ordem menor ou igual a i (DIAS et. al, 2004).

Utilizando-se da arquitetura MLP o algoritmo *Backpropagation* é baseado no gradiente descendente do erro, ou seja, na minimização do erro quadrado médio (EQM). Dependendo dos pesos iniciais, do número de neurônios intermediários e do coeficiente de aprendizado, pode acontecer de a RNA atingir um erro mínimo. Se esse mínimo local significar um EQM aceitável, o aprendizado está realizado, caso contrário é necessário continuar o treinamento (RIBEIRO, 2003). O aprendizado na rede *Backpropagation* é supervisionado, a rede produz sua própria saída, comparando-a com a saída desejada. Se houver diferença os pesos são modificados com o objetivo de diminuir esta diferença. Este algoritmo apresenta duas fases bem delineadas. Na primeira é realizado o treinamento da rede e na segunda ocorre a validação do treinamento, é quando a rede deve reconhecer os padrões que foram treinados. As etapas do treinamento são:

- Inicialização dos pesos com valores dentro de determinado intervalo;
- Apresentação do vetor de entrada com as respectivas saídas desejadas;

- Cálculo da saída;
- Adaptação dos pesos, começando pela camada de saída;
- Apresentação de novas entradas até que todos os padrões sejam apresentados e o erro esteja tão pequeno quanto se deseja.
- Fase de reconhecimento, a rede utiliza o último conjunto de pesos resultante do treinamento.

Ficando disponíveis para a rede os valores de entrada sem os valores de saída associados. O valor de saída correspondente ao padrão apresentado será calculado pela rede e se o padrão foi aprendido o valor será igual ao valor esperado.

Um exemplo de aplicação de RNA em análises de solo foi construído por Chao et al. em 1991 quando elaborou-se um modelo não linear para prever o coeficiente de absorção de poluentes orgânicos no solo. Esses coeficientes normalmente são calculados por uma relação linear logarítmica com algum parâmetro do composto orgânico. A rede apresentou uma arquitetura 3-2-1, com três neurônios na camada de entrada, dois na camada intermediária e um na camada de saída. A performance do modelo não linear fornecido pela RNA foi comparada com modelos lineares (regressões contra cada parâmetro e contra os dois conjuntamente). A RNA forneceu os melhores resultados: menor soma de resíduos e menor desvio na previsão das amostras.

Outro exemplo de aplicação na área agrônômica pode ser visto em Levine et al.(1994) que classificaram uma grande quantidade de solos em três categorias, de acordo com sua granulometria. Mouazen et al. (2010) compararam RNA com regressão por componentes principais (RCP) e mínimos quadrados parciais (PLS) para a previsão das propriedades de um solo selecionado. Eles descobriram que combinando PLS com RNA os modelos ofereciam melhores previsões em relação ao PLS com RCP. Viscarra Rossel e Behrens (2010) comparam o uso de PLS a uma série de algoritmos de mineração de dados e técnicas de seleção de recursos para as previsões de argila, carbono orgânico e pH. Seus resultados sugerem que os algoritmos de mineração de dados produziram resultados mais precisos do que o PLS. Em comparação a outras abordagens, Liu e Goering (1999) sugerem que o desempenho dos modelos de RNAs tem melhor possibilidade de prever a MO, especialmente quando as relações existentes entre os fatores são complexas e não-lineares.

2.11 SOFTWARE ESTATÍSTICO SAS

A estatística fornece métodos que auxiliam o processo de tomada de decisão. O software Statistical Analysis System – SAS, é uma das ferramentas disponíveis no mercado com esta finalidade para análises de dados. O *proc reg* é um dos procedimentos, entre outros de regressão do SAS que tem como principais funções:

- Especificação de múltiplos modelos;
- Métodos de seleção de modelos;
- Diagnósticos de regressão;
- Obtenção de valores preditos;
- Gráficos de resíduos.
- Diagnose de multicolinearidade;

Junto com o procedimento *proc reg* pode-se utilizar o método de seleção de modelo *stepwise*. Cada passo, após a entrada de uma das variáveis candidatas, são testadas as variáveis que estavam no modelo. Se uma ou mais delas apresentarem F parcial não significativo ($P > 0,05$), aquela que tiver menor valor de F parcial deve sair do modelo. Este sistema de retiradas é realizado com uma variável de cada vez, repetindo até não ter mais variáveis no modelo que apresentem resultados parciais não significativos. As variáveis que foram retiradas do modelo, não são mais candidatas a entrar. As variáveis remanescentes, candidatas a entrar no modelo, são colocadas uma de cada vez no modelo final e o processo continua com entradas e saídas até não existir mais variáveis novas para entrarem ou até as variáveis não atingirem o nível mínimo de significância para entrarem no modelo e as variáveis do modelo forem todas significativas (FERREIRA, 2007).

2.12 A FERRAMENTA WEKA

Desenvolvido em **Java**, dentro das especificações da **GPL** (*General Public License*) o **Weka**¹ é um *Software* livre do tipo *open source*, para mineração de dados (HALL et al., 2009).

O Weka pode ser definido como uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados, contendo ferramentas para pré-processamento,

¹ Disponível em: <http://www.cs.waikato.ac.nz/ml/weka/>

classificação, regressão, clusterização (agrupamento), regras de associação e visualização (WITTEN e FRANK, 2005). O sistema foi desenvolvido por um grupo de pesquisadores da Universidade de Waikato na Nova Zelândia e ao longo dos anos se consolidou como a “ferramenta” de MD mais utilizada em ambiente acadêmico.

O sistema pode ser utilizado no modo console ou por meio de uma interface gráfica chamada Weka Explorer (WITTEN e FRANK, 2005) que facilita sua utilização, abstraindo a complexidade dos procedimentos utilizados, fornecendo relatórios objetivos com os resultados das análises solicitadas, sejam elas, utilizando algoritmos de mineração de dados como de RNA. O algoritmo de RNA utilizado neste trabalho implementa uma rede do tipo *MLP* que pode ser utilizada tanto para classificação como para regressão.

Segundo Hall et al. (2009), o Weka possui implementada uma arquitetura de RNA baseado em MLP, para classificação de bases de dados. Esse módulo usa o algoritmo *Backpropagation* para classificar as informações da base de dados, criada por um algoritmo ou ambos os casos. A rede também pode ser monitorada e modificada durante o tempo de treinamento.

Quanto a métodos para validação dos experimentos de treinamento, no próprio manual do Weka (HALL et al., 2009), a ferramenta traz quatro opções. Nela o algoritmo classificador pode ser avaliado usando:

- *Use training set* – Determina a predição (regras) e testa com o próprio conjunto de treinamento submetido no classificador.
- *Supplied test set* – Determina a predição (regras) e testa em outro conjunto de teste inserido pelo usuário.
- *Crossvalidation* – O classificador é avaliado por validação cruzada, nesta opção o conjunto de testes é dividido em partes iguais e a predição é aplicada a cada parte separadamente.
- *Percentage slipt* - a ferramenta divide o conjunto de dados para aprendizado e teste de acordo com a porcentagem definida pelo usuário. (BOUCKAERT et al., 2008).

Bouckaert et al.(2008) comentam também que após o treinamento e validação, o simulador Weka traz no resultado do experimento uma série de informações, as quais eles apontam como sendo as mais importantes para serem interpretadas e saber o potencial de acertos do modelo RNA construído:

- Run information (informações sobre a execução).
- Classifier mode (informações sobre o classificador).
- Summary (lista de estatísticas, como acurácia geral do experimento de acordo com opção de validação).
- Acurácia por classe.

Os erros calculados pelo programa e mostrados ao usuário são:

- MAE: Erro médio absoluto (calcula a diferença entre os valores atuais e os preditos, é a média do erro da predição);

- RMSE: Erro médio da raiz quadrada (calculado pela média da raiz quadrada da diferença entre o valor calculado e o valor correto. É a raiz quadrada do MAE);

- RAE: Erro relativo absoluto (é o erro total absoluto, quanto menor o valor, significa maior precisão do modelo);

- RRSE: Erro relativo de raiz quadrada (reduz o quadrado do erro relativo na mesma dimensão da quantidade sendo predita, incluindo raiz quadrada);

A Figura 11 exemplifica uma saída do programa Weka, com as informações das opções escolhidas pelo usuário no cabeçalho, número de variáveis de entrada e os índices de acertos e erros para treinamento e validação do modelo. Neste exemplo as variáveis de entrada foram: IR528, IR1044 e IR2206 e o coeficiente de correlação na fase de treinamento foi de 0,93 e na fase de validação foi de 0,86:

Figura 11: Exemplo de uma saída do programa Weka.

```
Options: -L 0.4 -M 0.3 -V 0 -S 0 -E 20 -N 30 -H 1
```

```
Linear Node 0
```

```
Inputs  Weights
```

```
Threshold  0.7103188630817487
```

```
Node 1  -1.7953403397138106
```

```
Sigmoid Node 1
```

```
Inputs  Weights
```

```
Threshold  -0.16376614814733179
```

```
Attrib IR528  0.7942963874336662
```

```
Attrib IR1044  -2.5108076802218973
```

```
Attrib IR2206  0.22592786212128194
```

```
Class
```

Input

Node 0

Time taken to build model: 0.01 seconds

Time taken to test model on training data: 0 seconds

=== Error on training data ===

Correlation coefficient	0.9333
Mean absolute error	13.7897
Root mean squared error	19.2155
Relative absolute error	42.4953 %
Root relative squared error	43.2698 %
Total Number of Instances	25

=== Cross-validation ===

Correlation coefficient	0.8689
Mean absolute error	19.9479
Root mean squared error	23.4014
Relative absolute error	58.8623 %
Root relative squared error	50.6819 %
Total Number of Instances	25

3 MATERIAIS E MÉTODOS

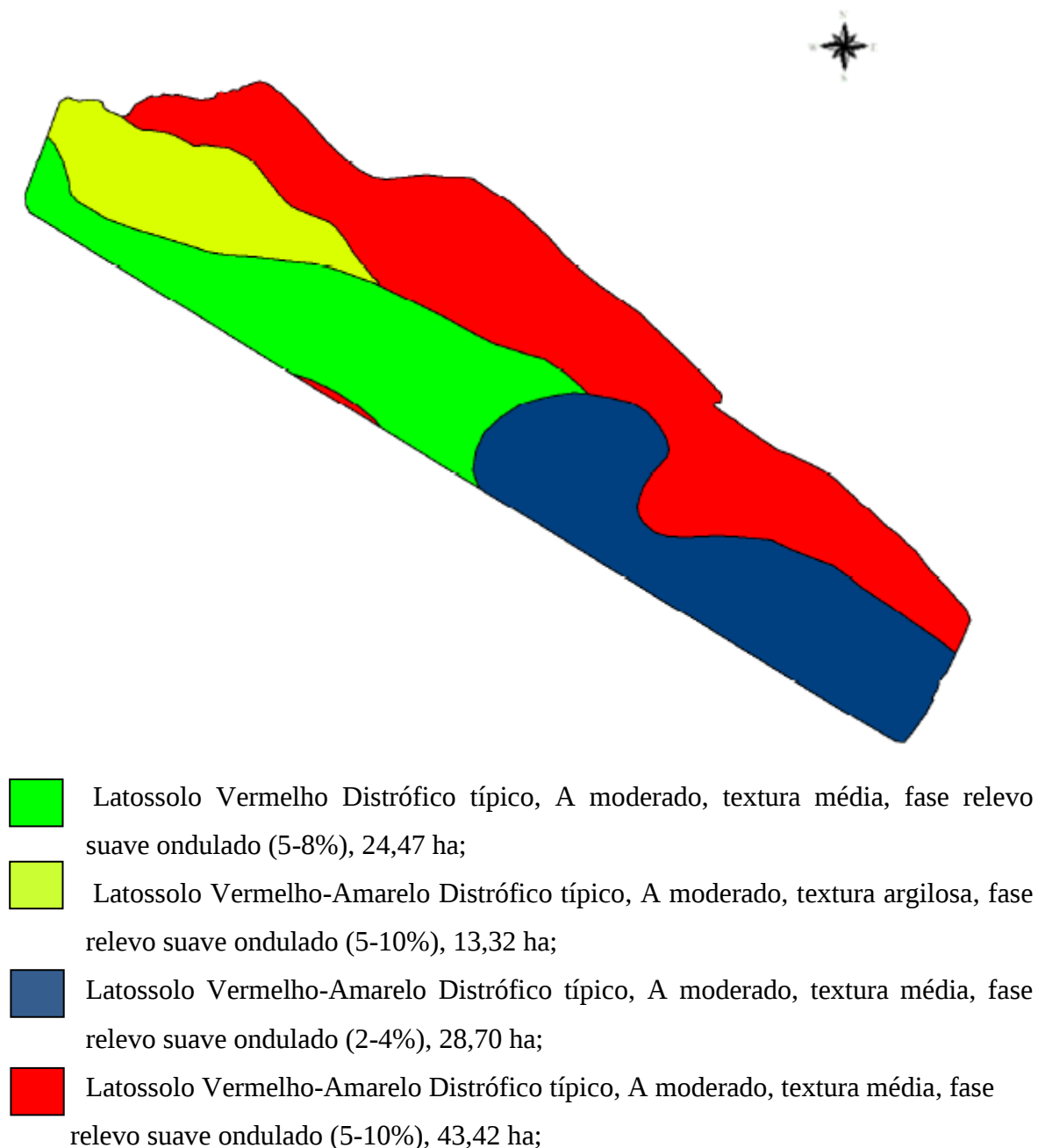
3.1 LOCAL DO ESTUDO

A região dos Campos Gerais do Paraná está localizada entre 24 e 26° Sul de latitude e 49 e 51° Oeste de longitude, com altitudes variando entre 600 e 1300 metros acima do nível do mar. Este estudo foi realizado nesta região, utilizando uma área que está localizada no município de Piraí do Sul (Figura 12), no Centro-Sul do estado do Paraná, Brasil. Sua posição geográfica tem como coordenadas 24° 22' 30" S, 50° 04' 00" O. A gleba, com uma extensão de 110 hectares, é composta predominantemente por Latossolos de textura média a argilosa (Figura 13).

Figura 12 - Figura com a Localização da propriedade, em destaque para a gleba utilizada.



Figura 13 - Mapa de classificação de solos da área.



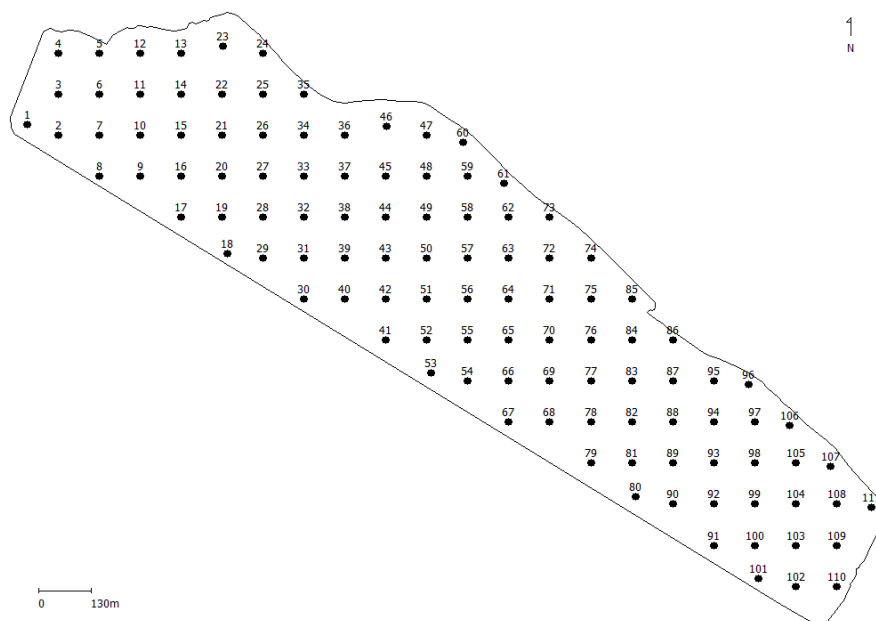
A precipitação pluvial média anual varia entre 1400 e 1800 mm, não apresentando estação seca bem definida. A estação mais chuvosa inicia em setembro e se estende até março, mas também ocorrem precipitações pluviais frequentes durante o inverno. Regionalmente, o mês de janeiro é o mais chuvoso, totalizando médias entre 150 e 210 mm, enquanto que o mês de agosto é o mês mais seco, com precipitação pluvial média entre 50 e 90 mm (TSUKAHARA et al., 2010).

Em relação às temperaturas históricas registradas nos Campos Gerais, a média anual indica valores entre 16 e 20°C no ano, variando de acordo com a latitude e altitude da região. Os meses de janeiro e fevereiro são os mais quentes do ano com temperaturas médias entre 20 e 23°C, e julho o mês mais frio, com temperaturas médias entre 12 e 16°C (TSUKAHARA et al., 2010).

3.2 COLETA E ANÁLISES DE SOLO

Para a coleta foi utilizando um grid georreferenciado (Figura 14) com uma amostra por hectare na profundidade de 00-20 cm. Para formar cada amostra foram retiradas 8 sub-amostras em torno do ponto e em seguida uniformizadas, totalizando 111 amostras de solo.

Figura 14 - Grid com os pontos identificando onde foram retiradas as amostras de solo.



Após a coleta, as amostras foram encaminhadas ao laboratório de análises físico químicas da **Fundação ABC**² para a determinação dos seguintes parâmetros:

1. Análise Química : Matéria Orgânica (MO).
2. Análises de Textura (g/kg): Argila, Silte e Areia.
3. Análise por vis-NIRS.

² Fundação ABC: Entidade que atua no desenvolvimento de pesquisa aplicada à agricultura, localizada em Castro-Pr, mantida pelas Cooperativas Capal, Batavo e Castrolanda.

3.3 PREPARO E ANÁLISE DOS ATRIBUTOS FÍSICOS E QUÍMICOS DO SOLO

O preparo das amostras consistiu em secagem em estufa com circulação forçada de ar a 40°C durante 12 horas. As amostras secas foram moídas em moinho martelo, homogeneizadas e peneiradas em peneira malha de 2 mm.

As amostras foram analisadas quanto ao teor de argila pelo método da pipeta (EMBRAPA, 1997), e a MO pelo método colorimétrico com leitura realizada em um espectrômetro na região do visível, descrito em Raij (2001).

3.4 PREPARO E ANÁLISE COM O EQUIPAMENTO NIRS

Os espectros de refletância difusa na região do infravermelho próximo foram obtidos em um espectrômetro marca FOSS, modelo XDS Near-infrared (Figura 15). O software utilizado foi o ISIscan versão 3.2 para aquisição dos espectros e o WinISI II, versão 1.5 para a montagem das curvas.

Figura 15 - Foto do espectrômetro NIRS, Marca Foss, modelo XDS.



A Figura 16 mostra a retirada da terra da caixa que contém a amostra, sendo colocada no compartimento que é inserido no equipamento FOSS.

Figura 16 - Preparo da amostra de solo para ser analisada pelo equipamento Foss.



A Figura 17 mostra em detalhes a amostra pronta para ser analisada no equipamento FOSS.

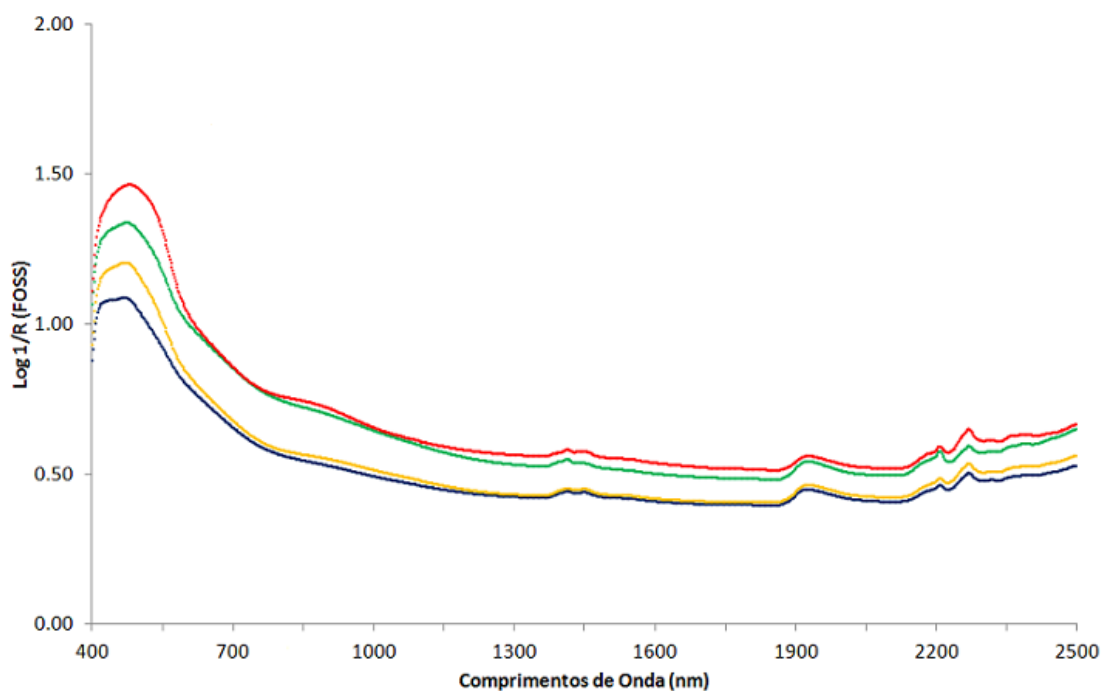
Figura 17 - Foto de uma amostra de solo pronta para ser analisada pelo equipamento Foss.



3.5 OBTENÇÃO DOS DADOS VIS-NIRS

Os espectros das amostras de solo foram obtidos a partir da leitura em nanômetros com incremento de 2 nm entre as varreduras, na faixa de comprimentos de onda de 400 a 2500 nm de forma automática. Todos os espectros de refletância difusa foram automaticamente convertidos a $\log(1/R)$ exemplificados por meio de quatro amostras de argila na Figura 18, e o “R” representa a refletância. Os teores de MO e argila, obtidos pelas análises de rotina, foram inseridos no software WinISI após a leitura das respectivas amostras.

Figura 18 – Exemplo da representação gráfica de espectros de reflectância de quatro amostras de argila.



3.6 PRÉ-PROCESSAMENTO

A detecção de “outliers”, que são valores extremos (positivos ou negativos), típicos é tão importante quanto as técnicas que serão aplicadas nas análises dos dados. 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. Nesta pesquisa, por não haver nenhum valor discrepante, não foi descartada nenhuma amostra.

3.7 ANÁLISE ESTATÍSTICA DOS DADOS

Os resultados das análises utilizando vis-NIRS geram muitas variáveis e, quanto mais variáveis são utilizadas no modelo, pode ser que a correlação em alguns casos seja melhor. Por outro lado, isso pode aumentar a quantidade de informações irrelevantes, o que pode chegar a uma superestimativa dos resultados e uma previsão ruim para outras amostras. Assim, escolher o número de variáveis que corresponda ao valor mínimo para a diferença de erros, é fundamental, pois a partir desse valor não se tem mais uma variância significativa, ou seja, a maior parte da informação estatística útil já está no modelo. Outro ponto importante quando se tem um número excessivo de variáveis é a utilização de um ambiente computacional de alto nível, para que operações sejam feitas com maior rapidez, eficiência e precisão.

Uma das razões para se utilizar análise de regressão é tentar prever valores de uma variável dependente de acordo com um conjunto de valores das variáveis independentes (AFIFI; AZEN, 1979). A respeito da regressão, Sounis (1971) diz que ela não tem a finalidade de usar uma função para representar exatamente os valores amostrados, mas, representar aproximadamente esses valores, sempre levando-se em consideração uma margem de erro.

Utilizou-se o software estatístico SAS, com análises de regressão baseadas no coeficiente de determinação (R^2) e no grau de significância para todas as amostras na geração dos modelos até 5%. Foram selecionadas as bandas espectrais que tiveram maior importância após análise destes fatores. Este procedimento foi realizado para selecionar o grupo de variáveis que mais correlacionavam linearmente com a variável de resposta.

Essas variáveis selecionadas foram utilizadas como entrada em uma RNA. Como esta RNA não produziu resultados significativos ao prever MO e argila, realizou-se um segundo passo transformando os dados com as funções: inverso, logaritmo e quadrado. Foram testadas todas as possíveis combinações de aplicação destas transformações, melhorando o resultado da RNA.

3.8 CONSTRUÇÃO DA RNA – TREINAMENTO E VALIDAÇÃO

A partir destes valores obtidos pelos modelos ajustados após as transformações, fez-se as entradas no software Weka utilizando o algoritmo MLP com validação cruzada 10 *folds*, mantendo-se quase todos os valores *default* do simulador com exceção de dois: número de épocas e número de neurônios na camada escondida.

Para facilitar os testes de seleção da melhor configuração da RNA, foram desenvolvidos três scripts. No primeiro (Quadro 1) a cada treinamento e validação, o número de épocas e neurônios foram alterados, como segue abaixo:

número de épocas: 30,50,100,500,1000,1500 e 2000.

número de neurônios: de 1 até 30.

A RNA inicia com 30 épocas e 1 neurônio, e a cada execução aumenta um neurônio chegando até 30 neurônios, iniciando novamente com 50 épocas e 1 neurônio, e assim sucessivamente, finalizando com 2000 épocas e 30 neurônios.

Primeiro script :

Quadro 1: Código do primeiro script

```
for %%e in (30 50 100 500 1000 1500 2000) do
    for /L %%n in (1, 1, 30) do
        java -weka.jar weka.classifiers.functions.MultilayerPerceptron -L 0.3 -
        'M 0.2 -V 0 -S -E 20 -t base_dados.csv
        -N %%e -H %%n > resultados\RNA-%%e-%%n.txt
```

A cada troca destes parâmetros, gerou-se um resultado da RNA com informações sobre treinamento e a validação em um arquivo, dando origem a uma lista de arquivos textos com nomes autoexplicativos para facilitar a identificação, como por exemplo: RNA-30-1.txt (30 épocas e 1 neurônio).

Para organizar estes resultados elaborou-se um segundo script (Quadro 2) com a função de retirar as informações importantes de cada arquivo texto e criar uma planilha, facilitando a filtragem para se chegar a melhor configuração da RNA.

Segundo script:

Quadro 2: Código do segundo script.

```

# Medidas
# Correl = Correlation coefficient
# MAE = Mean absolute error
# RMSE = Root mean squared error
# RAE = Relative absolute error %
# RRSE = Root relative squared error %
# Train = Dados de TREINAMENTO
# Val = Dados de VALIDAÇÃO
print
"Epocas,Neuronios,Train.Correl.,Train.MAE,Train.RMSE,Train.RAE,Train.RRSE,Val.Correl
.,Val.MAE,Val.RMSE,Val.RAE,Val.RRSE\n";
for my $e ( 30,50,100,500,1000,1500 2000) {
    for my $n ( 1 .. 30 ) {
        open my $in , "<", "RNA-$e$n..txt" ;
        my $txt = "";
        while( $linha = <$in ){
            $txt = $linha;
        }
        close $in;
        my @cor = ( $txt =~ /. *Correlation coefficient.* ([0-9]+\.[0-9]+).*/g );
        my @mae = ( $txt =~ /. *Mean absolute error.* ([0-9]+\.[0-9]+).*/g );
        my @rmse = ( $txt =~ /. *Root mean squared error.* ([0-9]+\.[0-9]+).*/g );
        my @rae = ( $txt =~ /. *Relative absolute error.* ([0-9]+\.[0-9]+).*/g );
        my @rrse = ( $txt =~ /. *Root relative squared error.* ([0-9]+\.[0-9]+).*/g );
        print
        "$e,$r,$cor[0],$mae[0],$rmse[0],$rae[0],$rrse[0],$cor[1],$mae[1],$rmse[1],$rae[1],$rrse[1]\n";
    }
}

```

Por meio de uma análise da planilha, verificou-se a arquitetura de RNA que apresentou melhor desempenho. E para mostrar os valores estimados por este modelo, fez-se um terceiro script (Quadro 3) tendo como entrada de dados as variáveis que continham os valores medidos pelo vis- NIRS e os atributos meta (MO ou argila). A Tabela 2 mostra um exemplo do resultado final.

Terceiro script:

Quadro 3: Código do terceiro script:

Java-cp	weka.jar	weka.classifiers.functions.MultilayerPerceptron	-T
BasedeDados.csv	ModeloGerado.model	-p 0	

Tabela 2: Exemplo da saída dos resultados estimados:

inst#	Actual	Predicted	Error
1	?	0.028	?
2	?	0.03	?
3	?	0.032	?
4	?	0.026	?
5	?	0.027	?
6	?	0.044	?
7	?	0.032	?
8	?	0.03	?
9	?	0.027	?
10	?	0.036	?
11	?	0.035	?
12	?	0.043	?
13	?	0.044	?
14	?	0.036	?
15	?	0.029	?
16	?	0.025	?
17	?	0.032	?
18	?	0.028	?
19	?	0.027	?
20	?	0.026	?
21	?	0.042	?
22	?	0.044	?
23	?	0.061	?
24	?	0.061	?
25	?	0.055	?

Para avaliar a metodologia quanto à previsão dos atributos, este estudo foi separado em duas etapas. Na primeira etapa utilizaram-se todas as mostras. A segunda etapa foi para validar esta metodologia. Nesta segunda etapa foram realizadas estimativas com um novo grupo de amostras. Foi retirado para calibração das 111 amostras um grupo de 25 amostras (escolhidas em diferentes faixas de teores de MO e argila) e gerou-se os modelos. A partir destes modelos fez-se a validação (estimativa) para outras 25 amostras (também escolhidas em diferentes faixas de teores de MO e argila), totalizando 50 amostras.

4 RESULTADOS E DISCUSSÕES

As amostras de solo utilizadas para a construção dos modelos de calibração apresentavam teor de MO na faixa de 15 a 43 g/dm³ e os teores de argila entre 123 a 534 g/kg.

A área da Fazenda deste estudo possui vários tipos de solos, como Latossolo Vermelho e Latossolo Vermelho-Amarelo, com textura variando entre média a argilosa com relevo entre 2 a 10 %.

Os resultados são apresentados em duas etapas. Na primeira etapa (Seção 4.1) foram feitos os treinamentos e validações com toda a base de dados e na segunda etapa (Seção 4.2) fez-se uma validação da metodologia dividindo a base em duas partes, uma para treinamento e outra para validação.

4.1 PRIMEIRA ETAPA – BASE DE DADOS COMPLETA

4.1.1 REGRESSÃO COM DADOS NÃO TRANSFORMADOS

Utilizou-se como atributos metas a MO e argila, sendo correlacionadas com todas as bandas (ou variáveis) resultantes do vis-NIRS no intervalo de 400 nm a 2500 nm. Os melhores modelos para esse conjunto foram de três bandas para MO e duas bandas para argila (Tabela 3), levando-se em consideração o ganho em R^2 com o maior grau de significância. Com este resultado elaborou-se os gráficos (Figuras 19 e 20) para mostrar esta correlação.

Tabela 3 - Bandas espectrais selecionadas na primeira regressão.

	Bandas Espectrais (nm)	R^2	Pr>F
MO	480, 1418 e 1846	0,75	<0,0001
Argila	454 e 480	0,81	<0,0001

Figura 19 - Gráfico de correlação com os valores de MO observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS com dados não transformados.

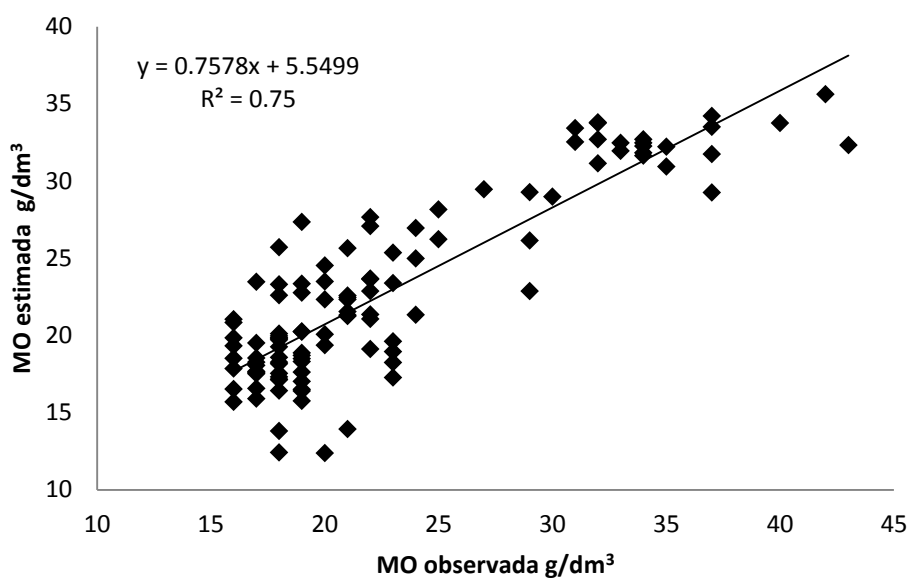
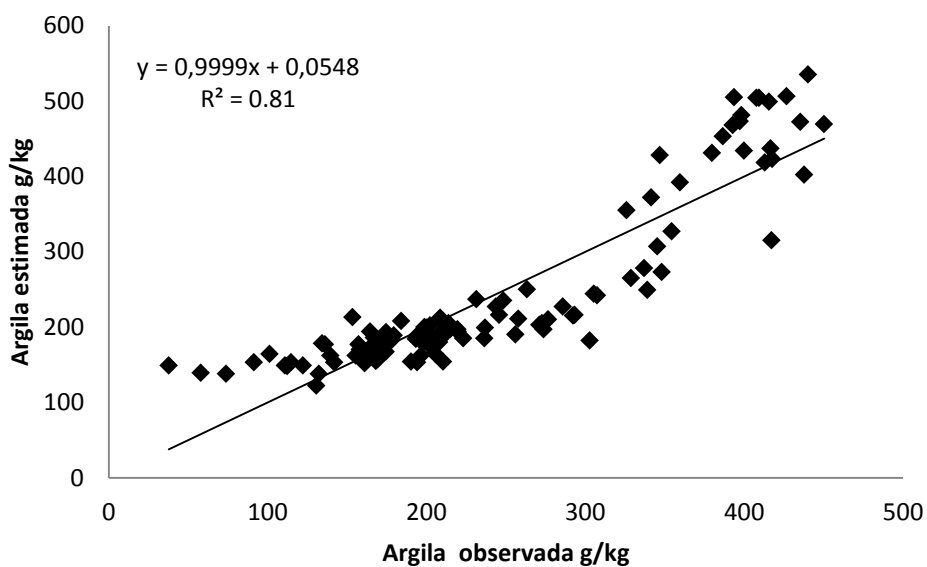


Figura 20 - Gráfico de correlação com os valores de argila observados nas análises laboratoriais versus os valores estimados pelos modelo vis-NIRS com dados não transformados.



Os resultados de MO e dos teores de argila, obtidos pelo vis-NIRS neste primeiro passo mostraram uma correlação acima de 75% para MO e de 81% para argila.

4.1.2 REGRESSÃO COM DADOS TRANSFORMADOS

Com as bandas selecionadas da primeira regressão, aplicou-se a segunda regressão com transformações cujas funções foram o inverso, quadrado e logaritmo neperiano. Os modelos de maior R^2 para esse conjunto foram de três bandas para MO: inverso da banda 480, ln da banda 1418 e inverso da banda 1846, e duas bandas para argila: inverso da banda 454 e inverso da banda 480 (Tabela 4). Para a seleção das bandas após o processo de transformação dos dados foram consideradas aquelas em que obteve-se o maior ganho em R^2 e significativos. Com este resultado elaborou-se os gráficos (Figura 21 e Figura 22) para mostrar esta correlação.

Tabela 4 - Bandas espectrais selecionadas na segunda regressão com os dados transformados.

	Bandas Espectrais (nm)	R^2	Pr>F
MO	1/ 480, LN 1418 e 1/1846	0,82	<0,0001
Argila	1/454 e 1/480	0,90	<0,0001

Figura 21 - Gráfico de correlação com os valores de MO observados nas análises laboratoriais versus os valores estimados pelo modelo vis- NIRS com dados transformados.

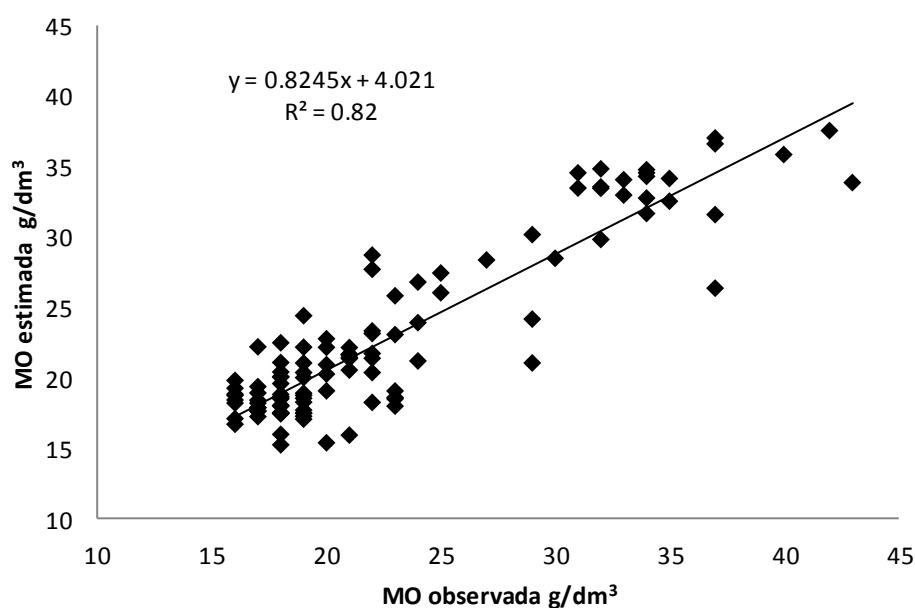
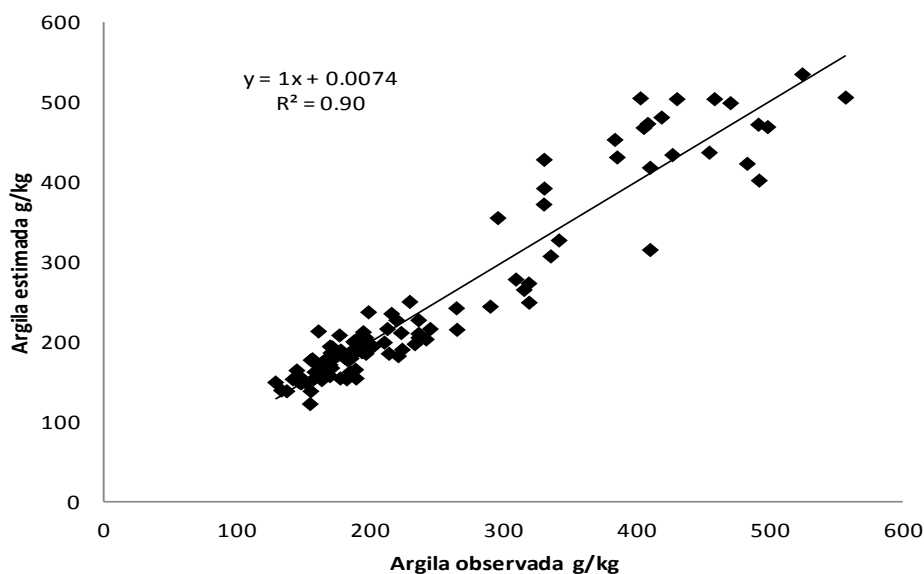


Figura 22 - Gráfico de correlação com os valores de argila observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS com dados transformados.



Os resultados obtidos com os dados transformados mostraram uma correlação de 82% para MO e de 90% para argila, aumentando o grau de determinação em relação a regressão com os dados não transformados.

4.1.3 CONFIGURAÇÃO DA RNA

Os dados transformados resultaram em um modelo mais ajustado em relação aos não transformados, deste modo, foram utilizados como entradas na RNA para gerar o modelo. Após aplicar os scripts (apresentados na seção 3.8), obter a configuração com o maior índice de acertos R^2 , fez-se o treinamento e validação da RNA utilizando o programa Weka.

As melhores arquiteturas de RNA obtidas estão descritas na Tabela 5, com o R^2 para MO chegando a 0,89 e para argila 0,94, sendo representadas pelas Figuras 23 (MO) e 24 (argila).

Tabela 5 - Valores com a configuração da RNA que obteve o melhor índice para estimar os valores de MO e argila no modelo gerado pela RNA.

	Épocas	Neurônios	R^2	MAE Validação	MAE Treinamento
MO	1000	21	0,89	2,346	2,283
Argila	500	9	0,94	25,235	23,766

Figura 23 - Representação gráfica da configuração da RNA para MO com 21 neurônios na camada escondida.

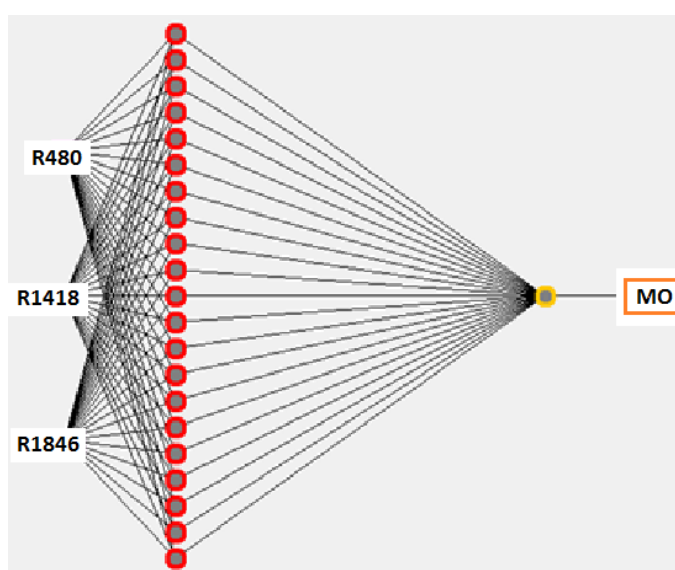
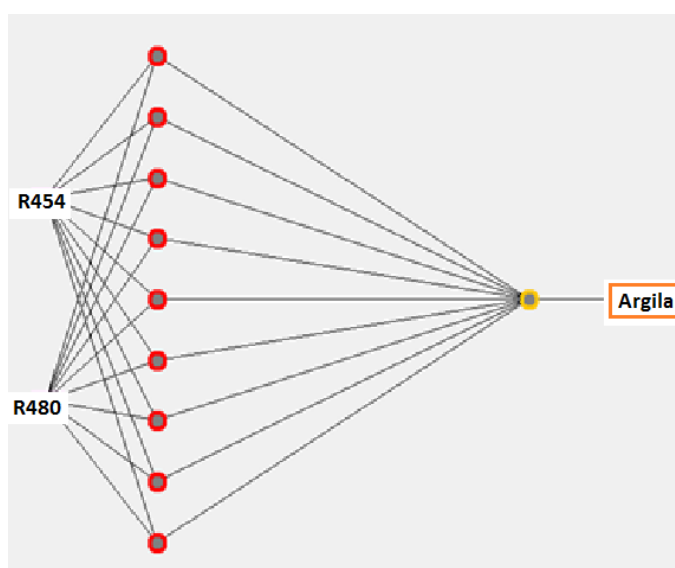


Figura 24 - Representação gráfica da configuração da RNA para argila com nove neurônios na camada escondida.



Observa-se por meio dos coeficientes de determinação que a cada passo da metodologia adotada os resultados foram obtendo um maior ganho.

Iniciando com os dados sem transformação obteve-se um índice de determinação R^2 de 0,75 para MO e 0,81 para argila. Nestes mesmos dados, aplicando-se as transformações, o resultado foi de 0,81 para MO e 0,90 para argila. Quando os dados transformados foram usados como entradas na RNA, chegou-se ao R^2 de 0,89 para MO e 0,94 para argila, evidenciando-se que houve ganho no R^2 ao testar a metodologia. Os resultados dos valores estimados pela RNA estão demonstrados no gráfico da Figura 25 para MO e no gráfico da Figura 26 para argila.

Figura 25 - Gráfico de dispersão com os valores de MO observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS.

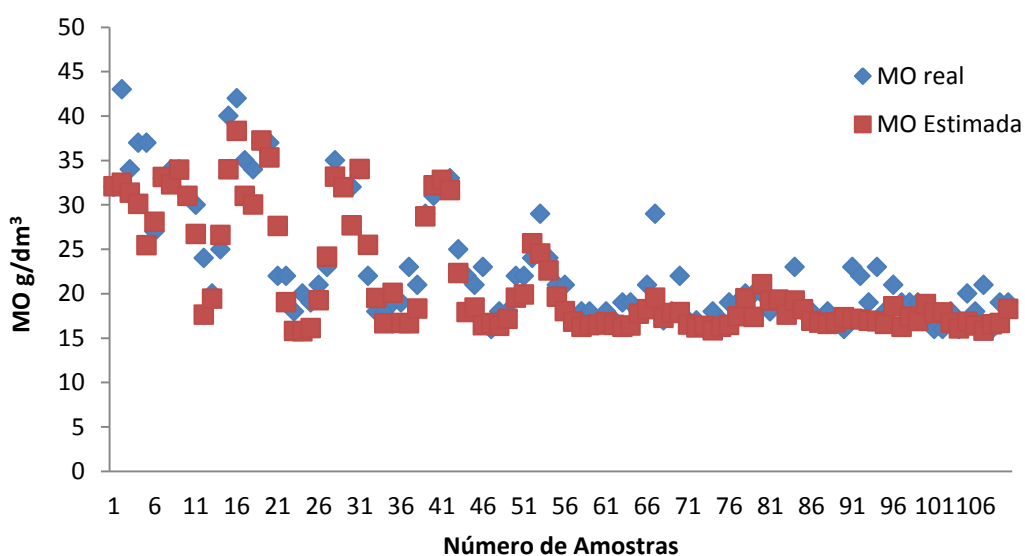
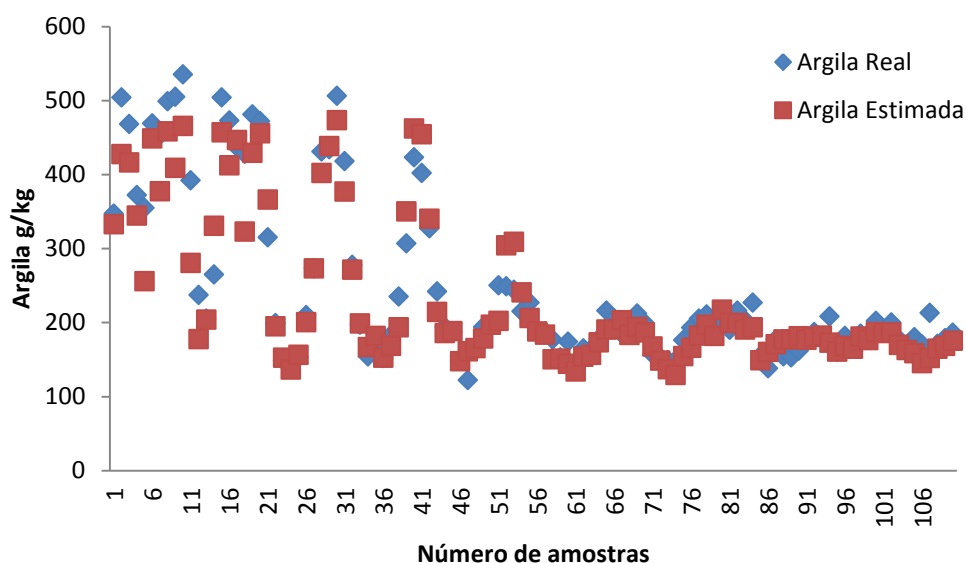


Figura 26 - Gráfico de dispersão com os valores de argila observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS.



Os resultados finais estimados pelos modelos gerados pela RNA para o teor de MO e de argila em amostras de solos apresentaram coeficientes de determinação de 0,89 e 0,94 com grau de significância menor que 0,0001, respondendo desta maneira que não existem diferenças entre o método convencional e o método vis-NIRS.

Os coeficientes de determinação sobre a estimativa do teor de MO e de argila foram superiores aos encontrados por Ben-Dor e Banim (1995) que conduziram investigações que demonstraram a capacidade da reflectância para fornecer algumas estimativas das propriedades físicas, químicas e biológicas de solos. Eles construíram modelos baseados em espectroscopia vis-NIRS e encontraram um coeficiente de determinação igual a 0,55 para MO e 0,56 para argila utilizando métodos de regressão.

No Brasil, situação semelhante foi analisada por Santos et al. (2010) por meio de modelos utilizando vis-NIRS, que encontraram um coeficiente de determinação igual a 0,70 para MO e 0,77 para argila, e FIDÊNCIO (2001) observou aumentos nos coeficientes de determinação pela substituição do método PLS por RNA, partindo de 0,88 para maiores que 0,9, respectivamente.

Chang et al. (2001), desenvolveram um estudo que estabeleceu, no caso da argila, resultados de correlação com coeficiente de determinação entre 0,8 e 1,0 ($0,8 < R^2 < 1,0$) indicando que esse atributo pode ser acuradamente quantificado pelo método vis-NIRS. Em outra pesquisa utilizando dados no comprimento de onda da região vis-NIRS do espectro, Moron e Cozzolino (2003) avaliaram o conteúdo de argila em solos do Uruguai. O resultado de R^2 obtido foi de 0,92 para argila. Em outro estudo, na Austrália, na mesma região do vis-

NIRS, Viscarra Rossel e McBratney (2008), indicaram que o valor médio de R^2 para validação cruzada deste atributo foi de 0,76.

Avaliados por vários autores como He e Song (2006) que utilizaram calibração multivariada e obtiveram correlação maior que 0,96 para MO utilizando modelos PLS em solo argiloso, e Meyer (1999) que obteve 0,91 para análise de argila, podemos dizer que o resultado deste trabalho com coeficiente de correlação de 0,89 e 0,94 para o teor de MO e argila respectivamente, está próximo ao encontrado por estes autores. Outro resultado similar para MO foi encontrado por Morimoto et al. (2004) que utilizando RNA não-supervisionada para fazer os modelos de predição e o algoritmo backpropagation obtiveram após treinamento e validação um coeficiente de determinação de 0,95.

As amostras de solos utilizadas neste trabalho foram retiradas utilizando uma profundidade única de 00-20 cm. Outros autores estudam a possibilidade deste item interferir nos índices de determinação. Em 2001 Lee et al. realizaram um estudo utilizando regressão linear múltipla com quatro tipos de solos diferentes na Flórida, em três profundidades diferentes e três épocas, totalizando 540 amostras. Nesse estudo, foram medidos vários atributos utilizando vis-NIRS, entre eles MO. Comentam que a reflectância variou dependendo da profundidade que a amostra foi retirada, resultando que a seleção do número de bandas foi em maior quantidade (até seis) para amostras retiradas com pouca profundidade (406nm, 492nm, 500nm, 506m, 522nm e 544nm) com R^2 de 0,76, e em menor quantidade quando retiradas com maior profundidade, chegando em apenas uma banda (600nm) com R^2 de 0,57. Este fato é explicado por Demattê et al. (2004) que dizem que o decréscimo no conteúdo de MO acentua as fortes relações entre a reflectância e o ferro total, favorecendo o aparecimento de uma banda de absorção bem definida ao redor de 900nm.

Krishman et al. (1980), ao estudarem a reflectância espectral de solos para identificar as bandas mais adequadas para predizer o conteúdo de MO do solo, concluíram que a região do visível proporcionou as melhores informações, com coeficiente máximo de correlação para os modelos estudados de 0,98 para as bandas na região de 623 e 564 nm, enquanto que na região do infravermelho, o coeficiente máximo de determinação para MO foi de 0,87.

Na Suécia, Wetterlind et al. (2008b) ao estimarem teores de MO com R^2 de 0,8, comentam sobre o grande potencial de análises utilizando vis-NIRS, por ser uma análise limpa, e de uma vasta aplicabilidade em várias áreas como por exemplo a AP.

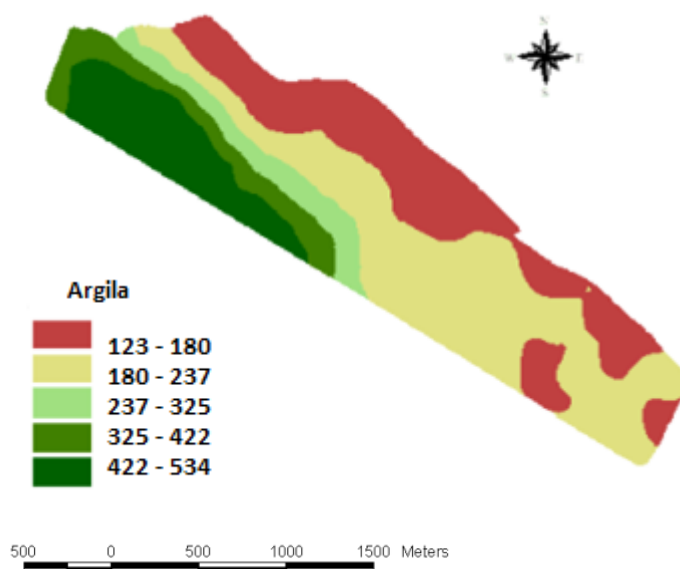
Assim, os trabalhos citados sugerem que os resultados obtidos foram satisfatórios para estimar MO e argila a partir do vis-NIRS. Este trabalho conseguiu produzir resultados também interessantes com a metodologia adotada. Foi verificado que a RNA forneceu

melhores resultados quando as variáveis de entrada foram transformadas. A seguir, são apresentados alguns resultados que permitiram validar a metodologia proposta.

4.2 SEGUNDA ETAPA - VALIDAÇÃO DA METODOLOGIA

Para validar a metodologia que foi utilizada na primeira etapa, com todas as amostras, selecionou-se 50 amostras, 25 para treinamento e outras 25 para validação. As amostras foram escolhidas aleatoriamente na região de maior variabilidade, separando 5 amostras em cada classe do mapa, que foi interpolado e classificado em 5 classes (Figura 27), pelo método de “quebras naturais” também chamado método de Jenks (JENKS, 1967). Este método divide as classes de frequência em agrupamentos naturais dos dados e define, a partir daí, a quantidade e os intervalos de dados que as classes terão dentro do tema em questão, ou, em outras palavras, pode ser entendido como um método que forma grupos que são homogêneos internamente e ao mesmo tempo maximiza a heterogeneidade entre as classes.

Figura 27 - Mapa da área dividida em cinco faixas de argila.



Utilizou-se como atributos metas a MO e argila, sendo correlacionados com todas as bandas (ou variáveis) resultantes do vis-NIRS no intervalo de 400 nm a 2500 nm.

Após a aplicação das regressões obteve-se a melhor configuração da RNA, e gerou-se o modelo para avaliar a capacidade de previsão de outras 25 amostras também escolhidas aleatoriamente.

4.2.1 REGRESSÃO COM DADOS NÃO TRANSFORMADOS

As variáveis selecionadas após a primeira regressão para esse conjunto de dados estão colocadas na Tabela 6.

Tabela 6: Bandas espectrais selecionadas na primeira regressão.

	Bandas Espectrais (nm)	R²	Pr>F
MO	576, 1850 e 1942	0,90	<0,0001, 0,0014 e 0,0340
Argila	566, 1706 e 2202	0,88	<0,0001, 0,0007 e 0,0067

Os resultados de MO e dos teores de argila, obtidos pelo vis-NIRS neste primeiro passo, mostraram uma correlação de 90% para MO e de 88% para argila.

4.2.2 REGRESSÃO COM DADOS TRANSFORMADOS

Com base nas bandas selecionadas no primeiro passo, aplicou-se a segunda regressão com transformações dos dados dos tipos: inverso, quadrado e logaritmo neperiano. O resultado está apresentado na Tabela 7.

Tabela 7 - Bandas espectrais selecionadas na segunda regressão com os dados transformados.

	Bandas Espectrais (nm)	R²	Pr>F
MO	1/MO, LN 576, 1/1850 e LN 1942	0,91	<0,0001, <0,0001 e <0,0016
Argila	LN(Argila),1/566, 1/1706 e (2202) ²	0,90	<0,0001, 0,0013 e 0,0140

Os resultados de MO e dos teores de argila, obtidos pelo vis-NIRS neste segundo passo com os dados transformados mostraram uma correlação de 91% para MO e de 90% para argila.

4.2.3 CONFIGURAÇÃO DA RNA

Os dados transformados resultaram em um modelo mais significativo em relação aos não transformados, e deste modo, foram utilizados como entradas na RNA para gerar o modelo. Aplicando-se os scripts (apresentados na seção 3.8), se obteve a configuração com o maior índice de acertos R², e fez-se o treinamento e validação da RNA utilizando o programa Weka. As melhores arquiteturas de RNA obtidas estão descritos na Tabela 8, com o R² para MO de 0,95 e argila 0,97.

As bandas transformadas resultaram em um índice de acerto mais significativo, sendo utilizadas como entradas na RNA para geração do modelo. As melhores configurações obtidas nos testes estão na Tabela 8.

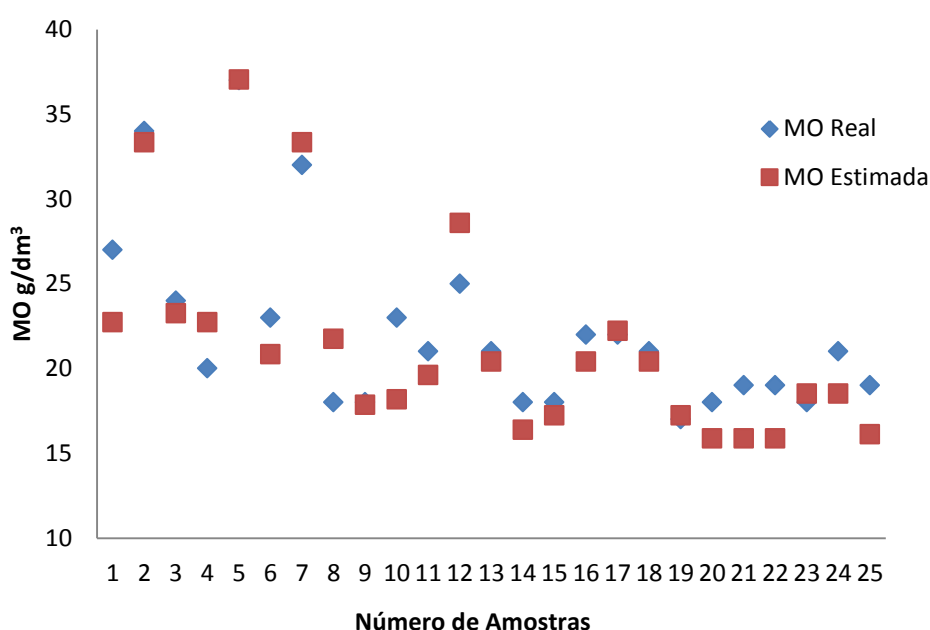
Tabela 8: Valores com a melhor configuração para a RNA para estimar os valores de MO e argila.

	Épocas	Neurônios	R²	MAE Validação	MAE Treinamento
MO	1500	1	0,95	0,002	0,002
Argila	1500	6	0,97	0,051	0,008

Ficou evidente que nesta metodologia, utilizando-se de técnicas de seleção de variáveis, por meio de regressões e transformações dos dados para servir de entrada para a geração da arquitetura da RNA o ganho no grau de determinação foi melhorado, neste caso, a MO passou de 0,89 para 0,95 e a argila de 0,94 para 0,97.

A Figura 28 abaixo mostra um gráfico de dispersão para MO entre os valores observados e os valores estimados.

Figura 28 - Gráfico de dispersão com os valores de MO observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS



Com o resultado da RNA para MO fez-se um estudo comparativo por meio de faixas de teor de MO estabelecidas em Siqueira et al. (1987) e Pauletti (2004) para observar o índice de acerto, demonstrado na Tabela 9. Os limites das faixas para MO são classificados: < 25 “Baixo”, de 25 a 50 “Médio” e > 50 “Alto”.

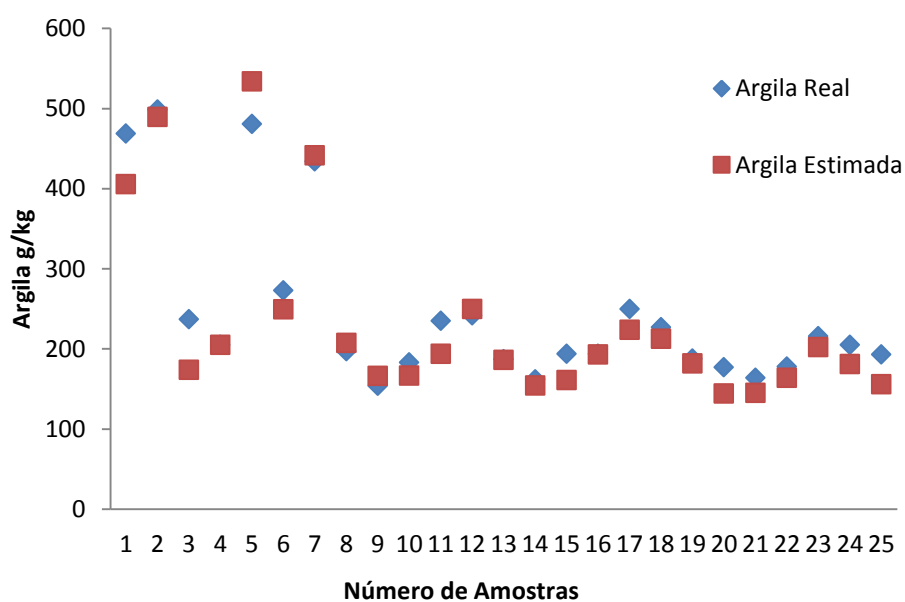
Tabela 9 - Diferenças entre o valor estimado pelo modelo vis-NIRS para o teor de MO nas amostras de solo utilizadas com o erro absoluto (EA) para cada amostra.

MO Estimada	MO Real	EA	Faixa Estimada	Faixa Real
23	27	4	Médio	Médio
33	34	1	Médio	Médio
23	24	1	Médio	Médio
23	20	-3	Médio	Médio
37	37	0	Médio	Médio
21	23	2	Médio	Médio
33	32	-1	Médio	Médio
22	18	-4	Médio	Médio
18	18	0	Baixo	Baixo
18	23	5	Baixo	Baixo
20	21	1	Baixo	Baixo
29	25	-4	Baixo	Baixo
20	21	1	Médio	Médio
16	18	2	Médio	Médio
17	18	1	Baixo	Baixo
20	22	2	Baixo	Baixo
22	22	0	Médio	Médio
20	21	1	Médio	Médio
17	17	0	Médio	Médio
16	18	2	Baixo	Baixo
16	19	3	Baixo	Baixo
16	19	3	Baixo	Baixo
19	18	-1	Baixo	Baixo
19	21	2	Baixo	Baixo
16	19	3	Baixo	Baixo

A taxa de acerto da RNA para classificar as faixas de MO foi de 100%, para este conjunto de dados de 50 amostras.

Em relação à argila a taxa de acerto não foi de 100%. Na Figura 29, mostra-se um gráfico de dispersão com o resultado da aplicação do modelo gerado pela RNA.

Figura 29 - Gráfico de dispersão com os valores de argila observados nas análises laboratoriais versus os valores estimados pelo modelo vis-NIRS.



Como pode ser observado na Tabela 10, considerando as faixas de teor de argila baseadas em Pauletti (2004), o índice de acerto foi de 92%. Os limites das faixas para argila foram classificados em: < 150 “Baixa”, de 150 a 350 “Média”, de 350 a 600 “Argiloso” e > 600 “Muito Argiloso”.

Tabela 10 - Diferenças entre o valor estimado pelo modelo vis-NIRS para o teor de argila nas amostras de solo utilizadas com o erro absoluto (EA) para cada amostra.

Argila Estimada	Argila Real	EA	Argila Estimada	Argila Real
406	469	63	Argiloso	Argiloso
490	499	9	Argiloso	Argiloso
174	237	63	Média	Média
205	205	0	Média	Média
534	481	-53	Argiloso	Argiloso
249	273	24	Média	Média
442	434	-8	Argiloso	Argiloso
208	197	-11	Média	Média
166	154	-12	Média	Média
167	183	16	Média	Média
194	235	41	Média	Média
250	242	-8	Média	Média
186	187	1	Média	Média
154	162	8	Média	Média
161	194	33	Média	Média
193	194	1	Média	Média
224	250	26	Média	Média
213	227	14	Média	Média
182	188	6	Média	Média
144	177	33	Baixo	Média
145	164	19	Baixo	Média
164	178	14	Média	Média
203	216	13	Média	Média
181	205	24	Média	Média
156	193	37	Média	Média

A validação permitiu verificar que esta metodologia pode ser utilizada com uma ferramenta na identificação de classes de MO e argila, com as vantagens de ser desenvolvida a partir da geração dos dados obtidos pela técnica vis-NIRS. O aumento no índice de acertos para previsão pelos modelos obtidos utilizando RNA podem ser decorrentes da qualidade dos dados, sem descartes. Segundo Ben-dor e Banin (1995) os métodos de calibração mais comuns para aplicações na quantificação dos atributos do solo são baseados em regressões lineares, ou seja, SMLR, PCR, e, pelo PLSR. Viscarra et. al, (2011), comenta que as RNAs tem sido pouco aplicadas para utilizar vis-NIRS na predição de elementos do solo.

Os resultados obtidos neste trabalho, tanto para MO como argila, mostram que é possível melhorar as estimativas destes componentes do solo por meio da associação de RNAs aos métodos estatísticos.

CONCLUSÕES

A metodologia desenvolvida neste trabalho para estimar teores de MO e argila, utilizando métodos estatísticos para seleção de variáveis e técnicas de RNAs para a predição de teores de MO e argila, levou a resultados satisfatórios, o que indica que a abordagem híbrida adotada melhorou a taxa de acerto de predição e sugere que o uso da espectroscopia tem grande potencial, trazendo importantes vantagens por ser uma técnica rápida e não poluente.

Sabendo-se dos ganhos relevantes pela técnica de análises utilizando vis-NIRS, principalmente na área ambiental, um dos desafios é de se gerar cada vez mais amostras para montar um banco de dados e desenvolver um sistema que possa trabalhar com todos os tipos de solos da região dos Campos Gerais no Paraná para minimizar o erro e melhorar a estimativa dos elementos que compõe o solo.

TRABALHOS FUTUROS

Algumas pesquisas que não foram possíveis de serem realizadas no curto período do mestrado e que podem ser feitas para melhorar as predições dos elementos estudados neste trabalho e de outros, são:

- a) Separação da MO da argila para estudos de comportamento das curvas espectrais;
- b) Geração de uma base de dados sobre reflectância do vis-NIRS para ajuste de modelos existentes e criação de novos modelos;
- c) Empregar outras técnicas de redução de dimensionalidade dos dados para aferir o impacto no resultado da RNA.

REFERÊNCIAS

- AFIFI, A. A.; AZEN, S. P. **Statistical analysis: a computer oriented approach**. [S.l.]: Academic Press, 1979.
- AL-ABBAS, A.H.; SWAIN, P.H.; BAUMGARDNER, M.F. **Relating organic matter and clay content to multispectral radiance of soils**. Soil Science, v.114, p.477-485, 1972.
- ALMEIDA, G.C.; KLOCK, U.; BOLZON, G.I. **Espectroscopia de Infravermelho Próximo e calibração Multivariada na Caracterização de Propriedades Físicas e Mecânicas da Madeira de Pinus spp**. Universidade Federal do Paraná, Engenharia Florestal, 2005.
- AMORIM, H. V. **Manual de métodos analíticos para o controle da produção de álcool e açúcar**. 2. ed. Piracicaba: Esalq-USP, 1996.
- ANJOS, L. H. C.; PEREIRA, M. G.; FONTANA, A. **Matéria orgânica e pedogênese**. In: SANTOS, G. de A.; CAMARGO, F.A.de O. (Ed.) **Fundamentos da matéria orgânica do solo: ecossistemas tropicais e subtropicais**. Porto Alegre: Metrópole, 2008. p. 65-86.
- ANDRIST-RANGEL, Y.; SIMONSSON, M.; ANDERSSON, S.; O' BORN, I.; AND HILLIER, S. **Mineralogical budgeting of potassium in soil: A basis for understanding standard measures of reserve potassium**. J. Plant Nutr. Soil Sci. 169, 605–615, 2006.
- AWITI, A.O; WALSH, M.G.; SHEPHERD, K.D; KINYAMARIO, J. **Soil condition classification using infrared spectroscopy: A proposition for assessment of soil condition along a tropical forest-cropland chronosequence**, Geoderma, Amsterdam, v. 143, p. 73-84, 2008.
- BAUMGARDNER, M. F; KRISTOF, S. J; JOHANNSEN ,C. J.; AND ZACHARY A. L.. **Effects of organic matter on the multispectral properties of soils**. Proceedings Indiana Academy of Science, Brookville, v.79, p.413-422, 1970.
- BEN-DOR, E.; BANIN, A. **Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties**. Soil Sci. Soc. Am. J. 59, 364–372. (1995a).
- BITTENCOURT, G. **Inteligência artificial**. Florianópolis: UFSC. 1998. 400p.
- BOUCKAERT, R.R.; E. FRANK; M. HALL; R. KIRKBY, P. REUTEMANN; A. SEEWALD; D. SCUSE. **WEKA Manual for version 3.6.0**. University of Waikato, Hamilton, Nova Zelandia, 2008
- BORGES NETO, W. **Parâmetros de qualidade de lubrificantes e óleo de oliva através de espectroscopia vibracional, calibração multivariada e seleção de variáveis**. 2005. 130 f. Tese (Doutorado em Química) - UNICAMP - Universidade Estadual de Campinas. Campinas, São Paulo, 2005.
- BROWN, D.J.; SHEPHERD, K.D.; WALSH, M.G.; DEWAYNE MAYS, M. & REINSCH, T.G. 2006. **Global soil characterization with VNIR diffuse reflectance spectroscopy**. Geoderma, 132, 273–290.

CAHN, M.D.; HUMMEL, J.W.; BROUER, B.H. 1994. **Spatial analysis of soil fertility for site-specific crop management**. Soil Science Society of America Journal 58: 1240-1248.

CAMARGO, O.A.; MONIZ, A.C.; JORGE, J.A.; VALADARES, J.M.A.S. **Métodos de Análise Química, Mineralógica e Física de Solos do Instituto Agronômico de Campinas**. Campinas, Instituto Agronômico, 2009. 77 p. (Boletim técnico, 106, Edição revista e atualizada)

CAMBARDELLA, C.A.; KARLEN, D.L. 1999. **Spatial analysis of soil version fertility parameters**. Precision Agriculture 1: 5-14.

CERQUEIRA, E.O; ANDRADE J.C.; POPPI R.J.. **Redes Neurais E Suas Aplicações Em Calibração Multivariada**. *Quim. Nova*, Vol. 24, No. 6, 864-873, 2001.

CHAO G.; GOVIND, R.; TABAK, H. H.; *Environ. Toxicol. Chem.* 1996, 15, 1089.

CHANG, C.W.; LAIRD, D.A.; MAUSBACH, M.J.; MAURICE, J.; HURBURGH, J.R. 2001. **Near-Infrared reflectance spectroscopy – principal components regression analyses of soil properties**. Soil Science Society of America Journal 65, 480–490.

CHEN, S.; HSIEH, K.W. AND CHANG W.H.. 1999. **Neural network analysis of sugar content in fruit juice**. ASAE Paper No. 99- 3083. St. Joseph, MI: ASAE.

DANIEL, K.W.; TRIPATHI, N.K.; HONDA, K.; APISIT, E., 2004. **Analysis of VNIR (400–1100 nm) spectral signatures for estimation of soil organic matter in tropical soils of Thailand**. International Journal of Remote Sensing 25, 643–652.

DALMOLIN R.S.D.; GONÇALVES C. N; KLAMT E; DICK D. P. **Ciência Rural, Santa Maria**, v.35, n.2, p.481-489, mar-abr, 2005. ISSN 0103-8478

DEMATTÊ J.A.M; CAMPOS R.C; ALVES M.C.; FIORIO PR, NANNI MR **Visible–NIR Reflectance: a new approach on soil evaluation**. Geoderma 121:95–112, 2004.

DIAS, A.H; MATHIAS I.M; VIRGENS FILHO J. S; RICKLI L.I. **Aplicação de redes neurais artificiais topologia MLP em simulação de dados radiométricos**. Revista Brasileira de Agrocomputação, Ponta Grossa - PR, v. 2, n. 2, 2004.

DU, C.; ZHOU, J., **Evaluation of soil fertility using infrared spectroscopy: a review**. DOI 10.1007/s10311-008-0166-x, 2009.

EBERHARDT, D. N.; VENDRAME, P. R. S; BECQUER, T.; GUIMARÃES, M. F. Influência da granulometria e da mineralogia sobre a retenção do fósforo em latossolos sob pastagens no cerrado. **Revista Brasileira de Ciência do Solo**, v. 32, p. 1009-1016, 2008.

EMBRAPA. Centro Nacional de Pesquisa de Solos. **Manual de métodos de análise de solo**, 2.ed, Rio de Janeiro, 1997.

EMBRAPA, Serviço Nacional de Levantamento e Conservação de Solos. **Levantamento de reconhecimento dos solos do Estado do Paraná**. Tomo I. Londrina, EMBRAPA-SNLCS/SUDESUL/IAPAR, p.414, 1984.

EMBRAPA. BUDIMAN N.; BRATNEY A.B.; SANTOS M.L.M; SANTOS H.G. **Revisão sobre Funções de Pedotransferência (PTFs) e Novos Métodos de Predição de Classes e Atributos do Solo**. EMBRAPA, I SSN 1517-262.7 Setembro, 2003.

ESSINGTON, M. E. **Soil and water chemistry: an integrative approach**. London: CRC. p.534, 2004

FERREIRA, M. M.; FERNANDES, B.; CURI, N. Mineralogia da fração argila e estrutura de latossolos da região sudeste do Brasil. **Revista Brasileira de Ciência do Solo**, v. 23, p. 507-514, 1999.

FERREIRA D. F. **Uso de Recursos Computacionais**. Universidade Federal de Lavras. FONTOURA,S.M.V; B.CIMÉLIO. **Manejo e Fertilidade de Solos em Plantio Direto**. Fundação Agrária de Pesquisa Agropecuária. Guarapuava 2007.

FIDÊNCIO, P.H. **Análise de Solos por Espectroscopia no Infravermelho Próximo e Aplicação de Métodos Quimiométricos**. 2001. 138f. Tese (Doutorado em Química Analítica). UNICAMP- Universidade Estadual de Campina, Campinas, 2001.

FIDÊNCIO, P. H.; POPPI, R. J.; DE ANDRADE, J. C.. **Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy**. Analytica Chimica Acta, Amsterdam, v. 453, p. 125-134, 2002.

HALL M.; FRANK E.; HOLMES G.; PFAHRINGER B.; REUTEMANN P.; WITTEN H.I. **The WEKA Data Mining Software: An Update**. SIGKDD Explorations, Volume 11, Issue 1, Pg. 10-18, ISSN: 1931-0145 EISSN: 1931-0153, 2009.

HAYKIN, S. **Neural networks. A comprehensive foudation**. New Jersey: Prentice Hall.842p, 1999.

HE,Y.;SONG, H. **Prediction of Soil Content Using Near-Infrared Spectroscopy**, SPIE Newsroom- International Society for Optical Engineering, [S.l.], v.2, p. 8-10, 2006.

HEWITSON B.C.; CRANE R.G. **Climate downscaling: techniques and application**. Clim. Res., vol. 7, pág. 85-95, 1996.

HILLEL, D. AND HILLEL, D. **Environmental Soil Physics**. Academic Press Ltd, London, 1998.

HOFFER, R. M.; AND JOHANNSSEN, C. J. **Ecological Potentials in Spectral Signature Analysis**. In: Johnson, P. (ed),*Remote Sensing in Ecology*, Chapter 1, 1-16. University of Georgia Press, Athens, Georgia.1969.

INGLEBY, H.R; CROWE, T.G. **Reflectance models for prediting organic matter is Saskatchewan soils**. ASAE paperN.99-3029. St.Joseph,Mich.:ASAE, 1999.

ISLAM, K.; SINGH, B.; MCBRATNEY, A. **Simultaneous Estimation of Several Soil Properties by Ultra-violet, Visible, and Near-infrared Reflectance Spectroscopy.** Australian Journal of Soil Research 41, 1193–1202, 2003.

ITO, Y.; HOSOKAMA, M.; LEE, H.; LIU, J. G. **Extraction of damaged regions using SAR data and neural networks.** Amsterdam: International Archives of Photogrammetric Engineering and Remote Sensing. v.33, part B1, p. 156-163, 2000.

JENKS, G. F. **The Data Model Concept in Statistical Mapping,** International Yearbook of Cartography 7: 186–190, 1967.

KITCHEN, N. R.. **Emerging technologies for real-time and integrated agriculture decisions, Computers and Electronics in Agriculture.** 61(2008) 1–3 LEVINE, E. R.; KIMES, D. S.; Sigillito, V. G.; Ecol. Model, 92, 101, 2008.

KRISHNAN, P. et al. Reflectance technique for predicting soil organic matter. **Soil Science Society of America Journal,** Madison, v.44, p.1282-1285, 1980.

LANDIM, P. M. B. **Análise estatística de dados geológicos.** São Paulo: Editora UNESP, 2003. 2. ed. Rev. E ampl. 253 p.

LARSON, W.E; F.J PIRCE. **The dynamics of soil quality as a measure of sustainable management.** Madison: ASA/SSSA.p.37-51, 1994.

LEE W. S.; MYLAVARAPU R.S.; CHOE J. S.; WHITNEY J.D. **Study on Soil Properties and Spectral Characteristics in Florida.** Paper Number: 01-1179. An ASAE Meeting Presentation, 2001.

LIU, J. AND C.E. GOERING. **Neural network for setting target corn yields.** ASAE Paper No. 99-3040. St. Joseph, MI: ASAE, 1999.

MARTENS, H., and NAES, T. **Multivariate Calibration** John Wiley & Sons, Chichester, UK. 419 pp, 1989.

MCCULLOCH W. S.; PITTS W. H. **A logical calculus of the ideas immanent in nervous activity.** Bulletin of Mathematical Biophysics, n. 5, 1943.

MEDEIROS, J, S. **Banco de dados geográficos e redes neurais artificiais: Tecnologias de Apoio 'a Gestao do Território.** Tese. (Doutorado em geografia física Universidade de SÃO Paulo). Faculdade de Filosofia, Letras e Ciencias Humanas - Departamento de Geografia. SP. 218p. 1999.

MELLO C.; POPPI R.J; ANDRADE J.C. ; CANTARELLA H. Analyst 124, 1999.

MEYER, J. H. **Use of NIR in the South African sugar industry with reference to soil fertility management.** South African Sugar Association Experiment Station, p. 1-13, KwaZulu- 1999.

MORGAN, M. T.; ESS, D. R.; **The Precision-farming Guide for Agriculturists. An Agriculture Primer.** Moline ,Ill.: John Deere Publisching, 1997.

MOREIRA, F. M. S.; SIQUEIRA, J. O. **Microbiologia e bioquímica do solo**. 2. ed. Lavras: Editora UFLA. p.729, 2006.

MORIMOTO, E.; SHIBUSAWA, S; KAHU, T; HIRAKO, S. **Soil Properties Prediction for Real-Time Soil sensor Based on Neural Network**, ASAE Paper No.70p004. Eds. Q. Zhang, M. Iida, A. Mizushima, 2004.

MOUAZEN, A. M.; KUANG, B.; De BAERDEMAEKER, J.; RAMON, H. **Comparison between principal component, partial least squares and artificial neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy**. *Geoderma* 158, 23-31, 2010.

MORON, A.; COZZOLINO, D. **The potential of near-infrared reflectance spectroscopy to analyze soil chemical and physical characteristics**. *Journal of Agricultural Engineering*, St. Joseph, v.140, p. 65-71, 2003. NIELSEN, H. R. *Neurocomputing*. San Diego: Addison-Wesley. 433p. 1989.

OECD; FAO. **OECD-FAO Agricultural Outlook 2010-2019**. Paris, FR, 2010. 247 p.

OBUKHOV, A. I.; ORLOV, D. S. **Spectral reflectivity of the Major Soil Groups and possibility of using diffuse reflection in soil investigations**. *Soviet Soil Science*, Washington, DC, v. 2, p. 174-184, 1964.

ODLARE M.; SVENSSON K.; PELL M. **Near infrared reflectance spectroscopy for assessment of spatial soil variation in an agricultural field**. doi:10.1016/j.geoderma.2004.09.013, 2005.

PAULETTI, V. **Nutrientes: Teores e Interpretações**. Fundação ABC, Castro-PR, 2004.

PIERCE, F. J.; NOWAK, P. **Aspects of Precision Agriculture**. *Advances in Agronomy*, San Diego, v. 67, p. 1-85, 1999.

POPPI R. J; SENA M. **AVALIAÇÃO DO USO DE MÉTODOS QUIMIOMÉTRICOS EM ANÁLISE DE SOLOS**. Departamento de Química Analítica - Instituto de Química - UNICAMP - CP 6154 - 13083-970 - Campinas – SP, 1999.

QUAGGIO, J. A.. **Acidez e Calagem em Solos Tropicais**. Instituto Agrônomo (IAC), 2000.

RAIJ, B.V. **Fertilidade do Solo e Adubação**. Editora Agronômica Ceres Ltda, 1991.

RAIJ B. V.; ANDRADE J.C; CANTARELLA H.; QUAGGIO J.A. **Análise Química do Solo para Avaliação da Fertilidade de Solos Tropicais**, pág. 177 a 180, Instituto Agrônomo, Campinas, SP., 2001.

RAIJ, B.V. **Fertilidade do Solo e Manejo de Nutrientes**. Piraciba Internacional Plant Nutrition Institute, 2001.

RESENDE, M.; CURI, N.; REZENDE, S. B.; CORRÊA, G. F. **Pedologia: base para distinção de ambientes**. 2.ed. Viçosa: NEPUT, 1997.

RIBEIRO, S. R. A.. **Integração de Imagens Landsat TM e dados auxiliares para a delimitação automática de unidades fisionômicas usando Redes Neurais Artificiais**. 160f. Tese (Doutorado em Ciencias Geodesicas) - Setor de tecnologia - Universidade Federal do Parana.Curitiba, 2003.

RESCK, D. V. S.; FERREIRA, E. A. B.; FIGUEIREDO, C. C.; ZINN, Y. L. Dinâmica da matéria orgânica no Cerrado. In: SANTOS, G.de A.; CAMARGO, F.A.de O. (Ed.) **Fundamentos da matéria orgânica do solo: ecossistemas tropicais e subtropicais**. Porto Alegre: Metrópole. p. 359-406, 2008.

ROBERT, P. C. **The Economic Feasibility of Precisium Agriculture**. In: SIMPÓSIO INTERNACIONAL DE AGRICULTURA DE PRECISÃO, 2., 2002, Viçosa. Viçosa: UFV, 2002.

SALIBA, P. R. **Desenvolvimento de Metodologias Analíticas utilizando espectroscopia visível e infravermelho próximo: classificação de açúcares, determinação de brix e acidez total em suco de laranja**. 78f. Tese (Doutorado em Química) Universidade Estadual de Campinas-SP, 1999.

SANTOS, G. A.; PEREIRA A. B.; KORNDÖRFER, G. H. **Uso do Sistema de Aanályses por Infravermelho Próximo (NIR) Para Análises de Matéria Orgânica e Fração Argila em Solos e Teores Foliares de Silício e Nitrogênio em Cana-de-Açúcar**. Biosci. J., Uberlândia, v. 26, n. 1, p. 100-108, Jan./Feb. 2010

SILVA, C. A. D.; FREITAS, E. C.; ARAÚJO, A. E.; ANDRADE, F. P. **Cultivo do Algodão Irrigado** ISSN 1678-8710 Versão Eletrônica Jan/2003 . Disponível em: <<http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Algodao/AlgodaoIrrigado/solos.htm>> Acesso em: 22 fev 2012.

SIQUEIRA, O. J. F.; SHERER, E.E; TASSINARI,G.; ANGHINONI, I.; PATELLA, J. F.; TEDESCO, M.J.; MILAN, P.A.; ERNANI, P.R. **Recomendações de Adubação e Calagem para os Estados do Rio Grande do Sul e de Santa Catarina**. Passo Fundo, EMBRAPA/CNPT, 1987, 100P.

SOUNIS, E. L. de M. **Bioestatística: princípios fundamentais, metodologia estatística: Aplicação às ciências biológicas**. Curitiba: UFPR, 1971.

SHIBUSAWA, S.; S.W. ANOM; SATO, S.; SASAO, A.; HIRAKO, S. **Soil mapping using the real-time soil spectrophotometer**. In: Grenier, G., Blackmore, S. (Eds.), ECPA 2001, Third European Conference on Precision Agriculture, vol. 1. AgroMontpellier, pp. 497– 508, 2001.

SCHUELLER, J. K. **A Review and Integrating Analysis of Spatially-Variable Crop Control of crop production**. Fertilizer Research, The Hague, v.33,p.1-34,1992.

SMITH, S.W. **Digital Signal Processing – a practical guide for engineers and scientists**. 1st ed. Newnes. 672p., 2003.

SUDDUTH, K.A. AND J.W. HUMMEL. **Evaluation of reflectance methods for soil organic matter sensing.** *Transactions of the ASAE* 34(4): 1900-1909., 1991.

STARK, J.; HARDY, K. **Chaos: Useful at Last?** *Science* 301: 1192-1193.,2003.

STENBERG, B. **Monitoring soil quality of arable land: Microbiological indicators.** *Acta Agric. Scand., Sect. B, Soil and Plant Sci.* 49, 1–24, 1999.

STENBERG, B.O.; VISCARRA ROSSEL, R.A.; MOUAZEN, A.M.; WETTERLIND, J. **Visible and Near Infrared Spectroscopy in Soil Science.** In: SPARKS, D.L. (Ed.). *Advances in Agronomy*, Burlington: Academic Press, v. 107, p. 163-215.,2010.

SHEPHERD, K. D.; WALSH, M. G. **Development of reflectance spectral libraries for characterization of soil properties.** *Soil Science Society of America Journal*, Madison, WI, v. 66, p. 988-998, 2002.

TERRA, F. DA SILVA. **Espectroscopia de reflectância do visível ao infravermelho médio aplicada aos estudos qualitativos e quantitativos de solos.** Tese (Doutorado) Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2011.

TRANGMAR, B.B.; YOST, R.S.; UEHARA, G. **Application of Geostatistics to Spatial Studies of Soil Properties.** *Advances in Agronomy* 38: 45-94.,1985.

TSUKAHARA R. Y.; CARAMORI P. H.; CAVIGLIONE J.H.; MARTORANO L.G; C. STRAUCH J.C.; GALDINO J. **Análise Climática da Região dos Campos Gerais, Pr.** Boletim Informativo, Fundação ABC , 2010.

UDELHOVEN, T.; EMMERLING, C.; JARMER, T. **Quantitative Analysis of Soil chemical Properties with Diffuse Reflectance Spectrometry and Partial-least Square Regression: a feasibility study.** *Plant Soil* 251, 319–329, 2003.

VISCARRA ROSSEL, R.A.; WALVOORT, D.J.J.; MCBRATNEY, A.B.; JANIK, L.J.; SKJEMSTAD, J.O. **Visible, Near infrared, Mid infrared or Combined Diffuse Reflectance Spectroscopy for Simultaneous Assessment of Various Soil Properties.** *Geoderma* 131 (2006) 59–75, 2006.

VISCARRA ROSSEL, R.A.; CATTLE, S.; ORTEGA, A.; FOUAD, Y. **In situ Measurements of Soil Colour, Mineral Composition and Clay Content by vis-NIR Spectroscopy.** *Geoderma* 150, 253–266, 2009.

VISCARRA ROSSEL, R. A. & MCBRATNEY, A. B. **Laboratory Evaluation of a Proximal Sensing Technique for Simultaneous Measurement of Soil Clay and Water Content.** *Geoderma*, 85, 19–39, 1998.

VISCARRA ROSSEL, R.A.; T. BEHRENS. **Using Data Mining to Model and Interpret Soil Diffuse Reflectance Spectra.** Published by Elsevier B.V. All rights reserved. doi:10.1016/j.geoderma.2010.12.025, 2010.

VISCARRA ROSSEL R.A; MOUAZEN A. M; WETTERLIND J. **Visible and near infrared spectroscopy in soil science.** In Donald L. Sparks, editor: *Advances in Agronomy*,

Vol. 107, Burlington: Academic Press, 2010, pp. 163-215 WEBSTER, R. & BUTLER, B.E. 1976. Soil Classification and Survey Studies at Ginninderra. Australian Journal of Soil Research, 14, 1–24, 2010.

VISCARRA ROSSEL, R. A; CHAPPELLA, P. DE CARITATB & N. J. MCKENZIEA. **On the soil information content of visible–near infrared reflectance spectra.** *European Journal of Soil Science*, **62**, 442–453. doi: 10.1111/j.1365-2389.2011.01372.x, 2011.

VISCARRA ROSSEL R.A.; MCBRATNEY A.B. **Diffuse reflectance spectroscopy as a tool for digital soil mapping.** In: HARTEMINK A.E, MCBRATNEY A.B ; MENDONÇA S. L. (Ed.). **Digital soil mapping with limited data.** Amsterdam: Sringer, p.165-172; p. 381-396.,2008.

WALINGA, I. **Spectrophotometric determination of organic carbon in soil.** *Communications in Soil Science and Plant Analysis*, New York, v.23, p.1935-1944, 1992.

WASSERMAN, P.D. **Neural Computing: Theory and Practice**, Van Nostrand Reinhold, New York, 1989.

WATANABE , A.M, BESSA, L.P.D, MARTINS, T.G.M.,ET AL. **Por que fazer análise de Solo ?.** Curitiba: Universidade Federal do Paraná, Projeto Solo Planta, 2002.

WETTERLIND, J.; STENBERG, B.; AND SÖDERSTRÖM, M. **The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale.** *Precision agriculture*. Volume: 9, Number: 1-2, pp. 57-69, 2008b.

WIEDA, R.; BORGELT, S. T. **Geoestatistical Analysis of Plant Nutrients From Sample Nested Grids.** St. Joseph, ASAE Paper MCR93-131, 14p. 1993.

WITTEN H. I; FRANK E. **“Data Mining – Practical Machine Learning Tools and Techniques”**, Elsevier, 2005.