

PEDRO AUGUSTO DE OLIVEIRA MORAIS

**Previsão e Classificação Textural do Solo  
Através da Análise Multivariada de Imagens**

Goiânia

2016

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

**1. Identificação do material bibliográfico:**      ☒ **Dissertação**      ☐ **Tese**

**2. Identificação da Tese ou Dissertação**

Nome completo do autor: Pedro Augusto de Oliveira Moraes

Título do trabalho: Previsão e classificação textural do solo através da análise multivariada de imagens

**3. Informações de acesso ao documento:**

Concorda com a liberação total do documento ☒ SIM      ☐ NÃO<sup>1</sup>

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Pedro Augusto de O. Moraes  
Assinatura do (a) autor (a) <sup>2</sup>

Data: 22 / 07 / 19

<sup>1</sup> Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

<sup>2</sup>A assinatura deve ser escaneada.

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS (TEDE) NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

**1. Identificação do material bibliográfico:**      ☒ **Dissertação**      ☐ **Tese**

### 2. Identificação da Tese ou Dissertação

Autor (a):		PEDRO AUGUSTO DE OLIVEIRA MORAIS			
E-mail:		pedro_augusto_04@hotmail.com			
Seu e-mail pode ser disponibilizado na página? <input checked="" type="checkbox"/> Sim <input type="checkbox"/> Não					
Vínculo empregatício do autor					
Agência de fomento:		Coordenação de Aperfeiçoamento de Pessoal de Nível Superior		Sigla:	CAPES
País:	Brasil	UF:		CNPJ:	
Título:		PREVISÃO E CLASSIFICAÇÃO TEXTURAL DO SOLO ATRAVÉS DA ANÁLISE MULTIVARIADA DE IMAGENS			
Palavras-chave:		CLASSIFICAÇÃO TEXTURAL, IMAGEM DIGITAL, PLS, SPA-RLM			
Título em outra língua:		PREDICTION AND SOIL TEXTURE CLASSIFICATION BY MULTIVARIATE IMAGE ANALYSIS			
Palavras-chave em outra língua:		TEXTURE CLASSIFICATION, DIGITAL IMAGE, PLS, MLR-SPA			
Área de concentração:		Química			
Data defesa: (08/03/2016)					
Programa de Pós-Graduação:		Programa de Pós-Graduação em Química			
Orientador (a):		ANSELMO ELCANA DE OLIVEIRA			
E-mail:		elcana@ufg.br			
Co-orientador (a):*					
E-mail:					

\*Necessita do CPF quando não constar no SisPG

### 3. Informações de acesso ao documento:

Concorda com a liberação total do documento    ☐ SIM      ☒ NÃO<sup>1</sup>

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC da tese ou dissertação.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses e ou dissertações, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

\_\_\_\_\_  
Assinatura do (a) autor (a)

Data: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

<sup>1</sup> Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

**PEDRO AUGUSTO DE OLIVEIRA MORAIS**

**Previsão e Classificação Textural do Solo Através da  
Análise Multivariada de Imagens**

Dissertação apresentada ao Instituto de Química da Universidade Federal de Goiás para obtenção do título de Mestre em Química pelo Programa de Pós-graduação em Química.

Área de Concentração: **Química**

Orientador: Prof. Dr. Anselmo Elcana de Oliveira

Goiânia

2016

Ficha catalográfica elaborada automaticamente  
com os dados fornecidos pelo(a) autor(a), sob orientação do Sibi/UFG.

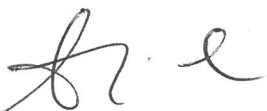
OLIVEIRA MORAIS, PEDRO AUGUSTO DE  
PREVISÃO E CLASSIFICAÇÃO TEXTURAL DO SOLO ATRAVÉS DA  
ANÁLISE MULTIVARIADA DE IMAGENS [manuscrito] / PEDRO  
AUGUSTO DE OLIVEIRA MORAIS. - 2016.  
CXC, 190 f.: il.

Orientador: Prof. Dr. ANSELMO ELCANA DE OLIVEIRA.  
Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto de  
Química (IQ) , Programa de Pós-Graduação em Química, Goiânia, 2016.  
Bibliografia. Apêndice.  
Inclui siglas, fotografias, abreviaturas, símbolos, gráfico, tabelas,  
lista de figuras, lista de tabelas.

1. CLASSIFICAÇÃO TEXTURAL. 2. IMAGEM DIGITAL. 3. PLS. 4.  
SPA - RLM. I. ELCANA DE OLIVEIRA, ANSELMO , orient. II. Título.

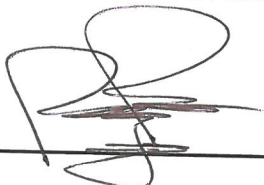
## Folha de Aprovação

Membros da Comissão Julgadora de Dissertação de Mestrado em Química, apresentada ao Instituto de Química da Universidade Federal de Goiás pelo discente Pedro Augusto de Oliveira Morais, em 08/03/2016:



---

Prof. Dr. Anselmo de Oliveira Morais (UFG)



---

Prof. Dr. Clarimar José Coelho (PUC-GO)



---

Prof. Dr. Nelson Roberto Antoniosi Filho (UFG)

*À minha família, amigos, e por mais estranho que pareça a mim mesmo, que nunca havia imaginado que chegaria até aqui.*

## Agradecimentos

Ao analista Diego Mendes de Souza do Centro Nacional de Pesquisa de Arroz e Feijão (CNPAF/EMBRAPA) por possibilitar a realização deste estudo, assim como a colaboração em cada etapa deste trabalho.

À pesquisadora Dra. Beata Eموke Madari do CNPAF pelo incentivo e colaboração.

Ao professor Dr. Anselmo Elcana de Oliveira pela orientação e paciência ao longo do curso.

Aos integrantes do laboratório de Análise Agroambiental (LAA/CNPAF), Silvio, Adilson, Roberto Carlos, Wesley, Ivã, Diego, Demétrio, Luismar, Marília, Renata, Jéssica, Anne, Thiago, Franciely, Beatriz e Rayssa por possibilitarem um ambiente de trabalho dinâmico e agradável.

Aos integrantes do Laboratório de Química Teórica e Computacional (LQTC/UFG), Vinicius, Eric, Gabriel, Leandro, Mariana e Deângelis, pelas discussões e descontrações.

Ao instituto de Química (IQ/UFG) pelo suporte.

Aos colegas, Kallita, Roger, Lucas, Danielly Cristina, Wanderson, Aleksander, Humberto, Thays, Danielly Rodrigues e Brunno pelo companherismo no decorrer do curso.

À Gabriela Matos pelo incentivo e apoio, principalmente nas etapas finais.

Aos meus amigos, Frederico, Ricardo, Eduardo, Danyell, Lázaro e Diogo, pela amizade e paciência.

À minha família, por sempre me depositar confiança, até quando eu duvidava de mim mesmo.

À Natália Bueno pela amizade e pelas longas discussões sobre a vida de um pós-graduando.

À Natália Dias pelo apoio e colaboração em várias traduções de texto.

À Veronica Lilja pela amizade, confiança e apoio, mesmo estando longe.

À Bárbara Diaz Pines Cabanes por me ensinar que nunca devemos desistir do que queremos.



*“O medo corta mais profundamente do que a espada.”*

*(George R. R. Martin)*

## Resumo

Morais, Pedro Augusto de Oliveira. **Previsão e Classificação Textural do Solo Através da Análise Multivariada de Imagens**. 2016. 190 f. Dissertação (Mestrado em Química) – Instituto de Química, Universidade Federal de Goiás, Goiânia, 2016.

A textura, ou granulometria, do solo é definida pela distribuição quantitativa das partículas minerais menores que 2 mm: areia, argila e silte. Esses indicadores físicos possibilitam a classificação do solo e orientam o manejo, a irrigação e a adição de insumos agrícolas. As metodologias usuais para análise textural são laboriosas, destrutivas, utilizam agentes químicos oxidantes e essa análise é bastante requisitada nos laboratórios de análise da fertilidade do solo. Logo, são imprescindíveis investigações e o desenvolvimento de metodologias alternativas que sejam operacionais e limpas. Nessa direção, esse estudo propõe a utilização de análise multivariada de imagens digitais para previsão da textura do solo e classificação. Para tanto, 60 amostras de solo diversificadas foram consideradas para análise textural pelo método da pipeta e para obtenção de imagens digitais no sistema de cor RGB (*Red*, *Green*, *Blue*) em formato Tiff. A correlação entre as imagens digitais e os teores de areia, argila e silte foi feita por Regressão por Quadrados Mínimos Parciais (PLS) e por Regressão Linear Múltipla associada com o Algoritmo das Projeções Sucessivas (SPA-RLM). Os melhores modelos apresentaram um índice de acerto de 100%. Portanto, a predição textural do solo através de imagens é uma técnica promissora por ser limpa, barata e operacional.

Palavras-chaves: Classificação textural, imagem digital, PLS, SPA-RLM.

## Abstract

Morais, Pedro Augusto de Oliveira. **Prediction and soil texture classification by multivariate image analysis**. 2016. 190 p. Dissertation (Master of Chemistry) – Institute of Chemistry, Federal University of Goiás, Goiânia, 2016.

The texture or grain size of the surface are defined by the quantitative distribution of the mineral particles smaller than 2 mm: sand, clay and silt. These physical indicators enable soil classification and guide the management, irrigation and the addition of agricultural inputs. Although the usual methods for textural analysis are laborious and destructive, using chemical oxidizing agents, this kind of analysis is quite required in soil fertility laboratories. Therefore, it is essential to research and develop alternative methodologies that are operational and clean. In this way, this study proposes the use of multivariate analysis of digital images to predict and classify soil texture. For this purpose, 60 samples of diverse soil were considered to textural analysis by the pipette method and for obtaining digital images in color system RGB (Red, Green, Blue) in Tiff format. The correlation between digital images and the percentage of sand, clay and silt is made by Partial Least Squares Regression (PLS) and Multiple Linear Regression algorithm associated with the Successive Projections (SPA-MLR). The best models had a 100 % success rate. Therefore, the prediction texture soil through images is a promising technique to be clean, inexpensive and operational.

Keywords: textural classification, digital image, PLS, MLR-SPA.

## Lista de figuras

Figura 1 – Diagrama de fases triangular de classificação textural do solo. . . . .	38
Figura 2 – Domínio da função $f(x, y)$ para uma imagem monocromática. . . . .	41
Figura 3 – Exemplo de histograma de uma imagem dispondo somente oito níveis de cinza. . . . .	42
Figura 4 – Espectro eletromagnético. . . . .	42
Figura 5 – Modelo de cores RGB. . . . .	45
Figura 6 – Modelo de cores HSI. . . . .	46
Figura 7 – Principais etapas do sistema de visão computacional. . . . .	47
Figura 8 – Representação da RLM. . . . .	54
Figura 9 – Decomposição da matriz $\mathbf{X}$ e do vetor $\mathbf{y}$ em LVs. . . . .	55
Figura 10 – Codificação de um histograma de frequência de uma imagem na con- formação de um cromossomo. De tal forma, que cada tonalidade de cor refere-se a um gene do cromossomo. . . . .	59
Figura 11 – Trado tipo holandês utilizado na coleta de amostras. . . . .	62
Figura 12 – Processador ultra-sônico utilizado na dispersão física das amostras. . .	64
Figura 13 – Lupa binocular Leica EZ4 D utilizada para a aquisição das imagens de solo. . . . .	65
Figura 14 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração ( $\bullet$ ) e validação ( $\blacktriangleleft$ ), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	69
Figura 15 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcenta- gem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração ( $\bullet$ ) e validação ( $\blacktriangleleft$ ), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	71
Figura 16 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcenta- gem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração ( $\bullet$ ) e validação ( $\blacktriangleleft$ ), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6, 3x. . . . .	76

Figura 17 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	81
Figura 18 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	81
Figura 19 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	87
Figura 20 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	94
Figura 21 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	94
Figura 22 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	101
Figura 23 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	103
Figura 24 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	105

Figura 25 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x. . . . .	109
Figura 26 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	127
Figura 27 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	128
Figura 28 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	129
Figura 29 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	130
Figura 30 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	131
Figura 31 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	132
Figura 32 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	133

Figura 33 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	134
Figura 34 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB. . . . .	135
Figura 35 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza. . . . .	136
Figura 36 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza. . . . .	136
Figura 37 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI. . . . .	137
Figura 38 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza. . . . .	137
Figura 39 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI. . . . .	138
Figura 40 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza. . . . .	138
Figura 41 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB. . . . .	139
Figura 42 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza. . . . .	140
Figura 43 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza. . . . .	140
Figura 44 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI. . . . .	141
Figura 45 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza. . . . .	141
Figura 46 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI. . . . .	142
Figura 47 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza. . . . .	142

Figura 48 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	143
Figura 49 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	144
Figura 50 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	145
Figura 51 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	146
Figura 52 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	147
Figura 53 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	148
Figura 54 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	149



Figura 55 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	150
Figura 56 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	151
Figura 57 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	152
Figura 58 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	153
Figura 59 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	154
Figura 60 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB. . . . .	155
Figura 61 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza. . . . .	156
Figura 62 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza. . . . .	156
Figura 63 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI. . . . .	157
Figura 64 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza. . . . .	157

Figura 65 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI. . . . .	158
Figura 66 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza. . . . .	158
Figura 67 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB. . . . .	159
Figura 68 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza. . . . .	160
Figura 69 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza. . . . .	160
Figura 70 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI. . . . .	161
Figura 71 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza. . . . .	161
Figura 72 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI. . . . .	162
Figura 73 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza. . . . .	162
Figura 74 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	163
Figura 75 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	164
Figura 76 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	165

Figura 77 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	166
Figura 78 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 6,3x. .	167
Figura 79 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 10,0x.	168
Figura 80 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	169
Figura 81 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	170
Figura 82 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 6,3x. . . . .	171
Figura 83 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 10,0x. . . . .	172
Figura 84 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	173

Figura 85 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	174
Figura 86 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	175
Figura 87 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	176
Figura 88 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	177
Figura 89 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	178
Figura 90 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	179
Figura 91 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	180

Figura 92 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . .	181
Figura 93 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . .	182
Figura 94 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	183
Figura 95 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	184
Figura 96 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	185
Figura 97 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	186
Figura 98 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x. . . . .	187

Figura 99 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x. . . . .	188
Figura 100 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x. . . . .	189
Figura 101 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x. . . . .	190

## Lista de tabelas

Tabela 1 – Relação de temperatura da água durante a análise e o tempo de sedimentação . . . . .	64
Tabela 2 – Resultados da análise granulométrica das amostras de solo . . . . .	68
Tabela 3 – Dados dos modelos PLS na predição dos teores de argila em amostras de solo, utilizando o algoritmo SPXY . . . . .	70
Tabela 4 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras de solo, utilizando o algoritmo SPXY . . . . .	72
Tabela 5 – Dados dos modelos PLS na predição dos teores de argila em amostras de solo, utilizando o particionamento igualmente espaçado . . . . .	73
Tabela 6 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras de solo, utilizando o particionamento igualmente espaçado . .	74
Tabela 7 – Parâmetros elegidos para a aplicação do algoritmo genético . . . . .	75
Tabela 8 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras de solo, utilizando o algoritmo SPXY . . . . .	77
Tabela 9 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras de solo, utilizando o particionamento igualmente espaçado . . . . .	78
Tabela 10 – Dados dos modelos PLS na predição dos teores de areia em amostras de solo, utilizando o algoritmo SPXY . . . . .	80
Tabela 11 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras de solo, utilizando o algoritmo SPXY . . . . .	82
Tabela 12 – Dados dos modelos PLS na predição dos teores de areia em amostras de solo, utilizando o particionamento igualmente espaçado . . . . .	84
Tabela 13 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras de solo, utilizando o particionamento igualmente espaçado . .	85
Tabela 14 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras de solo, utilizando o algoritmo SPXY . . . . .	86
Tabela 15 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras de solo, utilizando o particionamento igualmente espaçado . . . . .	88
Tabela 16 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o algoritmo SPXY . . . . .	89

Tabela 17 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado . . . . .	90
Tabela 18 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o algoritmo SPXY . . . . .	90
Tabela 19 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado . . . . .	91
Tabela 20 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o algoritmo SPXY . . . . .	91
Tabela 21 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o particionamento igualmente espaçado . . . . .	92
Tabela 22 – Dados dos modelos PLS na predição dos teores de argila em amostras argilosas, utilizando o algoritmo SPXY . . . . .	93
Tabela 23 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras argilosas, utilizando o algoritmo SPXY . . . . .	95
Tabela 24 – Dados dos modelos PLS na predição dos teores de argila em amostras argilosas, utilizando o particionamento igualmente espaçado . . . . .	97
Tabela 25 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras argilosas, utilizando o particionamento igualmente espaçado . . . . .	98
Tabela 26 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras argilosas, utilizando o algoritmo SPXY . . . . .	100
Tabela 27 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras argilosas, utilizando o particionamento igualmente espaçado . . . . .	102
Tabela 28 – Dados dos modelos PLS na predição dos teores de areia em amostras argilosas, utilizando o algoritmo SPXY . . . . .	104
Tabela 29 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras argilosas, utilizando o algoritmo SPXY . . . . .	106
Tabela 30 – Dados dos modelos PLS na predição dos teores de areia em amostras argilosas, utilizando o particionamento igualmente espaçado . . . . .	107
Tabela 31 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras argilosas, utilizando o particionamento igualmente espaçado . . . . .	108
Tabela 32 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras argilosas, utilizando o algoritmo SPXY . . . . .	110



Tabela 33 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras argilosas, utilizando o particionamento igualmente espaçado . . . . .	112
Tabela 34 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o algoritmo SPXY . . . . .	113
Tabela 35 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o algoritmo SPXY . . . . .	113
Tabela 36 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o algoritmo SPXY . . . . .	114
Tabela 37 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado . . . . .	114
Tabela 38 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o particionamento igualmente espaçado . . . . .	115
Tabela 39 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado . . . . .	115

## Lista de abreviaturas e siglas

CMY	<i>Cyan, Magenta, Yellow</i> , ou Ciano, Magenta e Amarelo
CV	<i>Cross Validation</i> , ou Validação Cruzada
DIP	<i>Digital Image Processing</i> , ou Processamento de Imagens Digitais
DPI	<i>Dots per Inch</i> , ou Pontos Por Polegada
GA	<i>Genetic Algorithm</i> , ou Algoritmo Genético
HIS	<i>Hue, Intensity, Saturation</i>
HSI	<i>Hue, Saturation, Intensity</i> , ou Matiz, Saturação e Intensidade
HSV	<i>Hue, Saturation, Value</i> , ou Matiz, Saturação e Valor
KS	Algoritmo Kennard-Stone
LAA	Laboratório de Análise Agroambiental
LV	Variável Latente
MIA	<i>Multivariate Image Analysis</i> , ou Análise Multivariada de Imagens
OSC	<i>Orthogonal Signal Correction</i> , ou Correção de sinal ortogonal
PCA	<i>Principal Analysis Component</i> , ou Análise por Componentes Principais
PLS	<i>Partial Least Square</i> , ou Mínimos Quadrados Parciais
RLM	<i>Multiple Linear Regression</i> , ou Regressão Linear Múltipla
RGB	<i>Red, Green, Blue</i> ou Vermelho, Verde e Azul
RMSE	<i>Root Mean Square Error</i> , ou Raiz Quadrada do Erro Médio Quadrático
RMSEC	<i>Root Mean Square Error of Calibration</i> , ou Raiz Quadrada do Erro Médio Quadrático de Calibração
RMSECV	<i>Root Mean Square Error of Cross-Validation</i> , ou Raiz Quadrada do Erro Médio Quadrático da Validação Cruzada

RMSEP	<i>Root Mean Square Error of Predition</i> , ou Raiz Quadrada do Erro Médio Quadrático de Previsão
RMSEV	<i>Root Mean Square Error of Validation</i> , ou Raiz Quadrada do Erro Médio Quadrático de Validação
ROI	<i>Region of Interest</i> , ou Região de Interesse
RPD	<i>Residual Prediction Deviation</i> , ou Desvio Residual de Previsão
SEP	<i>Standard Error of Prediction</i> , ou Erro Padrão de Revisão
SPA	<i>Sucessive Projection Algorithm</i> , ou Algoritmo das Projeções Sucessivas
SPXY	<i>Sample set Partitioning based on joint x-y distances</i>
SVC	Sistema de Visão computacional
TIFF	<i>Tagged Image File Format</i>
YIQ	<i>Luminance, In-phase, Quadrature</i>

## Sumário

<b>1</b>	<b>Introdução . . . . .</b>	<b>30</b>
<b>2</b>	<b>Objetivos . . . . .</b>	<b>33</b>
<b>2.1</b>	<b>Objetivos específicos . . . . .</b>	<b>33</b>
<b>3</b>	<b>Revisão Bibliográfica . . . . .</b>	<b>34</b>
<b>4</b>	<b>Metodologia . . . . .</b>	<b>38</b>
<b>4.1</b>	<b>A textura do solo . . . . .</b>	<b>38</b>
<b>4.1.1</b>	<b>Determinação da textura em amostras de solo . . . . .</b>	<b>39</b>
<b>4.2</b>	<b>MIA . . . . .</b>	<b>40</b>
<b>4.2.1</b>	<b>Imagens Digitais . . . . .</b>	<b>40</b>
<b>4.2.2</b>	<b>Cores . . . . .</b>	<b>42</b>
<b>4.2.3</b>	<b>Modelos de cores . . . . .</b>	<b>43</b>
4.2.3.1	Modelo RGB . . . . .	44
4.2.3.2	Modelo HSI . . . . .	45
<b>4.3</b>	<b>DIP . . . . .</b>	<b>46</b>
<b>4.3.1</b>	<b>Aquisição . . . . .</b>	<b>47</b>
<b>4.3.2</b>	<b>Pré-processamento . . . . .</b>	<b>47</b>
<b>4.3.3</b>	<b>Segmentação . . . . .</b>	<b>48</b>
<b>4.3.4</b>	<b>Extração de Características . . . . .</b>	<b>48</b>
<b>4.3.5</b>	<b>Reconhecimento e Interpretação . . . . .</b>	<b>48</b>
<b>4.3.6</b>	<b>Base de Conhecimento . . . . .</b>	<b>48</b>
<b>4.4</b>	<b>Quimiometria . . . . .</b>	<b>49</b>
<b>4.4.1</b>	<b>Organização dos dados . . . . .</b>	<b>50</b>
<b>4.4.2</b>	<b>Pré-processamento de dados . . . . .</b>	<b>50</b>
4.4.2.1	Centrar na Média . . . . .	51
4.4.2.2	OSC . . . . .	51
<b>4.5</b>	<b>Calibração Multivariada . . . . .</b>	<b>51</b>
<b>4.5.1</b>	<b>Algoritmo Kennard-Stone (KS) . . . . .</b>	<b>52</b>
<b>4.5.2</b>	<b>Algoritmo SPXY . . . . .</b>	<b>53</b>

4.5.3	RLM . . . . .	53
4.5.4	PLS . . . . .	54
4.5.5	Seleção de LVs . . . . .	55
4.5.6	Detecção de <i>outliers</i> . . . . .	56
4.5.7	Avaliação dos modelos de calibração multivariada através de métodos estatísticos . . . . .	57
4.5.8	Seleção das variáveis . . . . .	58
4.5.8.1	GA . . . . .	59
4.5.8.1.1	Codificação das variáveis . . . . .	59
4.5.8.1.2	Criação da população original . . . . .	60
4.5.8.1.3	Avaliação das respostas . . . . .	60
4.5.8.1.4	Cruzamentos . . . . .	60
4.5.8.1.5	Mutações . . . . .	60
4.5.8.2	SPA . . . . .	61
5	Parte experimental . . . . .	62
5.1	Amostragem . . . . .	62
5.1.1	Coleta de amostras . . . . .	62
5.1.2	Preparo das amostras . . . . .	62
5.2	Metodologia para análise textural do solo . . . . .	62
5.3	Aquisição das imagens . . . . .	64
5.4	Manipulação e tratamento das imagens . . . . .	65
6	Resultados e Discussões . . . . .	67
6.1	Análise textural do solo: método da pipeta . . . . .	67
6.2	MIA . . . . .	67
6.2.1	Determinação do teor de argila por calibração multivariada	67
6.2.2	Determinação do teor de areia por calibração multivariada	79
6.2.3	Classificação textural das amostras de solo por análise de imagens . . . . .	89
6.2.4	Determinação do teor de argila por calibração multiva- riada em amostras argilosas . . . . .	92

6.2.5	Determinação do teor de areia por calibração multivari- ada em amostras argilosas . . . . .	103
6.2.6	Classificação textural das amostras de solo por análise de imagens . . . . .	111
7	Conclusão . . . . .	116
	Referências <sup>1</sup> . . . . .	118
	Apêndice A–Modelos de regressão por mínimos quadrados PLS para predição dos teores de argila em amostras de solo	127
	Apêndice B–Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de argila em amostras de solo . . . . .	131
	Apêndice C–Relatório de desempenho do algo- ritmo genético aplicado a modelos PLS referentes as imagens com 6,3x de aumento óptico . . . . .	135
	Apêndice D–Relatório de desempenho do algo- ritmo genético aplicado a modelos PLS referentes as imagens com 10x de aumento óptico . . . . .	139
	Apêndice E–Modelos de regressão GA-PLS para predição dos teores de argila em amos- tras de solo . . . . .	143

<sup>1</sup> De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

Apêndice F	Modelos de regressão por mínimos quadrados PLS para predição dos teores de areia em amostras de solo	147
Apêndice G	Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de areia em amostras de solo . . . . .	151
Apêndice H	Relatório de desempenho do algoritmo genético aplicado a modelos PLS referentes as imagens com 6,3x de aumento óptico . . . . .	155
Apêndice I	Relatório de desempenho do algoritmo genético aplicado a modelos PLS referentes as imagens com 10x de aumento óptico . . . . .	159
Apêndice J	Modelos de regressão GA-PLS para predição dos teores de areia em amostras de solo . . . . .	163
Apêndice K	Modelos de regressão por mínimos quadrados PLS para predição dos teores de argila em amostras argilosas . . . . .	167
Apêndice L	Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de argila em amostras argilosas . . . . .	171

Apêndice M–Modelos de regressão GA-PLS para predição dos teores de argila em amostras argilosas . . . . .	175
Apêndice N–Modelos de regressão por mínimos quadrados PLS para predição dos teores de areia em amostras argilosas	179
Apêndice O–Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de areia em amostras ar- gilosas . . . . .	183
Apêndice P–Modelos de regressão GA-PLS para predição dos teores de areia em amostras argilosas . . . . .	187



# 1 Introdução

O Brasil detém atualmente uma das primeiras posições na exportação de produtos agrícolas (Estado de Minas, 2015), com uma produção de grãos estimada para safra de 2015/2016 de 211 milhões de toneladas, e uma área plantada prevista de 58,6 milhões de hectares. Ou seja, um aumento de aproximadamente 1,5 % e 1,1 % respectivamente, da produção de grãos e da área destinada a plantio, quando comparados com a safra passada (CONAB, 2015). Nesse contexto, observando a produtividade brasileira ao longo do período de 1975 a 2012, nota-se um incremento de produção anual médio de 3,5 %, que por sua vez, é superior a média mundial de 1,8 % (GASQUES et al., 2014). Nesse período, em especial a partir do ano de 2000, observaram-se os maiores índices de crescimento da produção. No período de 1999 a 2009, por exemplo, embora a cada ano tenha ocorrido um crescimento de 1,7% da área plantada, a produção cresceu 4,7% (LANDIM, 2010; FREITAS; MENDONÇA; LOPES, 2014). Esses valores são reflexos da potencialização da produção nas lavouras, resultado do desenvolvimento da tecnologia agrícola nacional, atualmente uma das mais avançadas do mundo (CORRÊA, 2009; Estado de Minas, 2015).

Considerando que os gastos com insumos podem representar mais que 25% do custo total da produção agrícola e que as análises físico-químicas do solo são ferramentas eficazes para a racionalização de fertilizantes e corretivos, essas análises possuem um papel importante na avaliação da fertilidade do solo (SANTOS et al., 2009). Elas fornecem informações para o uso e o manejo dos solos, para melhorar a utilização da água em sistemas produtivos, nos processos de calagem, gessagem, fosfatagem, adubações técnicas de conservação do solo, no desenvolvimento e rendimento das culturas agrícolas e na avaliação da dinâmica dos nutrientes e poluentes do solo (CORÁ et al., 2009; MAURI et al., 2011; RUIZ, 2005).

Geralmente, a avaliação da fertilidade do solo abrange 16 parâmetros:

- (i) Químicos: Fe, Zn, Cu, Mn, Ca, Mg, K, Al, acidez potencial, pH, matéria orgânica, carbono total (Ct) e nitrogênio total (Nt);
- (ii) Físicos (textura): areia, argila e silte.

A determinação dos parâmetros físicos do solo é viabilizada pela combinação de dois processos de dispersão: o químico e o mecânico. O tratamento químico possibilita a eliminação da matéria orgânica da amostra por oxidação e facilita a desagregação do solo (MAURI et al., 2011; CORÁ et al., 2009). Já o processo mecânico promove a dissociação

total das partículas, o que torna possível a separação da areia por peneiramento, seguida de sua determinação por gravimetria. Em seguida, a sedimentação sequencial do silte e da argila viabiliza a quantificação desses dois parâmetros físicos. Essa determinação pode ser realizada por pipetagem seguida de gravimetria ou densimetria (TAVARES FILHO; MAGALHAES, 2008; PANSU; GAUTHEYROU, 2003).

Diante do exposto, é crucial que a análise de solo seja eficiente e precisa, pois dela advêm inúmeras informações que, como já mencionadas, são aproveitadas na agricultura. Porém, na realização de algumas análises de solo são empregadas metodologias laboriosas, que além de utilizarem agentes químicos oxidantes, geram resultados tardios, com tempo de espera em média de dois dias, e nem sempre confiáveis (CORÁ et al., 2009; MAURI et al., 2011).

Como exemplo, a análise textural do solo, apesar de não produzir resíduos tóxicos e ser de baixo custo, em média de R\$ 10 reais por amostra, possui baixa produtividade. O Laboratório de Análise Agroambiental (LAA) da Embrapa Arroz e Feijão, localizado na Rodovia GO-462, Km 12, Santo Antônio de Goiás - GO, por exemplo, consegue realizar essa análise em apenas 30 amostras diariamente. Entretanto, somente no ano de 2015, esse laboratório recebeu mais de 15 mil amostras. Dessa forma, não seria possível atender a essas solicitações, caso em todas elas fossem requisitadas análise granulométrica.

Mais ainda, a associação da expansão da área plantada (FREITAS; MENDONÇA; LOPES, 2014), do incentivo governamental (PORTAL PLANALTO, 2015), e de deficiências nutricionais do solo implicou num aumento considerável da demanda por análises do solo. Logo, reforça-se a contínua necessidade de desenvolvimento de novos processos e tecnologias com as seguintes características: rapidez, operacionalidade e baixo impacto ambiental para reduzir a produção de resíduos e efluentes tóxicos (CORÁ et al., 2009; MAURI et al., 2011; PRADO, 2003; LENARDÃO et al., 2003).

Nesse sentido, as metodologias que utilizam imagens digitais vêm ganhando destaque em diversas áreas do conhecimento (BARBOSA, 2000) porque se apresentam como métodos limpos, rápidos, não destrutivos e de fácil processamento do diversificado conjunto de informações presentes em uma imagem (GOMES, 2001; GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

A análise multivariada de imagens (*Multivariate Image Analysis* - MIA) se constitui basicamente de três etapas: (i) o pré-processamento, (ii) a extração das informações, e (iii) a interpretação, ou calibração multivariada, a partir dos dados extraídos. A primeira

etapa consisti na otimização da qualidade da imagem como, por exemplo, o ajuste da luminosidade ou realce da cor através de filtros de imagem. A segunda é a conversão matemática da imagem em um conjunto de dados. Já a última etapa se apresenta como a utilização de ferramentas estatísticas para inferir sobre as amostras e fenômenos associados, ou para construir modelos multivariados para predição de parâmetros da amostra (GELADI; GRAHN, 1996; GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999; PETROU; PETROU, 2010; PRATS-MONTALBÁN; JUAN; FERRER, 2011). Obviamente, essa diversidade de possibilidades matemáticas sobre as informações imagéticas estão sendo viabilizadas graças aos avanços dos recursos computacionais (GOMES, 2001).

Neste trabalho são propostas metodologias para determinação da textura do solo através de imagens digitais combinadas a métodos quimiométricos de calibração multivariada. Em seguida, essas metodologias serão avaliadas e comparadas com os métodos clássicos quanto à precisão, exatidão, operacionalidade, geração de resíduos e custo.

## 2 Objetivos

Neste trabalho são propostas metodologias para determinação da textura do solo através de imagens digitais.

### 2.1 Objetivos específicos

Desenvolver modelos de calibração multivariados com o intuito de quantificar os constituintes texturais do solo: argila, silte, e areia.

Classificar os solos de acordo com os teores de argila, silte e areia preditos pelos modelos, e determinar os índices de acertos em relação a classificação dada pelo método de referência.

Avaliar as metodologias de análise de imagens propostas e compará-las com os métodos clássicos quanto à: precisão, operacionalidade, geração de resíduo, e custo.

### 3 Revisão Bibliográfica

A visão, um dos sentidos mais explorados pelo homem, possibilitou e possibilita ao ser humano estabelecer uma forte relação com o universo em que vive. Nesse sentido, a imagem, que por sua vez também corresponde a uma percepção visual, tem um papel bastante evidente no desenvolvimento humano desde o surgimento da humanidade, principalmente na comunicação, por exemplo, em representações artísticas, como pinturas e esculturas, e também em meios digitais, como a televisão e a internet (RAMOS, 2007). Além disso, as imagens desempenham um papel singular na construção do conhecimento, já que podem ser armazenadas, exploradas e transformadas nas nossas próprias mentes (COSTA, 2005).

Nesse contexto, nota-se a importância da imagem na sociedade, que além de ser amplamente utilizada na comunicação e na formação de saberes, ainda é utilizada no meio científico, por exemplo, no registro de evidências, e a partir do surgimento e desenvolvimento dos recursos computacionais, as imagens adquiriram aplicações em vários campos como em química, geografia, medicina e agricultura (GELADI et al., 1992; PRATS-MONTALBÁN; JUAN; FERRER, 2011).

As primeiras análises utilizando imagens datam do início da década de 60, por exemplo, Ledley (1964) propôs uma metodologia de análise de cromossomos através de imagens, muito mais rápida e padronizada do que os métodos convencionais utilizados no mesmo período.

Em medicina, Botvinick et al. (1976) propuseram a primeira técnica não invasiva de detecção de falsos aneurismas no ventricular esquerdo empregando a ventriculografia radioisotópica, que por sua vez, consiste na aquisição de imagens sequenciais que permitem mensurar a quantidade de batimentos cardíacos.

De modo semelhante, Strauss et al. (1971) propuseram um método de medição da fração de ejeção do ventrículo esquerdo sem a utilização de cateterismo cardíaco. Primeiramente, injeta-se um marcador por via intravenosa, em seguida, utilizando uma câmera de cintilação e um portão eletrônico acionado pelo eletrocardiograma obtiveram imagens do coração do paciente nas condições de relaxamento e contração. A partir das imagens obtidas, determinaram-se a fração de ejeção do ventrículo esquerdo calculando-se as áreas de comprimento do eixo do contorno ventricular nas fases finais da sístole e diástole.

Tal procedimento foi realizado em 20 pacientes, os resultados obtidos demonstraram boas correlações com os resultados fornecidos por cineangiocoronariografia.

O desenvolvimento do campo de análise de imagens, possibilitou sua aplicação na indústria. Goldstein e Nagler (1988) desenvolveram um método de inspeção, através da análise de imagens de peças metálicas em uma linha de produção, com uma eficiência de 300 inspeções por minuto. As peças foram avaliadas de acordo com vários tipos de defeitos previamente definidos, como defeitos estruturais e defeitos de textura (pintura, recozimento). Tal processo é realizado utilizando alguns algoritmos que são implementados através de um hardware de processamento de imagens.

Aliado ao desenvolvimento da análise de imagens, as metodologias de processamento e extração de informações de imagens, como as técnicas de segmentação, também expandiram-se (BILLINGSLEY, 1973). Algumas dessas técnicas utilizam unicamente os histogramas de frequência na escala cinza (PAL; PAL, 1993), viabilizando uma perspectiva geral sobre a imagem (RISEMAN; ARBIB, 1977).

Utilizando histogramas de frequência de cor, Geladi, Wold e Esbensen (1986) realizaram regressões multivariadas que correlacionavam os histogramas de imagens de microscópicos com alguma propriedade em massa de um material.

Ainda utilizando histogramas, Pataca (2006) desenvolveu um método de quantificação do índice de cor do mel, utilizando *scanner* de mesa. Após o processamento das imagens, obteve-se os histogramas de frequência de cores, que foram utilizados na construção do modelo de calibração por mínimos quadrados. Os resultados obtidos indicam que essa metodologia é praticável nessa aplicação.

A análise de imagens também foi empregada na análise de bebidas, por exemplo, na quantificação de *Lactobacillus* em leite fermentado (BORIN et al., 2007), na classificação de chás não processados (DINIZ et al., 2012), na classificação de refrigerantes dos tipos: Cola, Laranja e Guaraná, utilizando a análise de componentes principais (PCA) (GODINHO et al., 2008), e na classificação de cervejas, através da cor, utilizando um *scanner* de mesa (SILVA; GODINHO; OLIVEIRA, 2011).

Em biologia, Almeida et al. (2014) desenvolveram um método de classificação de bactérias utilizando uma *webcam*. Em análises de grãos, a análise de imagens foi empregada na estimação de umidade em frutos de café (PINTO et al., 2003), assim como na identificação de danos mecânicos em sementes de soja (FLOR et al., 2004), e na classificação de sementes de mamona em relação ao tipo de cultivar (VILAR et al., 2015).

Em determinações analíticas, Gomes et al. (2008) desenvolveram uma metodologia de determinação de ácido ascórbico em medicamentos utilizando um *scanner* comercial. Ainda utilizando *scanner* Paciornik et al. (2006) quantificaram teores de mercúrio em peixes, e Damasceno et al. (2015) determinaram o pH em matrizes de água potável.

Em análises de solo, Diniz et al. (2006) correlacionaram a cor do solo, determinada por colorímetro e pela carta de Munsell, com vários atributos do solo: matéria orgânica, argila e ferro extraído. Os autores observaram as correlações ( $r^2$ ) de 0,72, 0,64, e 0,92, respectivamente para a matéria orgânica, argila e ferro extraído.

De maneira semelhante, Demattê et al. (2011) desenvolveram um modelo matemático que estima o teor de matéria orgânica do solo através da sua correlação com a sua respectiva cor, que por sua vez, foi determinada por colorímetro no sistema Munsell de cores. Os modelos construídos apresentaram  $R^2$  de 0,66 para amostras úmidas, e 0,56 para amostras secas.

Utilizando câmera digital Rossel, Fouad e Walter (2008) quantificaram os teores de carbono orgânico e ferro extraído do solo. Após a aquisição das imagens em formato RGB, extraíram-se os histogramas e calcularam os valores médios de cada canal de cor, em seguida utilizando o método dos mínimos quadrados (PLS) correlacionaram-se tais valores com os teores de carbono orgânico e de ferro. Os autores demonstraram que tal metodologia pode ser utilizada na predição de tais atributos do solo, com as vantagens de ser simples, rápido, preciso e não destrutivo.

Moritsuka et al. (2014) desenvolveram um método de quantificação de carbono e nitrogênio totais, e ferro extraído em solos japoneses, utilizando um leitor de cores, que por sua vez determina a cor do solo no espaço de cores CEILAB. Determinaram-se boas correlações entre os canais do modelo de cor e os atributos do solo, o que permitiu as suas respectivas estimações de forma rápida e prática.

Ainda em análises de solo, Murti e Satyanarayana (1971) proporam um método de determinação de matéria orgânica, ferro, titânio, manganês e argila em amostras de solo a partir das suas respectivas cores, baseadas no sistema de cores de Munsell. Os resultados indicaram uma alta correlação entre a componente matiz e os teores de ferro e titânio, e entre a componente valor com os teores de matéria orgânica, argila e manganês.

Ibáñez-Asensio et al. (2013) construíram modelos multivariados com o intuito de correlacionar as três componentes da cor: luminosidade, matiz e saturação com alguns parâmetros do solo: matéria orgânica, areia, argila, pH, nitrogênio, fósforo e ferro. Os

resultados obtidos indicaram que com a exceção do pH e do fósforo, todos os outros parâmetros do solo avaliados apresentaram alguma correlação significativa com algum dos três componentes da cor.

Este estudo propõe uma nova metodologia para determinação dos constituintes texturais do solo: argila, silte e areia, através de imagens obtidas por uma Lupa com câmera digital acoplada. Modelos de calibração multivariados foram construídos com o intuito de correlacionar esses atributos do solo com os histogramas de frequência de suas respectivas imagens. Após a predição dos teores de argila e areia, os solos serão classificados de acordo com o triângulo de Atterberg.



## 4 Metodologia

### 4.1 A textura do solo

A textura do solo é representada pela distribuição granulométrica de partículas minerais menores que 2 mm, divididas de acordo com seus diâmetros em três frações: areia (0,05 a 2mm), silte (0,002 a 0,05 mm) e argila (menor que 0,002 mm) (CORÁ et al., 2009).

A análise textural classifica o solo de acordo com os seus constituintes texturais. Para tanto, emprega-se o diagrama de fases triangular de Atterberg, Figura 1, proposto pelo Soil Survey Staff do Departamento de Agricultura dos Estados Unidos, e que foi aceito e modificado pela Sociedade Brasileira de Ciência do Solo (SOUZA; CELLIGOI, 2012). Essa classificação do solo orienta o seu manejo e a sua conservação, bem como os processos de calagem, gessagem, fosfatagem, adubações no desenvolvimento e rendimento das culturas agrícolas e na avaliação da dinâmica dos nutrientes e poluentes do solo (CORÁ et al., 2009; MAURI et al., 2011).

A textura também é associada à distribuição do diâmetro dos poros do solo, delimitando a área de contato entre as partículas sólidas e a água. Consequentemente, ela influencia nas propriedades de retenção de água (GONZÁLEZ; CORONADO; ACOSTA, 2007; KLEIN et al., 2010).

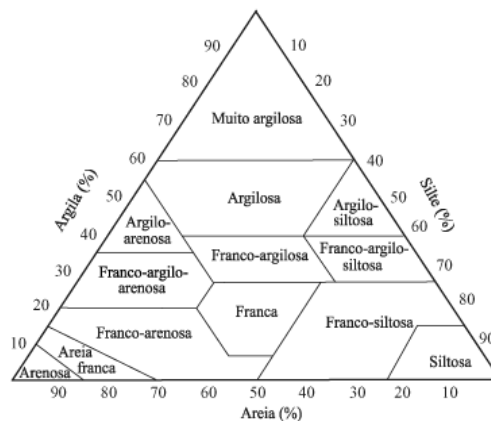


Figura 1 – Diagrama de fases triangular de classificação textural do solo.

Fonte: (SILVA et al., 2004)

#### 4.1.1 Determinação da textura em amostras de solo

A determinação da textura é realizada, normalmente, por meio de análises mecânicas que consistem em três etapas básicas: (i) dispersão química e física da amostra em meio aquoso, (ii) separação das frações por peneiramento e sedimentação e (iii) quantificação das frações texturais por gravimetria ou densitometria (ALEXANDRE; SILVA; FERREIRA, 2001; MAURI et al., 2011; TAVARES FILHO; MAGALHAES, 2008).

A primeira etapa dessa análise consiste na combinação do processo químico e da desagregação mecânica. O tratamento químico promove a elevação do potencial zeta, que está relacionado ao grau de repulsão ou atração das partículas. Quando esse potencial é baixo, a atração supera a repulsão e as partículas tendem a se agregar. De modo oposto, um potencial zeta alto proporcionará estabilidade, ou seja, a solução não sofrerá agregação devido à alta repulsão entre as partículas. A elevação desse potencial é realizada com o aumento do pH e pela redução de eletrólitos na solução (MAURI et al., 2011; RUIZ, 2005).

Além disso, o procedimento químico promove a precipitação de compostos de alumínio e cálcio, o que possibilita a remoção dos agentes cimentantes, flocculantes e alguns óxidos metálicos, e elimina a matéria orgânica por oxidação, sendo essa matéria orgânica uma das principais agentes agregadoras de argila (TAVARES FILHO; MAGALHAES, 2008; MAURI et al., 2011).

Já a dispersão física auxilia na quebra dos pequenos agregados e é realizada por intermédio da agitação lenta (DOURADO; SILVA; MARINHO, 2012), agitação rápida, por atenuação de raios gamas (VAZ; NAIME; MACEDO, 1997) ou por vibração, utilizando ondas mecânicas de ultrassom (GENRICH; BREMNER, 1972). Essa etapa é de fundamental importância, pois após seu emprego admite-se que as partículas minerais estejam totalmente desagregadas (MAURI et al., 2011; RUIZ, 2005; TAVARES FILHO; MAGALHAES, 2008).

A segunda e terceira etapas (ii, iii) compreendem a separação da areia por peneiramento e sua determinação por gravimetria. Em seguida, silte e argila são quantificados por sedimentação sequencial. Essa determinação pode ser realizada por densimetria, ou por pipetagem, seguida de gravimetria (GARCÍA; MEDINA; NÚÑEZ, 2008; TAVARES FILHO; MAGALHAES, 2008).

Apesar das análises mecânicas, como o método da pipeta e o método do hidrômetro, serem as mais utilizadas e, conseqüentemente, serem consideradas como métodos padrões, existem outras metodologias de análises texturais mais eficientes e precisas que associam a

sedimentometria à difração de raios X, difração de raios laser e sensores de pressão. Tais métodos também fazem uso das duas primeiras etapas citadas anteriormente, diferenciando-se, assim, o modo de quantificar a argila e o silte na sedimentação desses componentes (ALEXANDRE; SILVA; FERREIRA, 2001; GONZÁLEZ; CORONADO; ACOSTA, 2007; PANSU; GAUTHEYROU, 2003).

Dessa maneira, as metodologias usuais de determinação de argila, silte e areia são bastante difundidas por suas facilidades, operacionalidades e baixos custos. Elas são, entretanto, laboriosas, geram resultados tardios, possuem baixa produtividade e utilizam-se de agentes químicos oxidantes (CORÁ et al., 2009; MAURI et al., 2011).

## 4.2 MIA

### 4.2.1 Imagens Digitais

Uma imagem pode ser definida como uma representação de um objeto ou situação real (GELADI; GRAHN, 1996) e pode ser obtida através de diversos recursos como técnicas fotográficas, *scanners*, microscópicos, satélites de reconhecimento, equipamentos científicos e também por aparelhos de raios X (ALBUQUERGUE et al., 2004; RUSS; RUSS, 2008; RUSS, 2011).

No campo científico, uma imagem pode ser representada matematicamente como uma função  $f(x, y)$  que descreve a intensidade da luz em cada ponto de uma imagem 2D monocromática formada por duas dimensões: horizontal,  $x$ , e vertical,  $y$ . Essas dimensões  $x$  e  $y$  pertencem a intervalos reais contínuos, e o valor de  $f$  corresponde ao brilho ou ao nível de cinza referente em cada ponto. O domínio dessa função é ilustrado na Figura 2 (GELADI; GRAHN, 1996).

Em virtude da complexidade dessa função e da dificuldade em encontrá-la, o seu ajuste torna-se uma tarefa extremamente difícil. Dessa maneira, as imagens devem ser digitalizadas. Assim  $x$  se converte em  $i = 1, 2, 3, \dots, I$ , e  $y$  em  $j = 1, 2, 3, \dots, J$ . Dessa forma, a função  $f(x, y)$  torna-se  $g(i, j)$ , sendo que  $g$  fornece a média de intensidade em cada ponto  $(i, j)$  (GELADI; GRAHN, 1996).

Em outras palavras, o processo de digitalização transforma uma imagem contínua  $f(x, y)$  em uma série de matrizes bidimensionais. Dessa maneira, as coordenadas  $(i, j)$  de cada uma dessas matrizes identificam pontos na imagem e o valor do elemento da matriz

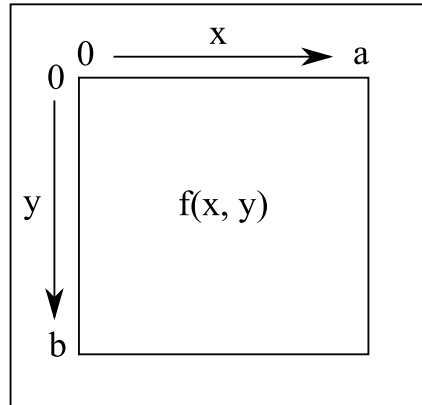


Figura 2 – Domínio da função  $f(x, y)$  para uma imagem monocromática.

corresponde ao nível de intensidade ou nível de cinza naquele ponto (GELADI; GRAHN, 1996).

Cada elemento da matriz digital é denominado elemento da imagem ou elemento da figura, comumente chamado de *pixel* ou *pel*, ambos derivados do termo *picture elements*. Após a digitalização, cada *pixel* adquire um valor inteiro pertencente ao intervalo  $[0, 2^n - 1]$ , correspondente à intensidade do nível de cinza naquele ponto. Assim, há uma relação de proporcionalidade entre os valores de  $n$  e os níveis de cinza; ou seja, um alto valor de  $n$  proporcionará uma larga faixa de níveis de cinza em uma imagem digital. Diante disso, é possível expressar graficamente a distribuição de frequências dos níveis de cinza em uma imagem, tal expressão gráfica é denominada histograma (GONZALEZ; WOODS, 2003).

O histograma de uma imagem representa a proporção ou a quantidade de *pixels* que manifestam um específico nível de cinza em uma determinada imagem. Em geral, esses dados são expostos em gráficos de barras, Figura 3, onde o eixo horizontal representa os vários níveis de cinza presentes em uma imagem, e o eixo vertical a frequência desses níveis. A análise de um histograma de uma imagem fornece várias informações sobre sua qualidade, contraste, e também sobre seu brilho (GELADI; GRAHN, 1996).

Da mesma forma que para uma imagem 2D monocromática, uma imagem 3D também pode ser descrita matematicamente e ainda ser digitalizada. Assim, uma função  $f(x, y, z)$  descreve tal imagem formada por três dimensões: (i)  $x$ , (ii)  $y$ , e (iii)  $z$ . Já a digitalização promove a transformação dos índices  $x$ ,  $y$  e  $z$  em respectivamente:  $i = 1, 2, 3, \dots, I$ ,  $j = 1, 2, 3, \dots, J$ , e  $h = 1, 2, 3, \dots, H$ . Dessa forma, a função de descrição torna-se  $g(i, j, h)$  e o elemento da matriz é denominado de *voxel* (GELADI; GRAHN, 1996).

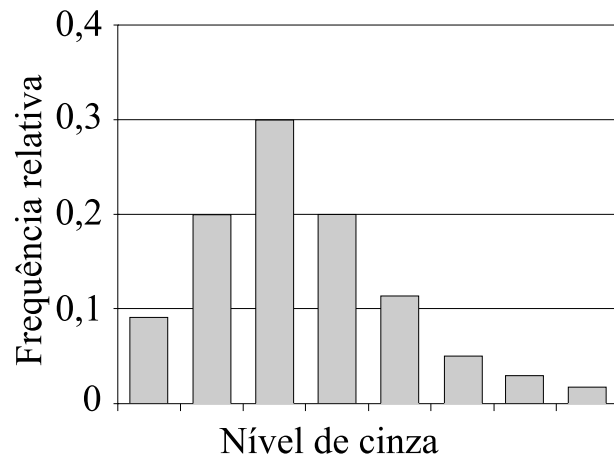


Figura 3 – Exemplo de histograma de uma imagem dispondo somente oito níveis de cinza.

Uma imagem colorida como já mencionado é formada por três dimensões, sendo cada dimensão composta por uma imagem monocromática vermelha, verde e azul. Assim, o produto dessa combinação é uma imagem RGB (*Red Green Blue*) (GELADI; GRAHN, 1996).

#### 4.2.2 Cores

A luz visível é um tipo de energia eletromagnética pertencente a uma estreita faixa do espectro eletromagnético, Figura 4, limitada pelas cores violeta, de menor comprimento de onda (400 nm), até a cor vermelha, de maior comprimento de onda (700 nm) (LU, 1997).

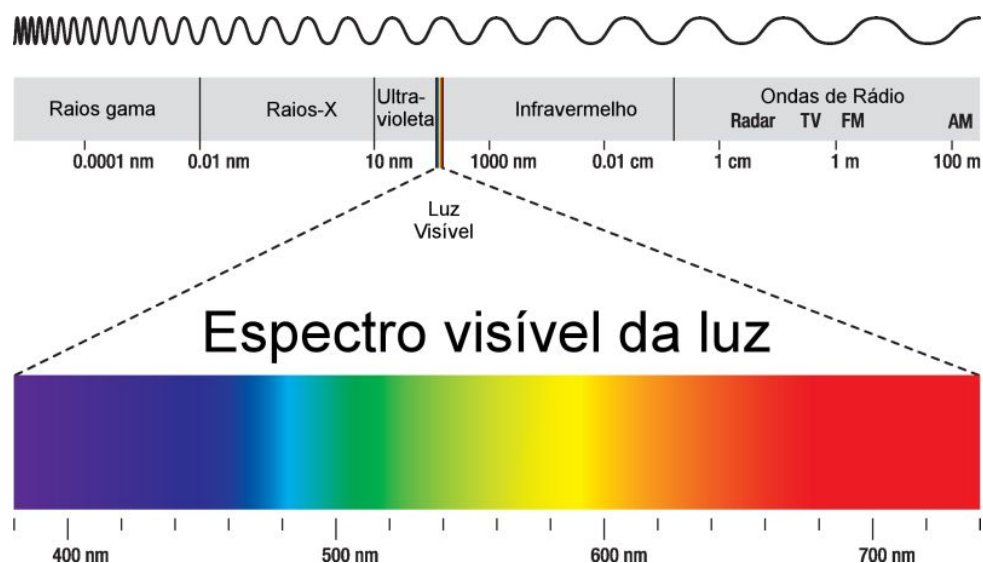


Figura 4 – Espectro eletromagnético.

Fonte: Info Escola <<http://www.infoescola.com/fisica/espectro-eletromagnetico/>>. Acesso em 23 dez. 2015

A percepção das cores é o resultado de um conjunto de interações: primeiramente da luz com o objeto, e em seguida da luz refletida por tal objeto com o sistema visual humano (MARQUES FILHO; VIEIRA NETO, 1999). Dessa forma, a cor de um objeto representa uma função de comprimento de onda de luz refletida por ele (LU, 1997). Ou seja, se um objeto é visto como vermelho, sobre a incidência de luz branca, significa que ele reflete somente o comprimento de onda referente à cor vermelha, absorvendo o restante das radiações eletromagnéticas (MARQUES FILHO; VIEIRA NETO, 1999).

Esse fenômeno de reconhecimento de cores pelo olho humano é devido à sua estrutura, formada por três membranas: *(i)* a córnea e a esclera, *(ii)* a coróide, e *(iii)* a retina. A retina é composta por células capazes de reconhecer as cores, denominadas de cones, e por células capazes de reconhecer a luminosidade, denominadas de bastonetes (GELADI; GRAHN, 1996; GONZALEZ; WOODS, 2003).

A quantidade de cones presentes em cada olho é de aproximadamente 6 milhões, classificados de acordo com sua sensibilidade nas zonas de cores: vermelho, verde e azul. Dessa forma, todas as cores perceptíveis ao ser humano, correspondem a uma combinação variável dessas três tonalidades de cores, também denominadas como cores primárias (GONZALEZ; WOODS, 2003; HUGHES et al., 2013)

A diferenciação de uma cor de outra baseia-se em três características: *(i)* brilho, *(ii)* matiz e *(iii)* saturação. Entretanto, a combinação em conjunto do matiz e da saturação pode ser denominada como cromaticidade. Logo, uma cor pode ser distinguida por sua cromaticidade e brilho (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

O brilho representa a quantidade de radiação luminosa recebida pelo olho, variando desde muito escuro (preto), até muito brilhante (branco). O matiz representa o comprimento de onda dominante frente a uma associação de espectros visíveis. Já a saturação exprime o grau de pureza do matiz, ou seja, representa o nível de pureza de uma cor. Dessa maneira, cores monocromáticas, como o vermelho, possuem saturação máxima. De maneira oposta, cor como lilás, formada pela mistura de violeta e branco apresenta níveis inferiores de saturação (GONZALEZ; WOODS, 2003).

### 4.2.3 Modelos de cores

Os modelos de cores, também denominados espaços ou sistemas de cores, são representações tridimensionais, onde cada ponto, determinado por três coordenadas,

representa uma cor. A finalidade dos modelos de cores é especificar as cores de maneira padronizada e organizada para uma fácil aceitação geral (HUGHES et al., 2013; MARQUES FILHO; VIEIRA NETO, 1999). Os modelos de cores são amplamente utilizados em monitores coloridos e impressoras, e também no processamento de imagens digitais (*Digital Image Processing - DIP*).

O modelo de cor mais utilizado é o RGB. Outros modelos de cores de recorrentes utilizações são: CMY (*Cyan, Magenta, Yellow*), utilizado principalmente em impressoras coloridas; YIQ (*Luminance In-phase Quadrature*), que corresponde ao formato NTSC em televisores coloridos; HSI (*Hue, Saturation, Intensity*) e HSV (*Hue, Saturation, Value*), geralmente utilizados na manipulação de imagens coloridas (IBRAHEEM et al., 2012; SHIH et al., 2001).

#### 4.2.3.1 Modelo RGB

O modelo de cores RGB pode ser representado como um cubo, Figura 5, onde cada vértice corresponde a uma cor. A cor preta encontrada no vértice simultâneo à origem  $(0, 0, 0)$  e a cor branca localizada no vértice mais afastado da origem  $(255, 255, 255)$ . Os demais vértices correspondem às cores primárias: vermelho  $(255, 0, 0)$ , verde  $(0, 255, 0)$  e azul  $(0, 0, 255)$ , e às cores secundárias amarelo  $(255, 255, 0)$ , magenta  $(255, 0, 255)$  e ciano  $(0, 255, 255)$ . Nesse modelo a escala cinza é representada pela diagonal principal do cubo, ou seja, a linha que se inicia na origem (preto) e que se prolonga até o vértice referente à cor branca (GONZALEZ; WOODS, 2003).

Esse modelo de cores pode ser classificado em: (i) espaço RGB linear, e (ii) espaço RGB não linear. O espaço RGB linear não é apropriado para realizar análises numéricas, assim seu uso na representação de imagens é bastante limitado. Entretanto, ele é utilizado em computação gráfica. Já o espaço RGB não linear, é amplamente utilizado na representação e no DIP. Os valores de R, G e B do espaço não linear variam entre 0 e 255 (IBRAHEEM et al., 2012; PLATANIOTIS; VENETSANOPOULOS, 2000).

A utilização desse modelo é bastante difundida no campo de DIP como no processamento de imagens multiespectrais obtidas por sensoramento remoto, onde a composição RGB, pertencente a esse conjunto de espectros, detém uma gama de informações relevantes e, dessa forma, possui grande utilidade e aplicação (FIGUEIREDO, 2005; SCHOWENGERDT, 1997).

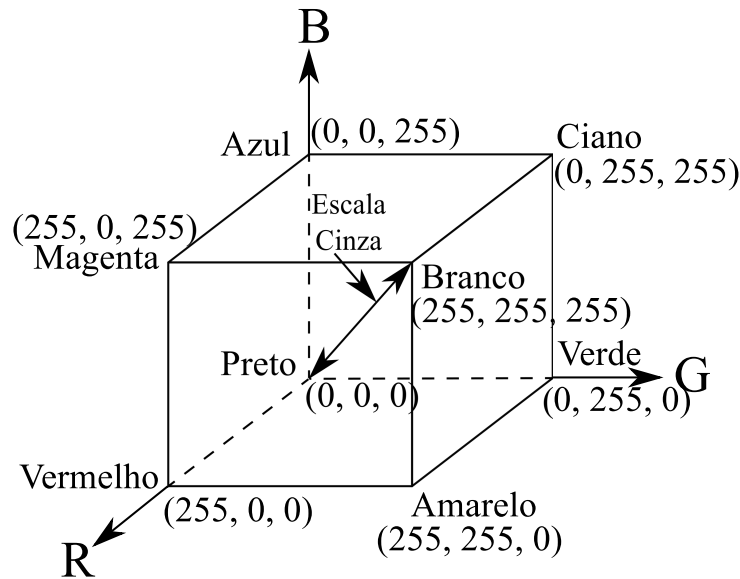


Figura 5 – Modelo de cores RGB.

Além disso, o modelo RGB também é utilizado em monitores coloridos, onde a cor é representada pela junção de 3 *bytes* por cada *pixel*, onde cada *byte* é contido no intervalo de 0 a 255, e referente a um plano de cor: (i) vermelho, (ii) verde e (iii) azul. Assim cada *pixel* pode ser representado por: 256x256x256 combinações, resultando em um total de 16.777.216 tonalidades de cor (GELADI; GRAHN, 1996).

#### 4.2.3.2 Modelo HSI

O modelo de cor HSI, também conhecido como HIS (*Hue, Intensity, Saturation*), representa três atributos: matiz (H), saturação (S) e intensidade (I), que também são reconhecidos pela visão humana. Esse modelo é utilizado em sistemas artificiais de reconhecimento visual, cujo funcionamento é semelhante à percepção de cores pelo olho humano. Assim, foi possível desenvolver dispositivos automatizados de colheita de frutas e vegetais, que reconhecem o grau de madurez dos mesmos em função da cor, da mesma forma que um ser humano. Além disso, esse modelo também é utilizado em inspeções de qualidade, que se baseiam na análise da cor (GONZALEZ; WOODS, 2003).

Esse modelo pode ser representado como um sólido geométrico, Figura 6, no qual cortes horizontais geram triângulos formados por vértices correspondentes às cores primárias, e o centro formado pela junção dessas três cores em iguais proporções. O resultado dessa combinação de cores estará mais próximo do branco ou do preto em função da altura do corte (GONZALEZ; WOODS, 2003).



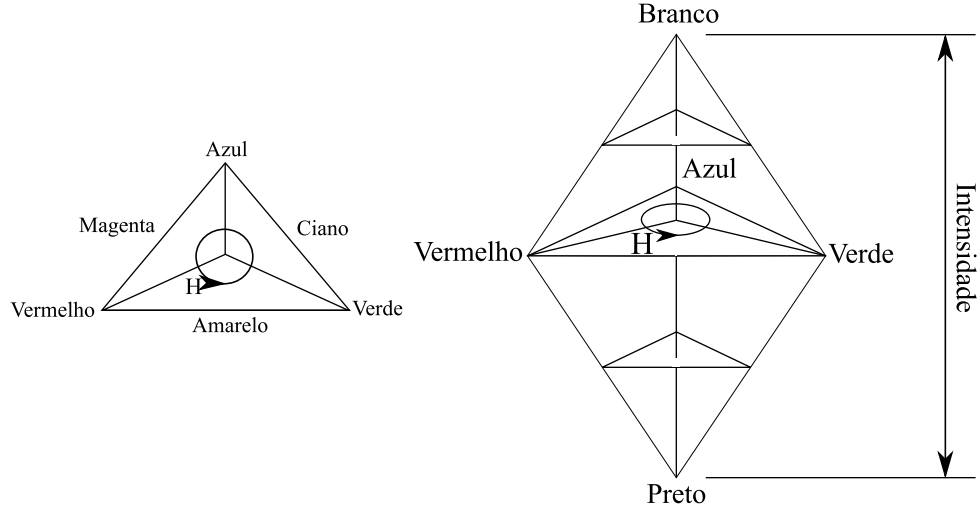


Figura 6 – Modelo de cores HSI.

A conversão do modelo RGB para o modelo HSI se dá através das Eqs. (1), (2), (3), essa conversão é possível (PM; CHEZIAN, 2013; MARQUES FILHO; VIEIRA NETO, 1999; TKALCIC; TASIC, 2003).

$$I = \frac{1}{3}(R + G + B) \quad (1)$$

$$S = 1 - \frac{3}{(R + G + B)}[\min(R, G, B)] \quad (2)$$

$$H = \cos^{-1} \left( \frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right) \quad (3)$$

### 4.3 DIP

O campo de DIP visa duas aplicações básicas: (i) gerar uma imagem a partir de outra que pode ser facilmente interpretada pelo ser humano, e (ii) produzir uma imagem a partir de outra mais acessível à análise computacional. De modo geral, denomina-se a primeira aplicação como processamento de imagens, e a segunda como visão por computador ou visão computacional.

O processamento de imagens é utilizado em várias áreas do conhecimento como na medicina com o aperfeiçoamento das imagens de raio X, e também de outras imagens de interesse médico, o que popularizou o diagnóstico médico por imagens (BARBOSA, 2000). Além disso, DIP também é utilizado no campo da geografia, tanto em sensoriamento remoto

como também em sistemas de geoprocessamento (FIGUEIREDO, 2005); em arqueologia, na restauração de fotografias antigas e danificadas (MENESES, 2013); em Física, na otimização de imagens geradas por técnicas de microscopia (CRUZ, 2007); em biologia (ALVES, 2006), e na Química, aonde o primeiro trabalho publicado nessa área no Brasil aplicava DIP e Análise de Componentes Principais (PCA) para o reconhecimento de padrões de bebidas refrigerantes (GODINHO et al., 2008), dentre outras áreas.

Já a visão computacional é utilizada em inúmeros processos industriais, principalmente no controle de qualidade (GOMES; LETA, 2012). A crescente aplicação dessa técnica se deve graças ao incremento da automação industrial (GONZALEZ; WOODS, 2003).

O sistema de visão computacional (SVC) envolve não só um extenso embasamento teórico, mas também compreende uma gama de softwares e hardwares. Esse processo é composto por uma série de etapas básicas, Figura 7, como: aquisição, pré-processamento, segmentação, extração de características, reconhecimento e interpretação. Entretanto, ressalta-se que nem todos os sistemas de visão computacional seguem todas essas etapas, ou seja, é possível obter resultados na saída de qualquer etapa, mas a maioria das técnicas se limita até o módulo de pré-processamento (GELADI; GRAHN, 1996; GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999; PETROU; PETROU, 2010).

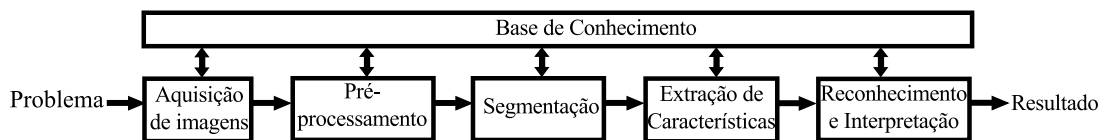


Figura 7 – Principais etapas do sistema de visão computacional.

#### 4.3.1 Aquisição

A primeira etapa do SVC é a aquisição de uma imagem analógica formada através de um sensor e, em seguida, convertida em uma imagem digital por meio de um digitalizador (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

#### 4.3.2 Pré-processamento

O objetivo dessa etapa é aperfeiçoar a imagem de tal forma que possibilite a máxima eficiência das etapas posteriores. Para tanto, utilizam-se técnicas que visam, principalmente,

o realce de contrastes, a minimização ou remoção de ruídos e imperfeições, e o isolamento de regiões de maior interesse. O resultado desse processo é uma imagem cuja qualidade seja superior à imagem de entrada (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

### **4.3.3 Segmentação**

A segmentação é uma etapa cuja principal função é a fragmentação da imagem em parcelas da região de interesse (ROI). Embora a descrição dessa fase seja simples, a sua implementação é bastante complexa devido ao fato de que a imprecisão na seleção da ROI influenciará desfavoravelmente a eficiência das etapas subsequentes (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999; LU, 1997; PETROU; PETROU, 2010; RUSS, 2011).

### **4.3.4 Extração de Características**

Nessa etapa são extraídas as informações relevantes da imagem. Para tanto, utilizam-se descritores que são capazes de determinar a informação requerida. Embora se processe, inicialmente, uma imagem, o resultado é um conjunto de dados referentes a essa imagem (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

### **4.3.5 Reconhecimento e Interpretação**

O reconhecimento é uma técnica que rotula um objeto a partir de dados fornecidos pelos descritores. Já a interpretação possui a função de fornecer significado aos objetos rotulados pelo reconhecimento (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

### **4.3.6 Base de Conhecimento**

Todas as etapas do SVC requerem informações sobre determinado problema a ser solucionado. Dessa forma, toda essa informação é arquivada em uma base de conhecimento, Figura 7, de dimensão e nível de complexidade variáveis de acordo com o problema

referido. Essa base de conhecimento não só orienta a atividade de cada etapa do SVC, como também influencia na comunicação entre elas. Além disso, em termos ideais, uma base de conhecimento deve possibilitar uma retroalimentação entre os estágios do SVC, de tal forma que cada fase possa reconhecer ou suspeitar de alguma falha no processo e encaminhá-la ao estágio que possivelmente a provocou, para um novo processamento. Entretanto, essa idealidade é de difícil implementação, e por isso não é largamente utilizada (GONZALEZ; WOODS, 2003; MARQUES FILHO; VIEIRA NETO, 1999).

## 4.4 Quimiometria

De modo geral, a quimiometria pode ser definida como uma ciência do ramo da química que adota métodos matemáticos, estatísticos e lógicos com a finalidade de: *(i)* projetar, planejar ou selecionar metodologias experimentais otimizados, *(ii)* proporcionar a maior quantidade de informação química possível através da análise dos dados químicos adquiridos e *(iii)* aprimorar os conhecimentos sobre processos de natureza química (FERREIRA et al., 1999; KRAMER, 1998; RODIONOVA; POMERANTSEV, 2007).

Dentre as aplicações da quimiometria, destaca-se a análise de dados, que é uma etapa importante em qualquer experimento e pode ser subdividida em univariada e multivariada. A análise multivariada leva vantagem em relação à univariada, porque considera as interações entre as variáveis, ou seja, reconhece os efeitos sinérgicos e antagônicos entre as variáveis (BARROS-NETO; SCARMINIO; BRUNS, 2001). Dessa forma, os métodos multivariados possibilitaram a extração de informações mais relevantes do que os métodos univariados. Graças a essas vantagens, os modelos estatísticos multivariados foram aplicados a dados químicos, o que foi imediatamente aceito e divulgado pela comunidade científica, representando então, o surgimento da quimiometria (BREITKREITZ; SOUZA; POPPI, 2014; KOWALSKI, 1975).

Embora a associação de química com estatística não seja uma combinação recente, ou seja, possivelmente iniciou-se com os trabalhos do químico cervejeiro W. S. Gosset (GOSSET, 1904), também conhecido pelo pseudônimo Student (BARROS-NETO; SCARMINIO; BRUNS, 2001), o desenvolvimento da quimiometria só foi significativo na década de 70, com o surgimento e popularização do computador no laboratório químico (KOWALSKI, 1980), principalmente associado aos instrumentos de medição (FERREIRA et al., 1999). A partir daí, graças ao crescimento e desenvolvimento dos recursos computacionais, a utilização

de computadores no tratamento e análise de dados químicos cresceu exponencialmente (ALCIATURI et al., 2010; RODIONOVA; POMERANTSEV, 2007).

Dentre as várias áreas da quimiometria, destaca-se a calibração multivariada, que consiste na geração de um modelo matemático, que por sua vez relaciona o sinal analítico emitido por um equipamento, com um parâmetro de interesse (FERREIRA et al., 1999; PIMENTEL; GALVÃO; ARAÚJO, 2008).

#### 4.4.1 Organização dos dados

A organização dos dados é uma etapa que antecede a calibração. Nesse estágio os dados são dispostos na forma de matrizes. Dessa forma, os dados analíticos são dispostos na matriz  $\mathbf{X}$  ( $n \times m$ ), onde cada linha e coluna representa, respectivamente, uma amostra e uma variável. Já o parâmetro de interesse, denominado de variável dependente, é disposto em um vetor  $\mathbf{y}$ , com número de linhas correspondente a quantidade de amostras  $n$ , e referente ao atributo ou característica que posteriormente será predita. A título de exemplo, em relação a dados de imagens, cada linha da matriz  $\mathbf{X}$ , Eq. (4), representa o histograma de uma imagem referente a uma amostra, e cada coluna uma tonalidade de cor (FERREIRA et al., 1999; KRAMER, 1998).

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (4)$$

#### 4.4.2 Pré-processamento de dados

Posteriormente à organização dos dados, realiza-se o pré-processamento dos dados, para tanto, utilizam-se vários métodos como: centrar na média, autoescalamento, normalização, correção de sinal ortogonal (OSC – *Orthogonal Signal Correction*) e algumas transformações matemáticas como, por exemplo, a primeira e a segunda derivadas. Esses métodos não só minimizam os ruídos, como também reduzem a informação irrelevante, e dessa forma melhoram a correlação entre as respostas analíticas e os atributos de interesse (FERREIRA et al., 1999; KRAMER, 1998; OTTO, 1999; RODIONOVA; POMERANTSEV, 2007).

#### 4.4.2.1 Centrar na Média

Centrar os dados na média é uma etapa de pré-processamento que simplifica a visualização dos dados. Para isso, primeiramente, calcula-se o valor médio para cada coluna ( $\bar{x}_j$ ), Eq. (5), e posteriormente, determina-se  $xcm_{i,j}$ , Eq. (6), que é a diferença entre a variável original ( $x_{i,j}$ ) e o valor médio para a respectiva coluna (OTTO, 1999).

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{i,j} \quad (5)$$

Onde  $n$  corresponde ao número de amostras.

$$xcm_{i,j} = x_{i,j} - \bar{x}_{i,j} \quad (6)$$

#### 4.4.2.2 OSC

A finalidade da técnica de pré-processamento OSC é aperfeiçoar o modelo de calibração multivariado, para isso ela remove da matriz  $\mathbf{X}$ , contendo os dados das amostras, somente os componentes que não apresentam nenhuma relação com o vetor  $\mathbf{y}$ , contendo os parâmetros de interesse. Para tanto, essa técnica assegura-se que a porção a ser removida seja ortogonal, ou praticamente ortogonal ao bloco  $\mathbf{y}$  (WOLD et al., 1998).

### 4.5 Calibração Multivariada

A calibração multivariada correlaciona, através de um modelo matemático, as respostas multivariadas geradas por um instrumento com um atributo físico-químico de interesse. Um modelo multivariado é gerado para prever determinada propriedade nas amostras a partir das suas respectivas respostas analíticas (FERREIRA et al., 1999; MARTENS; NÆS, 1991).

De modo geral, a calibração multivariada baseia-se em dois passos básicos: modelagem, que consiste na geração do modelo matemático que correlaciona a matriz das variáveis dependentes  $\mathbf{X}$  ao vetor de respostas  $\mathbf{y}$ ; e validação, que consiste na avaliação do modelo gerado. Nessa última, um novo conjunto de amostras, que não foi utilizado no

processo de calibração, é considerado, e suas propriedades de interesse são previstas pelo modelo construído anteriormente (FERREIRA et al., 1999; KRAMER, 1998).

Conforme mencionado, os dois estágios da calibração envolvem grupos de amostras diferentes, e essas devem ser selecionadas a partir do conjunto total de amostras. Para tanto, utilizam-se algumas metodologias de seleção como, por exemplo, o algoritmo Kennard-Stone (KS), e o algoritmo SPXY (*Sample set Partitioning based on joint  $x - y$  distances*) (GALVÃO et al., 2005).

Dentre os métodos de calibração multivariada aplicados em dados de natureza química, os mais adotados são: a regressão linear múltipla (RLM), regressão por componentes principais (PCR) e regressão por mínimos quadrados parciais (PLS) (FERREIRA et al., 1999).

#### 4.5.1 Algoritmo Kennard-Stone (KS)

A finalidade do algoritmo KS é selecionar, a partir de um conjunto de  $n$  amostras, um subconjunto representativo, que por sua vez seja uniforme por toda a extensão do bloco  $\mathbf{X}$ , que contém as informações das amostras. Para tanto, o KS adota uma metodologia que primeiramente seleciona um par de amostras  $(p, q)$ , com a maior distância Euclidiana,  $d_x(p, q)$ , entre si, Eq. (7) e, em seguida, seleciona a amostra que apresenta a maior distância entre uma das amostras recém selecionadas. Esse processo se repete gradualmente até a seleção da quantidade de amostras desejadas.

$$d_x(p, q) = \sqrt{\sum_{j=1}^J [x_p(j) - x_q(j)]^2}; \quad p, q \in [1, n] \quad (7)$$

Tratando-se de histogramas de imagens,  $x_p(j)$  e  $x_q(j)$  correspondem à intensidade referente ao  $j$ -ésimo índice de cores, respectivamente para a amostra  $p$ , e  $q$ . Já  $J$  corresponde a quantidade de índices de cores (GALVÃO et al., 2005; KENNARD; STONE, 1969).

### 4.5.2 Algoritmo SPXY

O algoritmo SPXY atua de forma semelhante ao KS entretanto, em seu cálculo de distância, ele considera não só o bloco  $\mathbf{X}$  como também o vetor  $\mathbf{y}$ . Dessa forma, a distância Euclidiana,  $d_y(p, q)$ , para um par de amostras  $(p, q)$  é determinada pela Eq. (8).

$$d_y(p, q) = \sqrt{(y_p - y_q)^2} = |y_p - y_q|; p, q \in [1, n] \quad (8)$$

Com o intuito de equiparar a significância de cada amostra, o SPXY primeiramente calcula as distâncias Euclidianas:  $d_x(p, q)$ , e  $d_y(p, q)$ , respectivamente através das Eqs. (7) e (8). Em seguida, esse algoritmo considera a soma das distâncias calculadas,  $dx$  e  $dy$ , divididas por seus respectivos valores máximos, Eq. (9) (GALVÃO et al., 2005).

$$d_{xy}(p, q) = \frac{d_x(p, q)}{\max_{p, q \in [1, n]} d_x(p, q)} + \frac{d_y(p, q)}{\max_{p, q \in [1, n]} d_y(p, q)}; p, q \in [1, n] \quad (9)$$

### 4.5.3 RLM

A RLM é um dos métodos mais simples de calibração multivariada. Esse método, correlaciona a matriz  $\mathbf{X}$ , contendo os dados das amostras, com o vetor  $\mathbf{y}$ , contendo as respostas. Essa correlação é apresentada pela Eq. (10), e está representada pela Figura 8 (MARTENS; NÆS, 1991).

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad (10)$$

Onde  $\mathbf{b}$  corresponde a um vetor que contém os coeficientes de regressão, e pode ser calculado, utilizando o método dos mínimos quadrados, Eq. (11).

$$\mathbf{b} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{y} \quad (11)$$

A RLM utiliza toda a informação presente na matriz  $\mathbf{X}$ , o que também implica na inserção de dados que muitas vezes, não apresentam nenhuma relevância ao modelo. Por isso, a aplicação da RLM só é vantajosa quando aplicada a dados sem interferentes, com baixo ruído, com nenhuma colinearidade, e que apresentem respostas lineares. Além disso, a RLM só é aplicável a um conjunto de dados, onde se tenha mais amostras do que



variáveis, ou seja, a matriz  $\mathbf{X}$ , deve conter mais linhas do que colunas, caso contrário, a matriz  $\mathbf{X}$  não poderá ser invertida, Eq. (11), e desse modo, o sistema não terá solução (MARTENS; NÆS, 1991).

$$\begin{array}{c} 1 \\ \boxed{\mathbf{y}} \\ n \end{array} = \begin{array}{c} m \\ \boxed{\mathbf{X}} \\ n \end{array} \begin{array}{c} 1 \\ \boxed{\mathbf{b}} \\ m \end{array} + \begin{array}{c} 1 \\ \boxed{\mathbf{e}} \\ n \end{array}$$

Figura 8 – Representação da RLM.

#### 4.5.4 PLS

O método PLS é uma ferramenta de referência e de ampla utilidade em processos estatísticos multivariados, nos quais as variáveis relacionadas possuem alta correlação (LI; MORRIS; MARTIN, 2002). Na química, por exemplo, essa técnica possui inúmeras aplicações tanto em química analítica (SJÖSTRÖM et al., 1983; GUOHUA; GUOAN, 1989) como também em química ambiental (ERIKSSON et al., 1995; NOUWEN et al., 1997; VERHAAR et al., 1994). Além disso, ela também é utilizada na inspeção de dezenas de processos industriais (KOURTI, 2005; KOURTI; LEE; MACGREGOR, 1996).

Essa técnica foi proposta por Herman Wold, um sueco estatístico, na década de 60 e foi utilizada primeiramente em previsões econômicas (WOLD, 1966a; WOLD, 1966b). Somente a partir dos anos 80, Svante Wold, filho de Herman Wold, juntamente com um grupo de pesquisadores, incluindo Harald Martens, defenderam a aplicação dessa técnica em dados químicos (WOLD; MARTENS; WOLD, 1983).

A finalidade da técnica PLS é maximizar a covariância entre a matriz  $\mathbf{X}$ , contendo as informações das amostras, e o vetor que armazena as respectivas concentrações,  $\mathbf{y}$ . Para tanto, esse modelo seleciona, a partir do espaço multivariado, um conjunto de variáveis latentes (LVs), as quais são combinações lineares das variáveis originais, e que por sua vez são decompostas em escores,  $\mathbf{t}$ , e pesos,  $\mathbf{w}$ . Essa decomposição, que ocorre de forma

independente tanto no bloco  $\mathbf{X}$  como no  $\mathbf{y}$ , é apresentada pelas Eqs. (12) e (13), e está representada pela Figura 9 (MARTENS; NÆS, 1991; OTTO, 1999).

$$\mathbf{X} = \mathbf{TP}^t + \mathbf{E} \quad (12)$$

$$\mathbf{y} = \mathbf{uq}^t + \mathbf{f} \quad (13)$$

onde  $\mathbf{T}$  e  $\mathbf{u}$  representam respectivamente a matriz e o vetor de escores para a matriz  $\mathbf{X}$  e para o vetor  $\mathbf{y}$ ;  $\mathbf{P}$  e  $\mathbf{q}$  compreendem os pesos relativos aos blocos  $\mathbf{X}$  e  $\mathbf{y}$ , respectivamente;  $\mathbf{E}$  é a matriz que contém os erros (resíduos) da matriz  $\mathbf{X}$ ; e  $\mathbf{f}$  é o vetor que abrange os erros (resíduos) referentes ao vetor  $\mathbf{y}$ .

Posteriormente, utilizando as matrizes:  $\mathbf{P}$  e  $\mathbf{W}$ , e o vetor  $\mathbf{q}$ , é possível encontrar o vetor  $\mathbf{b}$ , contendo os coeficientes de regressão do modelo, através da equação Eq. (14)

$$\mathbf{b} = \mathbf{W}(\mathbf{P}^t\mathbf{W})^{-1}\mathbf{q}^t \quad (14)$$

onde  $\mathbf{W}$  é a matriz que contém os pesos do modelo PLS, apresentado pela Eq. (15).

$$\mathbf{XW} = \mathbf{T} \quad (15)$$

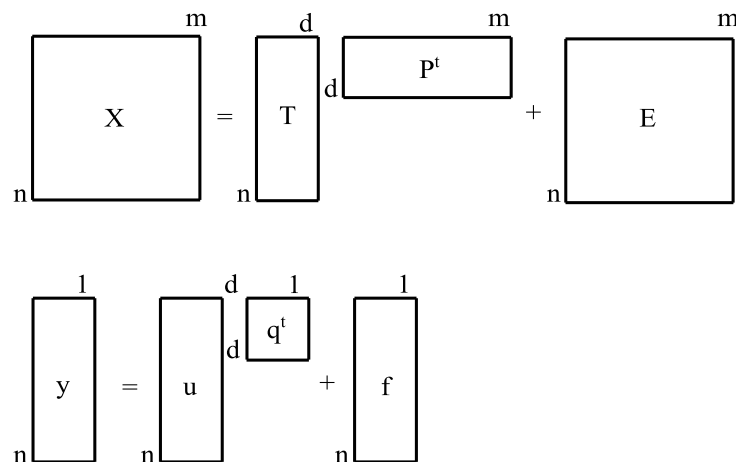


Figura 9 – Decomposição da matriz  $\mathbf{X}$  e do vetor  $\mathbf{y}$  em LVs.

#### 4.5.5 Seleção de LVs

Apesar da imensa quantidade de LVs que podem ser calculadas em um modelo PLS, é necessária uma descrição na escolha do número de LVs a serem incluídas no modelo,

com o intuito de evitar tanto o sobreajuste quanto o subajuste. Para isso, utiliza-se, por exemplo, a validação cruzada (CV), que primeiramente retira uma amostra do conjunto total, e em seguida, constrói um modelo de calibração. Posteriormente, o modelo criado é empregado para prever a composição dessa amostra removida. Esse processo se repete, até que todas as amostras sejam previstas pelo modelo (FERREIRA et al., 1999).

A técnica de CV descrita anteriormente é denominada *leave-one-out*, e é recomendada para um grupo de até 20 amostras. Dessa forma, se tratando de um conjunto maior de amostras, utiliza-se a técnica dos blocos contínuos (do inglês, *block continuous*), que realiza o mesmo processo que a *leave-one-out*, diferenciando-se, somente, que ao invés de uma amostra retirada, retira-se um grupo de amostras, com tamanho elegível (BRERETON, 2007; FERREIRA et al., 1999).

Terminada a previsão de todas as amostras, calcula-se, para cada LV, a raiz quadrada do erro médio quadrático da validação cruzada (RMSECV), Eq. (16), que mensura a eficiência das previsões do modelo gerado. Assim, seleciona-se o número de LVs de acordo com o menor RMSECV, ou quando se chega ao ponto em que a adição de uma nova variável latente não provoca uma alteração relevante ao modelo (FERREIRA et al., 1999; MARTENS; NÆS, 1991).

$$RMSECV = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{prev})^2}{n_{cal}}} \quad (16)$$

sendo  $y_{ref}$  e  $y_{prev}$ , respectivamente, os valores de referência e previstos para a variável de interesse, e  $n_{cal}$  o número de amostras do conjunto de calibração.

#### 4.5.6 Detecção de *outliers*

A homogeneidade das amostras representa um parâmetro de qualidade em modelos de regressão, dessa forma, amostras anômalas, também denominadas de *outliers*, devem ser detectadas, investigadas, e em alguns casos removidas do modelo, já que elas podem prejudicar o processo de calibração. A detecção de *outliers* pode ser realizada analisando o *leverage* de cada amostra, que em outras palavras representa a influência de tal amostra no modelo. Assim, uma amostra com valor de *leverage* baixo, significa que ela influencia pouco no modelo. Já uma amostra com alto *leverage*, proverá uma alta influência, o que pode

ser desfavorável ao modelo de calibração. De modo geral, a *leverage* pode ser calculada através da Eq. (17).

$$h_{ii} = \frac{1}{n_{cal}} + (\mathbf{x}_i - \bar{\mathbf{x}})^t (\mathbf{X}^t \mathbf{X})^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \quad (17)$$

onde  $\mathbf{x}_i$  refere-se a  $i$ -ésima linha da matriz  $\mathbf{X}$ , contendo os dados das amostras;  $\bar{\mathbf{x}}$  é a média de todas as linhas da matriz  $\mathbf{X}$ ;  $n_{cal}$  corresponde ao número de amostras de calibração; e  $\mathbf{X}^t \mathbf{X}$  é a matriz de variância-covariância.

Geralmente, uma amostra é reconhecida como *outlier*, quando seu respectivo  $h_{ii}$  é superior ao  $h_{crit}$ , que por sua vez pode ser calculado através da Eq. (18).

$$h_{crit} = \frac{3k}{n_{cal}} \quad (18)$$

onde  $k$  corresponde ao número de LVs (FERREIRA et al., 1999; MARTENS; NÆS, 1991).

#### 4.5.7 Avaliação dos modelos de calibração multivariada através de métodos estatísticos

Posteriormente à construção de modelos de calibração multivariados, vários parâmetros quimiométricos podem ser calculados com o intuito de avaliar a eficiência desses modelos. A título de exemplo, a raiz do erro quadrático médio de calibração (RMSEC - *Root Mean Square Error of Calibration*), e a raiz do erro quadrático médio de previsão (RMSEP - *Root Mean Square Error of Prediction*), que podem ser calculados, respectivamente, através das Eqs. (19) e (20). Tais equações relacionam os valores verdadeiros (fornecidos pelo método de referência) com os valores obtidos pelo modelo (FERREIRA et al., 1999; GALVÃO et al., 2007).

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{cal})^2}{n_{cal}}} \quad (19)$$

onde  $y_{ref}$  e  $y_{cal}$  correspondem respectivamente aos valores de referência e os valores previstos pelo modelo de calibração, e  $n_{cal}$  o número de amostras do conjunto de calibração.

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{val})^2}{n_{val}}} \quad (20)$$

onde  $y_{val}$  são os valores previstos, e  $n_{val}$  o número de amostras do conjunto de validação.

Além do RMSEC e do RMSEP, outros parâmetros podem ser determinados a fim de mensurar a qualidade dos modelos construídos, por exemplo: o coeficiente de correlação de Pearson ao quadrado ( $R^2$ ), o erro padrão de revisão (SEP - *Standard Error of Prediction*), Eq. (21) (MARTENS; NÆS, 1991), e o Desvio Residual de Previsão (RPD - *Residual Prediction Deviation*) (IGNE; HURBURGH, 2010; FEARN, 2002), que por sua vez, corresponde ao quociente do desvio padrão das amostras pertencentes ao conjunto de validação ( $DP_{val}$ ) pelo SEP, Eq. (23).

$$SEP = \sqrt{\frac{\sum_{i=1}^{n_{val}} (y_{ref} - y_{val} - bias)^2}{n_{val} - 1}} \quad (21)$$

onde *bias* (OTTO, 1999) corresponde pela diferença entre o valor de referência,  $y_{ref}$ , fornecido pelo método de referência, e o valor previsto pelo modelo,  $y_{val}$ , Eq. (22).

$$bias = \frac{\sum_{i=1}^{n_{val}} (y_{ref} - y_{val})}{n_{val}} \quad (22)$$

$$RPD = \frac{DP_{val}}{SEP} \quad (23)$$

O RPD é uma figura de mérito bastante utilizada na comparação de modelos de calibração. Tal parâmetro, avalia a qualidade do modelo construído de acordo com a predição das amostras do conjunto de validação. Essa avaliação baseia-se em uma escala numérica de tal forma que quanto maior o valor do RPD, melhor o modelo. Por exemplo, um modelo de calibração com RPD maior que 5 pode ser utilizado na análise de controle de qualidade. Da mesma maneira, que um modelo com RPD maior que 6,5 pode ser utilizado no monitoramento de processos. Já um modelo com RPD superior a 8 pode ser utilizado em qualquer aplicação (WILLIAMS; NORRIS et al., 1987; IGNE; HURBURGH, 2010).

#### 4.5.8 Seleção das variáveis

Os modelos de calibração multivariada são aplicados a inúmeros conjuntos de dados multivariados, com a finalidade de prever uma ou mais propriedades, obtendo muitas vezes resultados satisfatórios. Entretanto, com o intuito de minimizar a quantidade de variáveis, aperfeiçoar o modelo, e proporcionar melhores interpretações, utilizam-se algumas técnicas de seleção de variáveis, por exemplo, o algoritmo genético (GA - *Genetic Algorithm*), e o algoritmo de projeções sucessivas (SPA - *Successive Projections Algorithm*).

#### 4.5.8.1 GA

O GA foi desenvolvido com o intuito de otimizar sistemas. Esse algoritmo simula de modo matemático toda a estruturação da evolução biológica das espécies proposta por Darwin (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

O GA quando aplicado a dados de natureza química, seleciona um conjunto de variáveis da matriz  $\mathbf{X}$ , que correlacionados com o vetor  $\mathbf{y}$ , proporcionem os menores erros de previsão. Para tal, o GA procede com cinco etapas básicas (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999):

##### 4.5.8.1.1 Codificação das variáveis

Nesta etapa, cada variável da matriz  $\mathbf{X}$ , agora denominada de gene, é codificada em números binários, sendo o número 1 referente à variável selecionada, e o número 0 para variável rejeitada. Em uma imagem, por exemplo, cada índice de cor representaria uma variável, e essa, por sua vez, corresponderia a um gene. Logo, é possível codificar um histograma de frequência de uma imagem, como representado na Figura 10 (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

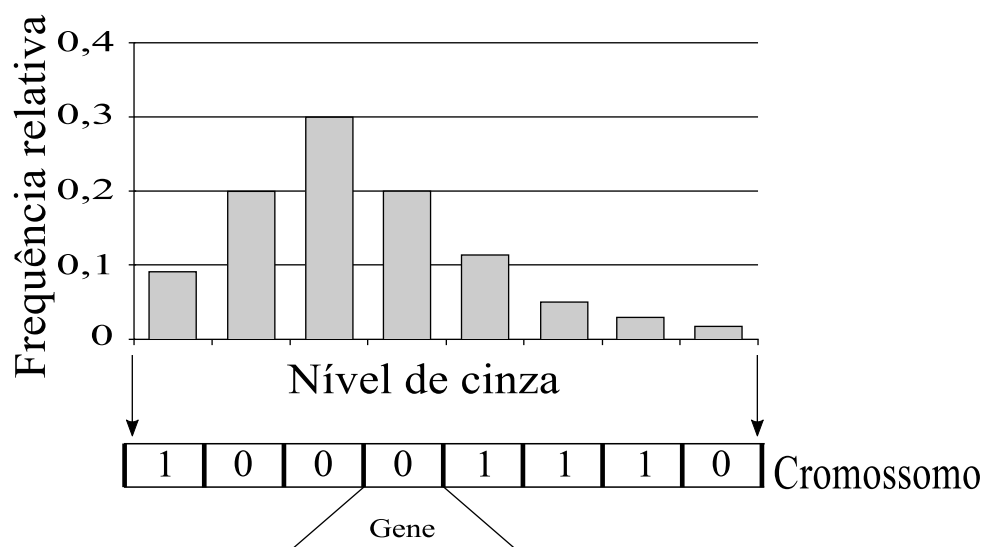


Figura 10 – Codificação de um histograma de frequência de uma imagem na conformação de um cromossomo. De tal forma, que cada tonalidade de cor refere-se a um gene do cromossomo.

#### 4.5.8.1.2 Criação da população original

Nesta etapa, uma população de cromossomos artificiais, formados por  $p$  genes, é gerada de forma aleatória, sendo  $p$  correspondente ao número de variáveis (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

#### 4.5.8.1.3 Avaliação das respostas

A avaliação das respostas é a etapa fundamental do GA. Nesta etapa, avalia-se a resposta fornecida por cada cromossomo criado. De modo geral, essa resposta corresponde à diferença entre o valor de referência e o valor previsto pelo modelo de regressão. Com tal avaliação, determinam-se quais cromossomos apresentam as melhores respostas, e estes, por sua vez, participarão da etapa de cruzamento (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

#### 4.5.8.1.4 Cruzamentos

Nesta etapa, ocorre o cruzamento, de modo aleatório, entre os cromossomos pertencentes a população vigente. Assim, uma nova população é formada (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

No cruzamento, sucede a troca de material genético entre os cromossomos. De modo que as características sobressalentes são transmitidas para as próximas descendências. Por isso, é comum obter, ao passo de algumas gerações, populações com características preponderantes (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

#### 4.5.8.1.5 Mutações

As mutações são pequenas alterações no sequenciamento genético de alguns indivíduos. Tais alterações, contanto que sejam benéficas, são transmitidas para as próximas gerações (ANDERSEN; BRO, 2010; COSTA FILHO; POPPI, 1999).

#### 4.5.8.2 SPA

O SPA é um método de seleção de variáveis desenvolvido com o objetivo de otimizar as calibrações multivariadas, através da redução da multicolinearidade dos dados. Logo, este método é bastante aplicável, em especial, na RLM, que apesar de ser uma técnica de regressão simples, é prejudicada quando o conjunto de dados apresenta problemas de colinearidade.

Anteriormente a aplicação do SPA, o conjunto de dados referentes a calibração, devem ser divididos em dois grupos: (i) calibração, e (ii) validação, também denominado de grupo teste. Esses dois grupos, por suas vez, devem ser organizados na forma de matrizes:  $\mathbf{Xc}(n_c \times m)$ , contendo os dados das amostras de calibração,  $\mathbf{Xv}(n_v \times m)$ , contendo os dados das amostras de validação e também na forma de vetores:  $\mathbf{yc}(n_c \times 1)$ , contendo as respostas referentes ao conjunto de calibração, e  $\mathbf{yv}(n_v \times 1)$ , contendo as respostas referentes ao conjunto de validação, onde  $n_c$ , e  $n_v$  correspondem, respectivamente, ao número de amostras de calibração e de validação, e  $m$  corresponde ao número de variáveis (GALVÃO et al., 2007; GALVÃO FILHO, 2010).

O mecanismo do SPA inicia-se na partição da matriz de calibração  $\mathbf{Xc}(n_c \times m)$  em  $m$  subconjuntos com  $n_c$  variáveis cada uma. Cada subconjunto inicia-se com somente uma variável, em seguida seleciona-se, a partir das variáveis remanescentes, qual variável ou coluna apresenta a maior projeção ortogonal com a variável inicial. Esse processo se repete selecionando as próximas variáveis até a seleção do número limite de colunas,  $n_c$ . Posteriormente, constrói-se para cada subconjunto um modelo de RLM e calcula-se para cada modelo construído o erro quadrático de predição (RMSEV), Eq. (24). Em seguida, seleciona-se o subconjunto que apresentar o menor valor de RMSEV. Por último, realiza-se a eliminação de algumas variáveis do subconjunto recém selecionado que não promovem nenhum ganho significativo ao modelo de predição, através do teste  $F$  (GALVÃO et al., 2007; GALVÃO FILHO, 2010).

$$RMSEV = \sqrt{\frac{\sum_{i=1}^n (y_{ref} - y_{val})^2}{n_v}} \quad (24)$$

onde  $y_{ref}$  e  $y_{val}$  correspondem, respectivamente, aos valores de referência e aos valores previstos, e  $n_v$  corresponde o número de amostras do conjunto de validação.



## 5 Parte experimental

### 5.1 Amostragem

#### 5.1.1 Coleta de amostras

Foram obtidas 60 amostras de solo coletadas a uma profundidade de até 10 cm. Dois tipos de trados foram empregados: (i) rosca, e (ii) holandês, Figura 11. As coletas foram realizadas em Pirenópolis, Santo Antônio de Goiás, e Brazabrantes, municípios do estado de Goiás; em Belém do São Francisco, no estado do Pernambuco; e em Nova Canaã do Norte, no estado do Mato Grosso.



Figura 11 – Trado tipo holandês utilizado na coleta de amostras.

Fonte: Mercado Físico Rural – MF Rural <<http://www.mfrural.com.br/detalhe.asp?cdp=28937&nmoca=trado-galvanizado-para-coletar-amostra-de-terra-no-solo>>. Acesso em 23 mar. 2015

#### 5.1.2 Preparo das amostras

Após a coleta, as amostras foram secadas em estufa a 45 °C por 48 horas. Posteriormente à secagem, as amostras foram moídas em moinhos de martelo para solos, tipo Tecnosolo TE 330. Depois peneiradas em malha de 2 mm. Por último, as amostras foram ensacadas e identificadas.

### 5.2 Metodologia para análise textural do solo

Após a coleta e preparo das amostras, determinaram-se seus teores de areia, silte e argila utilizando o método da pipeta. Para tanto, pesaram-se em um béquer 10 g de amostra

de solo seco, em seguida adicionaram-se os dispersantes químicos: 10 mL de hidróxido de sódio  $1 \text{ mol L}^{-1}$ , 10 mL de hexametáfosfato de sódio  $1 \text{ mol L}^{-1}$ , e 2 mL de água oxigenada 9%. Posteriormente, realizou-se a dispersão física utilizando um processador ultra-sônico UP400S (400W, 24kHz), Figura 12.

Após as duas etapas de dispersão, que ocorrem concomitantemente, o conteúdo do béquer é passado por uma peneira de malha de 0,05 mm, acoplada a um funil e uma proveta de 1,0 L. O peneiramento retém a areia que é transferida para um béquer de massa conhecida,  $m_{ar}$ , e encaminhado a estufa à  $105^\circ\text{C}$  por 48 horas. Após a secagem, determina-se a massa do béquer contendo a areia seca,  $m_{bar}$ . O conteúdo que passou pela peneira foi coletada em uma proveta e o volume final foi ajustado com água para 1,0 L. A determinação da argila foi realizada retirando-se uma alíquota à profundidade de 5 cm do conteúdo da proveta, à um tempo específico, Tabela 1, referente ao tempo de sedimentação das partículas de silte em uma dada temperatura. Tal alíquota é transferida para um béquer de massa conhecida,  $m_b$ , que permanece em estufa à  $105^\circ\text{C}$  por 48 horas, em seguida, determina-se a massa do béquer com o conteúdo de argila seca,  $m_a$ . Por fim, determinaram-se as porcentagens de areia, argila e silte, através das Eqs. (26), (27) e (25).

$$Argila(\%) = 100 \times \frac{(m_a - m_b) \times 1000}{V_a \times m} \quad (25)$$

$$Areia(\%) = 100 \times \frac{(m_{ar} - m_{bar})}{m} \quad (26)$$

$$Silte(\%) = 100 - Areia(\%) - Argila(\%) \quad (27)$$

Onde  $m$  corresponde a massa de amostra de solo e  $V_a$  corresponde ao volume alíquotado à 5 cm.

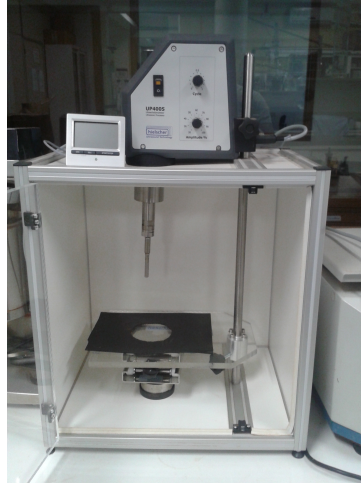


Figura 12 – Processador ultra-sônico utilizado na dispersão física das amostras.

Tabela 1 – Relação de temperatura da água durante a análise e o tempo de sedimentação

Temperatura (°C)	Tempo (h)	Temperatura (°C)	Tempo (h)
10	5 11'	20	4 00'
11	5 03'	21	3 54'
12	4 55'	22	3 48'
13	4 47'	23	3 43'
14	4 39'	24	3 38'
15	4 33'	25	3 33'
16	4 26'	26	3 28'
17	4 20'	27	3 24'
18	4 12'	28	3 19'
19	4 06'	29	3 15'

Fonte: (PANSU; GAUTHEYROU, 2003)

### 5.3 Aquisição das imagens

Em uma placa de petri foram adicionados 2 g da amostra de solo. Em seguida, essa placa contendo a amostra foi disposta no foco de uma Lupa binocular Leica EZ4 D com câmera digital incorporada, Figura 13, com dois diferentes aumentos: (i) 6,3x e (ii) 10x. Para cada amostra obtiveram-se seis imagens em triplicata para cada aumento óptico da Lupa, resultando em um total de 360 imagens. Essas imagens foram digitalizadas no sistema de cor RGB, resolução 2048 x 1536 *pixels*, e em formato TIFF (*Tagged Image File Format*).



Figura 13 – Lupa binocular Leica EZ4 D utilizada para a aquisição das imagens de solo.

## 5.4 Manipulação e tratamento das imagens

Utilizou-se o software MATLAB<sup>TM</sup> R2014b, versão 8.4.0 (*Mathworks, Natick, EUA*), com o emprego do pacote PLS Toolbox 8.1 (*Eigenvector Research, EUA*) para o pré-tratamentos de imagens e de dados, redução da dimensionalidade dos dados, conversão de modelos de cores, e também para realizar as calibrações multivariadas. Para tanto, as imagens foram armazenadas como matrizes tridimensionais ( $2048 \times 1536 \times 3$ ). Logo, extraíram-se, de cada imagem, histogramas de frequência dos canais RGB, HSI, e intensidade de escala cinza, em seguida calcularam-se os histogramas médios de cada triplicata de imagem. Os histogramas extraídos foram empilhados em sete matrizes diferentes: (i)  $\mathbf{X}_{\text{RGB}}$ , contendo somente os histogramas dos canais RGB, (ii)  $\mathbf{X}_{\text{HSI}}$ , contendo os histogramas dos canais HSI, (iii)  $\mathbf{X}_{\text{cinza}}$ , compreendendo os histogramas de intensidade de escala cinza, (iv)  $\mathbf{X}_{\text{RGBc}}$  compreendendo os histogramas dos canais RGB e da escala cinza, (v)  $\mathbf{X}_{\text{HSIc}}$ , compreendendo os histogramas dos canais HSI e da escala cinza, (vi)  $\mathbf{X}_{\text{RGBHSI}}$  compreendendo os histogramas dos canais RGB e HSI; e (vii)  $\mathbf{X}_{\text{RGBcHSI}}$ , compreendendo os histogramas dos canais RGB, HSI e da escala cinza.

Para divisão de amostras em conjuntos de calibração e validação, utilizaram-se o algoritmo SPXY (*Sample set Partitioning based on joint x-y distances*) (GALVÃO et al., 2005), e também o particionamento igualmente espaçado, de modo que o conjunto de calibração compreenda os dois limites extremistas de argila ou areia. Assim definiram-se 42 amostras para calibração, e 18 amostras para a validação. Posteriormente, os dados foram centrados na média e aplicou-se o método OSC como etapa de pré-processamento.

Em seguida, utilizaram-se os métodos de regressão multivariada PLS, e regressão linear múltipla associada ao algoritmo das projeções sucessivas (SPA-RLM) para correlacionar a matriz  $\mathbf{X}$ , contendo as informações das amostras de solo, com o vetor  $\mathbf{y}$ , correspondente aos teores de argila ou de areia, determinados experimentalmente pelo método da pipeta.

## 6 Resultados e Discussões

### 6.1 Análise textural do solo: método da pipeta

Os resultados das composições de argila, silte e areia, como também a classificação textural para as 60 amostras de solo analisadas pelo método da pipeta são apresentados na Tabela 2. De acordo com esses resultados, observa-se a predominância de amostras mais argilosas, pertencentes aos grupos muito argiloso e argiloso (Figura 1).

### 6.2 MIA

Após a obtenção das imagens, extraíram-se os histogramas de frequência dos canais de cores RGB, HSI, e escala cinza, e suas respectivas combinações (*i*) RGB, (*ii*) escala cinza, (*iii*) HSI, (*iv*) RGB + escala cinza, (*v*) HSI + escala cinza, (*vi*) RGB + HSI e (*vii*) RGB + HSI + escala cinza. Em seguida, para cada espaço de cor e suas combinações, construiu-se uma matriz  $\mathbf{X}$  diferente, que por sua vez é composta por 60 linhas, sendo cada linha referente a uma amostra, e cada coluna correspondendo a um índice de cor.

Já os dados físicos do solo foram representados pelo vetor  $\mathbf{y}$ , que primeiramente corresponde à porcentagem de argila, e em seguida pelos teores de areia, ambos determinadas pelo método da pipeta, Tabela 2.

Como etapa de pré-processamento de dados, todas as matrizes  $\mathbf{X}$  contendo os histogramas de frequência das amostras, foram centradas na média. Assim, para cada modelo de cor e suas combinações, construíram-se um modelo SPA-RLM e um modelo PLS, desse modo, obtiveram-se 7 modelos PLS e 7 modelos SPA-RLM para cada aumento óptico, 10 e 6,3x. Tal procedimento foi realizado tanto para o constituinte textural argila, como para areia, logo, totalizam-se 28 modelos PLS e 28 modelos SPA-RLM.

#### 6.2.1 Determinação do teor de argila por calibração multivariada

Cada matriz  $\mathbf{X}$  formada anteriormente juntamente com o vetor  $\mathbf{y}$  contendo os teores de argila foram divididas, utilizando o algoritmo SPXY, em dois conjuntos: calibração,  $\mathbf{Xc}$  e  $\mathbf{yc}$ , ambos referentes às 42 amostras, e validação,  $\mathbf{Xv}$  e  $\mathbf{yv}$ , ambos correspondentes à 18 amostras. Também realizou-se um particionamento de amostras de modo que o conjunto

Tabela 2 – Resultados da análise granulométrica das amostras de solo

Amostra	Areia (%)	Argila (%)	Silte (%)	Classificação
1	27,190	72,513	0,297	Muito Argiloso
2	32,196	67,519	0,285	Muito Argiloso
3	31,587	67,615	0,798	Muito Argiloso
4	32,868	67,049	0,083	Muito Argiloso
5	27,933	68,665	3,402	Muito Argiloso
6	27,329	70,341	2,330	Muito Argiloso
7	27,915	67,285	4,800	Muito Argiloso
8	34,992	58,912	6,096	Argiloso
9	35,776	60,547	3,677	Muito Argiloso
10	37,263	62,303	0,434	Muito Argiloso
11	28,200	70,946	0,854	Muito Argiloso
12	29,518	69,473	1,009	Muito Argiloso
13	28,899	68,729	2,372	Muito Argiloso
14	27,514	70,240	2,246	Muito Argiloso
15	31,327	67,789	0,884	Muito Argiloso
16	29,205	69,132	1,663	Muito Argiloso
17	39,624	57,327	3,049	Argiloso
18	41,138	54,638	4,224	Argiloso
19	41,140	54,841	4,019	Argiloso
20	37,707	59,007	3,286	Argiloso
21	37,424	58,441	4,135	Argiloso
22	40,427	57,530	2,043	Argiloso
23	40,203	57,458	2,339	Argiloso
24	40,512	57,154	2,334	Argiloso
25	38,868	60,973	0,159	Muito Argiloso
26	40,992	58,517	0,491	Argiloso
27	40,163	57,677	2,160	Argiloso
28	40,417	59,583	0,000	Argiloso
29	38,020	59,435	2,545	Argiloso
30	42,278	55,706	2,016	Argiloso
31	41,456	56,719	1,825	Argiloso
32	41,212	56,922	1,866	Argiloso
33	40,255	57,849	1,896	Argiloso
34	19,073	80,461	0,466	Muito Argiloso
35	60,339	37,792	1,869	Franco Argilo Arenosa
36	64,435	34,131	1,434	Franco Argilo Arenosa
37	63,755	35,948	0,297	Franco Argilo Arenosa
38	60,620	37,539	1,841	Franco Argilo Arenosa
39	61,565	35,479	2,956	Franco Argilo Arenosa
40	62,146	37,804	0,050	Franco Argilo Arenosa
41	70,130	29,688	0,182	Franco Argilo Arenosa
42	36,293	60,872	2,835	Muito Argiloso
43	41,600	58,279	0,121	Argiloso
44	25,858	72,272	1,870	Muito Argiloso
45	60,147	38,737	1,116	Franco Argilo Arenosa
46	74,733	23,528	1,739	Franco Argilo Arenosa
47	73,126	26,611	0,263	Franco Argilo Arenosa
48	43,190	56,320	0,490	Argiloso
49	69,440	30,492	0,068	Franco Argilo Arenosa
50	73,793	20,439	5,768	Franco Argilo Arenosa
51	34,082	59,284	6,634	Argiloso
52	33,704	58,671	7,625	Argiloso
53	18,019	81,522	0,459	Muito Argiloso
54	35,465	56,645	7,890	Argiloso
55	36,165	56,512	7,323	Argiloso
56	84,207	7,371	8,422	Areia franca
57	85,117	7,618	7,265	Areia franca
58	98,779	0,601	0,620	Areia
59	2,024	77,933	20,043	Muito Argiloso
60	73,117	26,783	0,100	Franco Argilo Arenosa

de calibração compreenda as amostras com o maior e menor teor de argila ou areia, e que seja igualmente espaçado.

Após a determinação da quantidade de LVs adequada, de acordo com os menores valores de RMSEP, para as imagens 6, 3x e 10x, correlacionaram-se as matrizes  $\mathbf{Xc}$ 's com o vetor  $\mathbf{yc}$ , contendo os teores de argila, utilizando o método dos mínimos quadrados PLS, e a regressão linear múltipla associada com o algoritmo das projeções sucessivas (SPA-RLM). Desse modo, construíram-se 7 modelos PLS e 7 modelos SPA-RLM, ou seja, um modelo PLS e SPA-RLM para cada modelo de cor, e suas respectivas combinações. Tal procedimento foi realizado para os dois modos de divisão entre calibração e validação. As características estatísticas dos modelos constam nas Tabelas 3, 4, 5 e 6.

De acordo com as Tabelas 3 e 5, o modelo PLS com melhor desempenho refere-se ao conjunto HSI + escala cinza das imagens de 10x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 14 apresenta a correlação entre os valores de argila previstos por esse modelo e os valores de argila determinados pelo método de referência. As demais correlações constam no Apêndice A.

Segundo as Tabelas 4 e 6, o modelo SPA-RLM com melhor desempenho refere-se ao conjunto HSI das imagens de 10x, assim a Figura 15 apresenta a correlação entre os valores de argila previstos por esse modelo e os valores de argila determinados pelo método de referência. As demais correlações constam no Apêndice B.

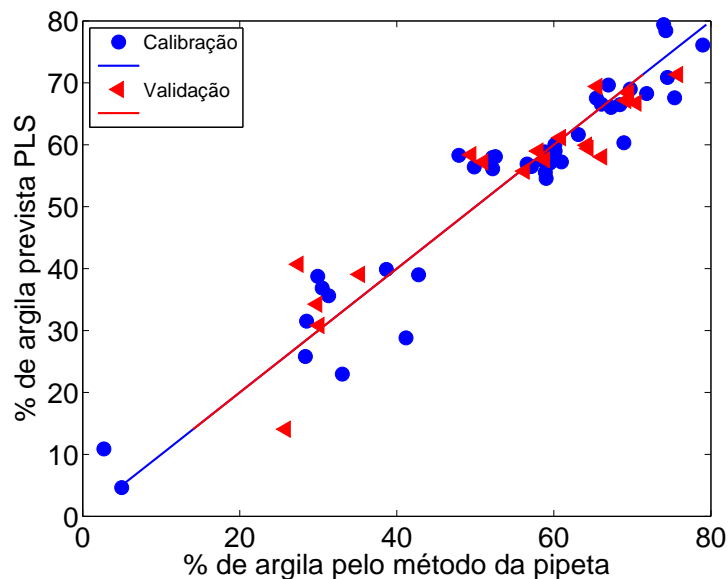


Figura 14 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◀), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.



Tabela 3 – Dados dos modelos PLS na predição dos teores de argila em amostras de solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	n° de LVs	Especificações do modelo PLS					
			RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
6,3x	RGB	5	6,470	0,9006	7,059	0,7267	16,686	1,283
	Escala Cinza	1	4,745	0,9737	7,966	0,7528	18,622	1,534
	RGB + Cinza	7	5,729	0,9218	7,506	0,6968	17,136	1,263
	HSI	6	3,881	0,9632	4,981	0,8229	20,380	2,374
	HSI + Cinza	2	6,127	0,9107	6,889	0,6437	13,595	1,562
	RGB + HSI	7	6,184	0,9093	5,869	0,8035	13,548	1,739
	RGB + HSI + Cinza	3	5,869	0,9170	5,773	0,8062	23,310	2,024
10,0x	RGB	8	4,741	0,9401	4,992	0,8229	34,296	1,827
	Escala Cinza	2	2,408	0,9851	5,304	0,8241	12,818	1,832
	RGB + Cinza	5	6,148	0,9012	5,497	0,8491	30,841	2,132
	HSI	5	4,354	0,9517	5,435	0,8999	15,574	3,052
	HSI + Cinza	7	3,698	0,9659	4,751	0,8489	13,046	2,421
	RGB + HSI	3	3,830	0,9627	5,062	0,8750	18,367	2,518
	RGB + HSI + Cinza	10	4,065	0,9588	5,354	0,8646	12,883	2,099

De acordo com a Tabela 3, não houve diferença estatística entre os valores de RMSEPs dos modelos PLS referentes as imagens de 6,3x por meio do teste  $F_{(P=0.05,17,17)}$ , com exceção do modelo correspondente a escala cinza. Já nos modelos PLS correspondentes as imagens de 10x, não foi observado diferença estatística entre os valores de RMSEP para nenhum dos modelo.

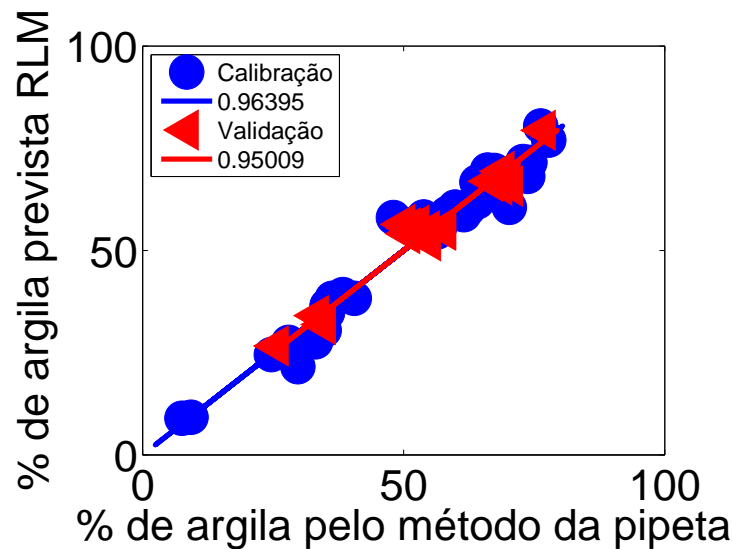


Figura 15 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

Segundo a Tabela 4, todos os modelos SPA-RLM construídos apresentam valores de RMSEP estatisticamente iguais de acordo com o teste  $F_{(P=0.05,17,17)}$ , com exceção do modelo referente ao espaço de cor RGB + escala cinza (6,3x) e o modelo referente a escala cinza, pertencente as imagens de 10x.

De acordo com a Tabela 5, os modelos PLS correspondentes aos espaços de cores RGB, HSI, HSI + escala cinza, RGB + HSI e RGB + HSI + escala cinza apresentaram valores de RMSEP iguais por meio do teste  $F_{(P=0.05,18,18)}$ . Nas imagens de 10x, os modelos PLS correspondentes aos espaços de cores RGB + HSI e HSI + escala cinza apresentaram valores de RMSEP iguais.

Segundo a Tabela 6, todos os modelos SPA-RLM referentes as imagens de 6,3x apresentam valores de RMSEP estatisticamente iguais através do teste  $F_{(P=0.05,18,18)}$ , com exceção dos modelos referentes aos espaços de cor HSI e RGB + HSI. Em relação as imagens de 10x, não houve diferença estatística entre os erros de previsão dos modelos

Tabela 4 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras de solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RPD
<b>6,3x</b>	RGB	5,017	0,9403	5,293	0,7321	1,717
	Escala Cinza	4,688	0,9451	5,524	0,8034	2,208
	RGB + Cinza	5,704	0,9224	8,183	0,7396	1,147
	HSI	7,126	0,8767	5,668	0,8324	2,160
	HSI + Cinza	4,437	0,9531	4,973	0,7639	2,054
	RGB + HSI	5,869	0,9183	6,517	0,6663	1,579
	RGB + HSI + Cinza	3,981	0,9618	4,407	0,8899	3,014
<b>10,0x</b>	RGB	6,028	0,9031	5,602	0,8196	1,620
	Escala Cinza	5,301	0,9279	8,089	0,6622	1,227
	RGB + Cinza	5,278	0,9272	5,686	0,7795	1,904
	HSI	3,762	0,9639	3,870	0,9501	4,027
	HSI + Cinza	5,049	0,9364	4,030	0,8817	2,798
	RGB + HSI	4,086	0,9576	4,747	0,8907	2,771
	RGB + HSI + Cinza	4,934	0,9393	4,359	0,8667	2,462

Tabela 5 – Dados dos modelos PLS na predição dos teores de argila em amostras de solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo PLS						
		nº de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
6,3x	RGB	4	6,420	0,8827	7,900	0,8124	18,692	2,150
	Escala Cinza	1	2,998	0,9744	11,668	0,7916	17,546	1,577
	RGB + Cinza	3	8,600	0,7895	9,193	0,7886	29,689	2,171
	HSI	2	7,428	0,8429	6,842	0,8873	18,205	2,689
	HSI + Cinza	2	6,748	0,8704	6,373	0,9045	20,758	2,930
	RGB + HSI	3	5,250	0,9216	7,123	0,8907	29,554	2,455
	RGB + HSI + Cinza	7	5,807	0,9040	5,594	0,9137	18,833	3,013
10,0x	RGB	7	5,480	0,9145	6,195	0,8731	11,092	2,720
	Escala Cinza	1	3,602	0,9630	7,496	0,8877	8,928	2,608
	RGB + Cinza	6	6,540	0,8783	8,434	0,8536	16,187	2,019
	HSI	8	5,163	0,9241	6,202	0,8692	16,220	2,762
	HSI + Cinza	5	4,580	0,9402	3,689	0,9522	12,167	4,569
	RGB + HSI	7	4,193	0,9499	4,747	0,9249	16,730	3,563
	RGB + HSI + Cinza	1	10,762	0,6704	11,241	0,6228	14,082	1,566

Tabela 6 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras de solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RPD
<b>6,3x</b>	RGB	2,603	0,9807	8,967	0,7379	1,910
	Escala Cinza	6,282	0,8877	9,771	0,8534	2,152
	RGB + Cinza	3,384	0,9674	9,417	0,7121	1,815
	HSI	3,013	0,9741	23,177	0,8077	0,735
	HSI + Cinza	2,822	0,9773	12,333	0,7551	1,367
	RGB + HSI	2,802	0,9776	15,095	0,8702	1,128
<b>10,0x</b>	RGB + HSI + Cinza	2,840	0,9770	12,236	0,8703	1,378
	RGB	2,686	0,9794	16,300	0,7404	1,131
	Escala Cinza	3,374	0,9676	9,823	0,7752	1,779
	RGB + Cinza	4,628	0,9390	9,369	0,8083	1,824
	HSI	3,058	0,9734	33,186	0,0298	0,513
	HSI + Cinza	2,627	0,9803	22,839	0,0338	0,7416
	RGB + HSI	5,848	0,9027	16,959	0,1983	1,010
	RGB + HSI + Cinza	7,846	0,8248	14,813	0,5767	1,138

correspondentes aos espaços de cores: escala cinza, RGB + escala cinza, RGB + HSI e RGB + HSI + escala cinza.

De acordo com as Tabelas 3 e 5, observa-se que de modo geral, a alteração do método de divisão entre calibração e validação, não alterou de forma significativa os resultados de predição. Entretanto, em relação as imagens de 10x, todos os valores de RMSEP calculados para esses dois modos de subdivisão são estatisticamente iguais, com exceção dos conjuntos RGB + escala cinza e RGB + HSI + escala cinza.

De modo análogo, conforme as Tabelas 4 e 6, comparando os valores de RMSEPs calculados para as imagens de 6, 3x para os dois modos de subdivisão de amostras, observou-se diferença estatística para todos os conjuntos, com exceção dos modelos escala cinza e RGB + escala cinza. Em relação as imagens de 10x, todos os valores de RMSEP calculados para esses dois modos de subdivisão são estatisticamente diferentes, com exceção do conjunto escala cinza.

Com o intuito de minimizar o número de variáveis e dessa forma aprimorar os modelos de calibração, aplicou-se o algoritmo genético (GA) para seleção de variáveis em modelos PLS empregando os parâmetros da Tabela 7.

De acordo com a Tabela 7, o GA foi empregado juntamente com a validação cruzada por blocos contínuos com o intuito de avaliar as respostas de cada indivíduo (RMSECV). Utilizou-se 4000 gerações com a finalidade de mensurar os valores de RMSECV ao longo de inúmeras gerações.

Os relatórios de desempenho do método de seleção de variáveis para cada modelo de cor e suas respectivas combinações são apresentados pelas Figuras 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 e 47, que constam nos Apêndices C e D.

Tabela 7 – Parâmetros elegidos para a aplicação do algoritmo genético

Parâmetros	Dados
Tamanho da população	64 cromossomos
Variáveis iniciais	230
Método de regressão	PLS
Validação cruzada	Blocos contínuos
Amostras por bloco	5
Possibilidade de cruzamento	90%
Probabilidade de mutação	0,5%
Número máximo de LVs	20
Número máximo de gerações	4000

Os parâmetros estatísticos dos modelos GA-PLS constam nas Tabelas 8 e 9.

De acordo com as Tabelas 8 e 9, o modelo GA-PLS com melhor desempenho refere-se ao conjunto RGB + HSI + escala cinza das imagens de 10x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 16 apresenta a correlação entre os valores de argila previstos por esse modelo e os valores de argila determinados pelo método de referência. As demais correlações constam no Apêndice E.

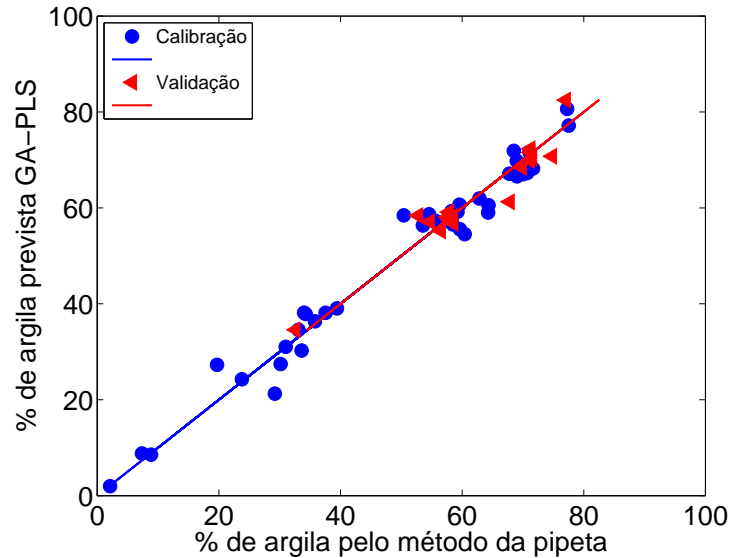


Figura 16 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6, 3x.

A aplicação do GA como técnica de seleção de variáveis possibilitou de modo geral um aprimoramento dos modelos, principalmente, observando os valores de RMSECV. Além disso, em todos os modelos de cores referentes as imagens de 6, 3x, houve uma significativa diminuição dos valores de previsão (RMSEP), de acordo com o teste  $F_{(P=0.05,17,17)}$ , quando comparados com os modelos PLS anteriores sem a utilização de GA, com exceção do espaço de cor HSI. Já nas imagens de 10x somente os modelos referentes à RGB + escala cinza, RGB + HSI e RGB + HSI + escala cinza apresentaram RMSEP com diferença estatística com o modelo anterior sem a aplicação da técnica de seleção de variáveis. Esses resultados são indicativos da viabilidade dessa técnica de seleção à matrizes de imagens.

De acordo com as Tabelas 5 e 9, a utilização do GA, de modo geral, aperfeiçoou os modelos de calibração PLS, principalmente promovendo a diminuição dos valores de RMSEP. Entretanto, somente observou-se diferença estatística entre os valores de RMSEP antes e após a aplicação de GA para os conjuntos RGB, RGB + escala cinza, HSI, HSI +

Tabela 8 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras de solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo GA-PLS							
		n° de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD	
6,3x	RGB	14	3,138	0,9755	3,903	0,9220	10,283	3,109	
	Escala Cinza	9	3,673	0,9673	4,650	0,7835	6,272	2,063	
	RGB + Cinza	10	2,390	0,9862	3,236	0,8694	3,954	2,405	
	HSI	10	5,544	0,9251	5,451	0,8241	12,389	2,171	
	HSI + Cinza	9	3,233	0,9745	3,724	0,9081	5,172	3,079	
	RGB + HSI	10	2,808	0,9804	3,214	0,9270	4,248	3,552	
	RGB + HSI + Cinza	10	2,396	0,9859	2,984	0,9430	3,620	4,161	
10,0x	RGB	13	3,026	0,9767	3,396	0,9022	6,904	2,773	
	Escala Cinza	9	3,511	0,9685	5,831	0,8194	5,317	1,880	
	RGB + Cinza	13	2,782	0,9803	2,862	0,9417	5,665	3,865	
	HSI	12	3,542	0,9664	3,770	0,9439	5,962	3,581	
	HSI + Cinza	10	3,474	0,9694	4,751	0,9004	5,807	2,879	
	RGB + HSI	12	3,299	0,9720	2,923	0,9506	5,435	4,382	
		RGB + HSI + Cinza	12	3,316	0,9721	2,993	0,9265	5,981	3,258



Tabela 9 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras de solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo GA-PLS						
		n° de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
6,3x	RGB	7	4,973	0,9296	6,759	0,8501	6,944	2,517
	Escala Cinza	4	6,279	0,8659	10,914	0,7917	5,251	1,501
	RGB + Cinza	7	3,133	0,9666	3,659	0,9559	6,982	4,386
	HSI	7	4,902	0,9316	5,538	0,9462	5,762	3,070
	HSI + Cinza	8	4,004	0,9544	3,552	0,9612	6,054	4,819
	RGB + HSI	8	3,584	0,9644	3,871	0,9566	4,146	4,382
10,0x	RGB + HSI + Cinza	8	3,019	0,9740	3,412	0,9630	3,745	4,942
	RGB	6	5,497	0,8972	5,175	0,9412	7,450	3,184
	Escala Cinza	8	3,842	0,9498	7,867	0,9281	9,074	2,261
	RGB + Cinza	4	9,321	0,7527	8,672	0,8573	4,997	2,076
	HSI	8	4,639	0,9268	3,905	0,9548	8,201	4,112
	HSI + Cinza	9	3,740	0,9524	4,810	0,9298	7,849	3,772
	RGB + HSI	4	7,194	0,8240	8,440	0,8829	10,851	1,935
	RGB + HSI + Cinza	8	3,930	0,9474	5,073	0,9669	7,230	4,152

escala cinza, RGB + HSI e RGB + HSI + escala cinza das imagens de 6, 3x. Já nas imagens de 10x, houve diferença estatística entre os valores de RMSEP antes e posteriormente a aplicação de GA, somente para os conjuntos escala cinza, HSI, RGB HSI e RGB + HSI + escala cinza.

Conforme as Tabelas 8 e 9, não houve diferença estatística entre os valores de RMSEP determinados utilizando o algoritmo SPXY e o particionamento igualmente espaçado, na divisão de amostras entre calibração e validação, por meio do teste  $F_{(P=0.05,17,17)}$ , com as exceções dos conjuntos RGB e escala cinza das imagens de 6, 3x, e dos conjuntos RGB, escala cinza, HSI + escala cinza e RGB + HSI + escala cinza das imagens de 10x.

Segundo as Tabelas 3, 4, 5, 6, 8 e 9, o modelo de calibração multivariado com o menor valor de RMSEP, aplicado nas predições dos teores de argila, corresponde ao modelo GA-PLS referente aos espaços de cores RGB + escala cinza das imagens de 10x. Entretanto todos os modelos construídos apresentaram valores de RMSEP estatisticamente superiores ao desvio laboratorial de 1,817 % por meio do teste  $F_{(P=0.05,17,17)}$ .

### 6.2.2 Determinação do teor de areia por calibração multivariada

A determinação do teor de areia por calibração multivariada foi realizada de forma semelhante à argila, entretanto, agora o vetor  $y$ , representa as porcentagens de areia. Assim, construíram-se 6 modelos PLS e 6 modelos SPA-RLM para cada aumento óptico e para cada subdivisão entre calibração e validação. As características estatísticas dos modelos constam nas Tabelas 10, 11, 12 e 13.

De acordo com as Tabelas 10 e 12, o modelo PLS com melhor desempenho refere-se ao conjunto RGB + escala cinza das imagens de 10x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 17 apresenta a correlação entre os valores de areia previstos por esse modelo e os valores de areia determinados pelo método de referência. As demais correlações constam no Apêndice F.

Segundo as Tabelas 11 e 13, o modelo SPA-RLM com melhor desempenho refere-se ao conjunto HSI das imagens de 6, 3x, assim a Figura 18 apresenta a correlação entre os valores de areia previstos por esse modelo e os valores de areia determinados pelo método de referência. As demais correlações constam no Apêndice G.

De acordo com a Tabela 10, os valores de RMSEP calculados para cada espaço de cor das imagens de 6, 3x, são iguais estatisticamente por meio do teste  $F_{(P=0.05,17,17)}$ , com

Tabela 10 – Dados dos modelos PLS na predição dos teores de areia em amostras de solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	n° de LVs	Especificações do modelo PLS					
			RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RMSECV	RPD
<b>6,3x</b>	RGB	13	3,126	0,9768	3,533	0,9527	12,921	3,973
	Escala Cinza	1	3,562	0,9703	5,486	0,8459	24,572	2,451
	RGB + Cinza	12	2,418	0,9861	2,704	0,9633	22,253	5,200
	HSI	12	4,020	0,9609	4,920	0,9089	15,802	3,301
	HSI + Cinza	10	5,384	0,9283	5,694	0,9180	14,182	3,242
	RGB + HSI	10	4,900	0,9430	5,000	0,7853	13,548	2,087
	RGB + HSI + Cinza	13	2,874	0,9800	3,313	0,9198	14,268	3,492
<b>10,0x</b>	RGB	9	3,885	0,9634	2,833	0,9269	17,516	3,420
	Escala Cinza	6	2,147	0,9888	2,761	0,9598	13,587	4,731
	RGB + Cinza	4	2,311	0,9740	2,467	0,9501	30,251	4,203
	HSI	8	4,390	0,9549	5,265	0,8334	12,553	2,261
	HSI + Cinza	1	13,085	0,5892	9,546	0,6455	56,742	1,271
	RGB + HSI	6	5,081	0,9384	4,930	0,8862	20,305	2,403
	RGB + HSI + Cinza	6	5,064	0,9388	4,879	0,8852	18,156	2,427

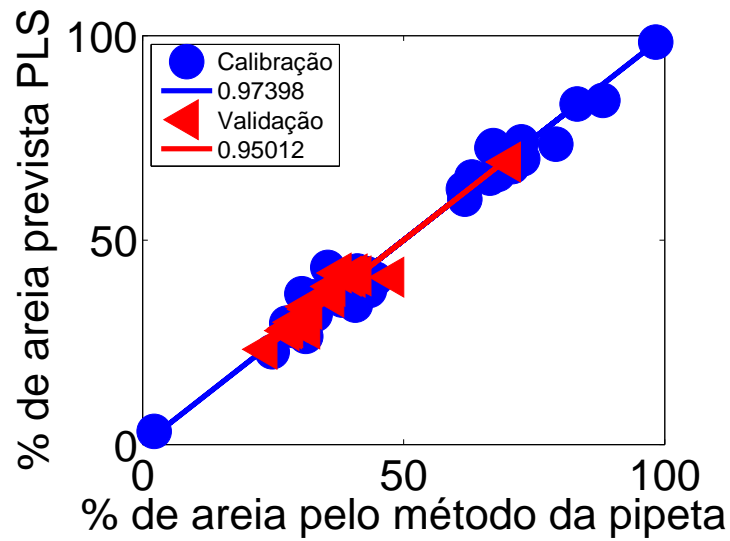


Figura 17 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

exceção dos modelos de cores escala cinza, HSI, HSI + escala cinza e RGB + HSI. Ainda na Tabela 10, observa-se que os RMSEPs para os três modelos de cores, RGB, escala cinza, e RGB + escala cinza das imagens 10x são iguais estatisticamente, por meio do teste  $F_{(P=0.05,17,17)}$ .

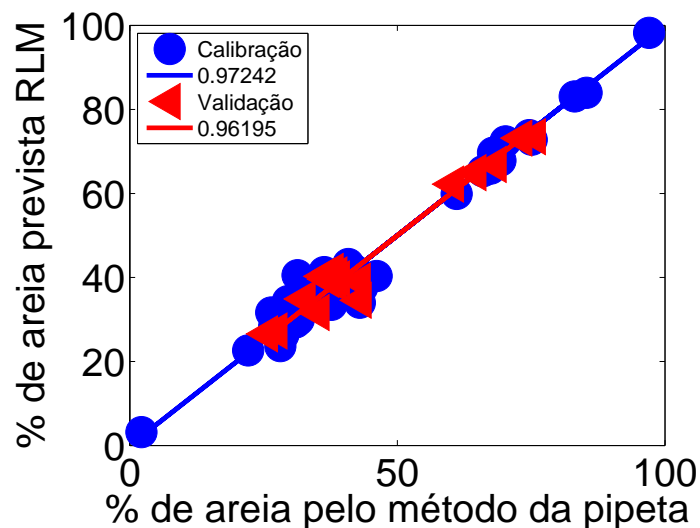


Figura 18 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

Tabela 11 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras de solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RPD
<b>6,3x</b>	RGB	3,998	0,9620	4,606	0,9080	3,152
	Escala Cinza	9,368	0,7948	7,796	0,6655	1,711
	RGB + Cinza	4,019	0,9616	4,832	0,9172	3,432
	HSI	3,377	0,9724	3,088	0,9619	4,987
	HSI + Cinza	2,989	0,9779	5,856	0,9118	2,668
	RGB + HSI	2,641	0,9831	5,193	0,9011	3,140
<b>10,0x</b>	RGB + HSI + Cinza	4,189	0,9575	4,748	0,9024	3,197
	RGB	6,028	0,9031	5,602	0,8196	1,620
	Escala Cinza	5,301	0,9279	8,089	0,6622	1,227
	RGB + Cinza	5,519	0,9277	5,610	0,7256	1,743
	HSI	2,447	0,9856	5,824	0,7995	2,030
	HSI + Cinza	2,972	0,9788	5,041	0,8560	2,427
<b>10,0x</b>	RGB + HSI	3,264	0,9746	4,811	0,8475	2,502
	RGB + HSI + Cinza	4,016	0,9615	5,225	0,8343	2,293

Segundo a Tabela 11 todos os modelos referentes as imagens de 10x apresentam precisão estatisticamente iguais, já nas imagens de 6,3x, somente os modelos: escala cinza, e HSI + escala cinza, apresentaram diferença quanto à precisão de acordo com o teste  $F_{(P=0.05,17,17)}$ .

De acordo com a Tabela 12, não se observou diferença estatística entre os valores de RMSEP dos conjuntos RGB, RGB + escala cinza, HSI, HSI + escala cinza e RGB + HSI + escala cinza das imagens 6,3x, por meio do teste  $F_{(P=0.05,17,17)}$ . Em relação as imagens de 10x, houve diferença estatística entre os valores de RMSEP somente para o modelo RGB + HSI + escala cinza.

Conforme a Tabela 13 todos os modelos referentes as imagens de 6,3x apresentam precisão estatisticamente iguais, com exceção do conjunto RGB + escala cinza. Já nas imagens de 10x, somente os modelos: escala cinza, HSI e RGB + HSI + escala cinza, não apresentaram diferença quanto precisão de acordo com o teste  $F_{(P=0.05,17,17)}$ .

Da mesma forma que para argila, utilizou-se o algoritmo genético para seleção de variáveis na construção do modelo PLS, empregando os mesmos parâmetros apontados na Tabela 7.

Os relatórios de desempenho do método de seleção de variáveis para cada modelo de cor e suas respectivas combinações são apresentados pelas Figuras 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72 e 73, que constam nos Apêndices H e I.

Os parâmetros estatísticos dos modelos GA-PLS constam nas Tabelas 14 e 15.

De acordo com as Tabelas 14 e 15, o modelo GA-PLS com melhor desempenho refere-se ao conjunto RGB + HSI + escala cinza das imagens de 6,3x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 19 apresenta a correlação entre os valores de areia previstos por esse modelo e os valores de areia determinados pelo método de referência. As demais correlações constam no Apêndice J.

Segundo a Tabela 14, houve diferença estatística somente para os modelos RGB, HSI e escala cinza, por meio do teste  $F_{(P=0.05,17,17)}$  referentes as imagens de 6,3x. Já o conjunto de imagens de 10x, somente observou-se diferença estatística para os modelos escala cinza e HSI.

Conforme as Tabelas 10 e 14, a aplicação do GA não ocasionou reduções significativas nos erros de predição no conjunto de imagens de 6,3x, pois somente observou-se diferença estatística entre os valores de RMSEP anteriormente e após a aplicação de GA para o modelo HSI + escala cinza. Já nas imagens de 10x, a aplicação de GA, proporcionou uma

Tabela 12 – Dados dos modelos PLS na predição dos teores de areia em amostras de solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	n° de LVs	Especificações do modelo PLS					
			RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
<b>6,3x</b>	RGB	6	6,821	0,8676	7,889	0,8124	18,747	2,152
	Escala Cinza	2	1,814	0,9918	11,966	0,5475	21,367	1,336
	RGB + Cinza	7	5,414	0,9166	7,195	0,8572	23,229	2,392
	HSI	9	5,568	0,9117	6,747	0,9159	17,701	2,501
	HSI + Cinza	9	5,472	0,9148	5,789	0,9258	17,378	2,950
	RGB + HSI	3	9,282	0,7868	8,650	0,7177	17,599	1,872
	RGB + HSI + Cinza	8	5,683	0,9081	5,515	0,9063	18,833	3,064
<b>10,0x</b>	RGB	8	5,782	0,9049	6,386	0,8679	10,438	2,646
	Escala Cinza	3	1,389	0,9952	12,134	0,6037	21,168	1,310
	RGB + Cinza	6	6,540	0,8783	8,434	0,8536	16,187	2,019
	HSI	8	5,162	0,9241	6,202	0,8691	16,220	2,762
	HSI + Cinza	1	13,085	0,5892	9,546	0,6455	56,742	1,271
	RGB + HSI	3	6,569	0,8932	6,390	0,8469	22,017	2,556
	RGB + HSI + Cinza	10	4,509	0,9421	3,905	0,9653	16,444	4,449

Tabela 13 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras de solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RPD
<b>6,3x</b>	RGB	4,474	0,9504	13,311	0,5522	1,191
	Escala Cinza	2,982	0,9780	16,507	0,406	1,012
	RGB + Cinza	4,019	0,9616	4,832	0,9172	3,432
	HSI	2,729	0,9816	9,053	0,7027	1,792
	HSI + Cinza	3,478	0,9700	13,290	0,6392	1,260
	RGB + HSI	3,663	0,9667	8,869	0,7398	1,832
<b>10,0x</b>	RGB + HSI + Cinza	3,042	0,9771	9,286	0,7672	1,847
	RGB	2,505	0,9844	21,647	0,3369	0,7405
	Escala Cinza	2,971	0,9781	8,964	0,7501	1,823
	RGB + Cinza	2,273	0,9872	16,791	0,4102	0,948
	HSI	2,994	0,9778	10,172	0,65045	1,559
	HSI + Cinza	2,874	0,9796	21,577	0,3240	0,7496
	RGB + HSI	6,569	0,8932	6,390	0,8669	2,556
	RGB + HSI + Cinza	3,532	0,9691	10,588	0,6639	1,499



Tabela 14 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras de solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	n° de LVs	Especificações do modelo GA-PLS					
			RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
<b>6,3x</b>	RGB	9	3,587	0,9709	3,949	0,9053	7,956	3,229
	Escala Cinza	11	4,101	0,9606	4,185	0,9036	7,627	3,210
	RGB + Cinza	10	2,814	0,9815	3,356	0,9462	4,161	4,303
	HSI	12	4,504	0,9506	5,021	0,9118	7,200	3,174
	HSI + Cinza	14	3,158	0,9758	3,520	0,9259	5,058	3,604
	RGB + HSI	13	2,785	0,9809	2,971	0,9649	5,114	5,071
	RGB + HSI + Cinza	8	2,494	0,9853	2,589	0,9609	4,153	4,953
<b>10,0x</b>	RGB	12	3,054	0,9780	3,824	0,8670	12,165	2,524
	Escala Cinza	11	3,956	0,9626	4,420	0,8995	7,850	2,770
	RGB + Cinza	14	2,99	0,9786	3,406	0,9410	4,944	3,970
	HSI	11	4,079	0,9599	4,365	0,9099	6,845	3,208
	HSI + Cinza	10	3,796	0,9665	3,830	0,8449	5,875	2,523
	RGB + HSI	13	3,034	0,9783	3,081	0,9571	5,995	3,871
	RGB + HSI + Cinza	12	3,344	0,9732	2,871	0,9595	5,850	4,316

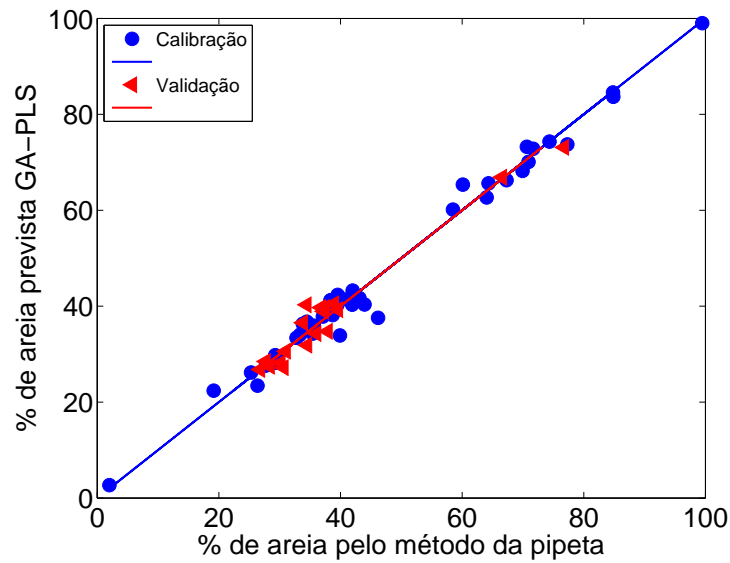


Figura 19 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

redução significativa dos valores de RMSEP para os modelos escala cinza, HSI + escala cinza e RGB + HSI + escala cinza.

Conforme a Tabela 15, os conjuntos RGB, RGB + escala cinza, HSI + escala cinza, RGB + HSI e RGB + HSI + escala cinza pertencentes as imagens de 6, 3x apresentaram precisão estatisticamente iguais através do teste  $F_{(P=0.05,18,18)}$ . Nas imagens de 10x somente o conjunto de RGB apresentou valor de RMSEP diferente dos demais, por meio do teste  $F_{(P=0.05,18,18)}$ .

Comparando os valores de RMSEP dos modelos PLS antes e após a aplicação de GA, Tabelas 15 e 12, observou-se diferença estatística somente para os modelos referentes ao espaço de cor RGB, escala cinza, RGB + escala cinza, RGB + HSI e RGB + HSI + escala cinza pertencentes as imagens 6, 3x, o que indica que nesse caso, a utilização de GA aprimorou, de modo geral, as previsões dos modelos de calibração PLS. Já nas imagens de 10x, houve diferença estatística somente para os modelos referentes aos espaços de cor HSI, RGB + HSI e RGB + HSI e escala cinza. Nesse último caso, a aplicação do GA não aperfeiçoou as previsões dos modelos de calibração PLS.

De acordo com as Tabelas 14 e 15, comparando os valores de RMSEP determinados utilizando o algoritmo SPXY e o particionamento igualmente espaçado, na divisão de amostras entre calibração e validação, observou-se diferença estatística por meio do teste  $F_{(P=0.05,17,18)}$  para todos os modelos de cores, exceto o modelo RGB, referentes as imagens

Tabela 15 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras de solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo GA-PLS					
		n° de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV
<b>6,3x</b>	RGB	8	3,624	0,9605	3,865	0,9454	6,672
	Escala Cinza	7	5,549	0,9074	6,377	0,9088	10,800
	RGB + Cinza	9	2,744	0,9774	2,989	0,9752	8,783
	HSI	7	5,553	0,9073	5,150	0,9219	9,668
	HSI + Cinza	9	3,288	0,9675	3,876	0,9567	5,856
	RGB + HSI	9	2,772	0,9769	2,799	0,9794	7,458
	RGB + HSI + Cinza	9	2,767	0,9770	3,438	0,9655	5,803
<b>10,0x</b>	RGB	9	3,753	0,9634	4,597	0,9332	7,426
	Escala Cinza	3	11,011	0,6355	11,249	0,6552	10,772
	RGB + Cinza	4	10,489	0,7146	9,779	0,8074	5,044
	HSI	2	9,035	0,7546	11,224	0,7681	8,480
	HSI + Cinza	3	8,282	0,7938	10,745	0,7916	11,802
	RGB + HSI	4	8,104	0,8025	10,256	0,8000	9,736
	RGB + HSI + Cinza	10	3,333	0,9711	4,005	0,9534	6,921
							4,282
							2,625
							5,783
							3,318
							4,296
							5,904
							4,901
							3,742
							1,485
							1,878
							1,482
							1,538
							1,612
							4,625

de 6,3x. Em relação as imagens de 10x, não se observou diferença estatística em nenhum modelo, com exceção do espaço de cor escala cinza.

Segundo as Tabelas 10, 11, 12, 13 e 14, o modelo de calibração multivariado com o menor valor de RMSEP, aplicado nas predições dos teores de areia, corresponde ao modelo PLS referente aos espaços de cores RGB + escala cinza das imagens de 10x. Entretanto todos os modelos construídos apresentaram valores de RMSEP estatisticamente superiores ao desvio laboratorial de 0,931 % por meio do teste  $F_{(P=0.05,17,17)}$ .

Finalizada a estimação dos constituintes texturais por calibrações multivariadas, os solos foram classificados quanto à textura (Figura 1) com base nos valores previstos de argila e de areia obtidos pelos modelos PLS e SPA-RLM construídos.

### 6.2.3 Classificação textural das amostras de solo por análise de imagens

Através das estimações de argila e areia previstas pelos modelos PLS, GA-PLS e SPA-RLM, 18 amostras do conjunto de validação foram classificadas quanto à textura e comparadas com a classificação textural (Figura 1) obtida com base nos resultados do método da pipeta. Dessa forma, as Tabelas 16, 17, 18, 19, 20 e 21 apresentam os índices de acertos de cada modelo comparada com a classificação em base na metodologia de referência.

Tabela 16 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Índice de acerto(%)
<b>6,3x</b>	RGB	72,22
	Escala cinza	66,67
	RGB + escala cinza	83,33
	HSI	66,67
	HSI + escala cinza	77,78
	RGB + HSI + escala cinza	83,33
<b>10,0x</b>	RGB	100,00
	Escala cinza	88,89
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

De acordo com as Tabelas 16, 17, 18, 19, 20 e 21, os índices de acertos foram satisfatórios para a maioria dos modelos construídos, principalmente os modelos PLS,

Tabela 17 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	72,22
	Escala cinza	72,22
	RGB + escala cinza	83,33
	HSI	77,78
	HSI + escala cinza	77,78
	RGB + HSI + escala cinza	83,33
<b>10,0x</b>	RGB	100,00
	Escala cinza	88,89
	RGB + escala cinza	100,00
	HSI	88,89
	HSI + escala cinza	77,78
	RGB + HSI + escala cinza	100,00

Tabela 18 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o algoritmo SPXY

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	72,22
	Escala cinza	66,67
	RGB + escala cinza	83,33
	HSI	66,67
	HSI + escala cinza	77,78
	RGB + HSI + escala cinza	83,33
<b>10,0x</b>	RGB	100,00
	Escala cinza	88,89
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

Tabela 16, referentes ao aumento óptico de 10x, onde somente o modelo correspondente a escala cinza não obteve 100,00 % de acerto. Diante desses resultados, observa-se a aplicabilidade dessa metodologia desenvolvida a rotina laboratorial, primeiramente, porque ela é operacional, de baixo custo e não gera resíduo.

Embora, a aplicação do GA tenha aprimorado a maioria dos modelos PLS, nota-se, comparando as Tabelas 16 e 18, e as Tabelas 17 e 19, que não houve um incremento significativo do índice de classificação.

Apesar dos bons resultados obtidos na classificação textural do solo, constataram-se que os modelos constituídos não realizaram boas predições para amostras arenosas (com pelo menos 40% de areia), nem para amostras com altos teores de argila (acima de 77%),

Tabela 19 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	66,67
	Escala cinza	66,67
	RGB + escala cinza	83,33
	HSI	66,67
	HSI + escala cinza	66,67
	RGB + HSI + escala cinza	83,33
<b>10,0x</b>	RGB	100,00
	Escala cinza	66,67
	RGB + escala cinza	83,33
	HSI	83,33
	HSI + escala cinza	83,33
	RGB + HSI + escala cinza	100,00

Tabela 20 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o algoritmo SPXY

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	77,78
	Escala cinza	83,33
	RGB + escala cinza	72,22
	HSI	100,00
	HSI + escala cinza	72,22
	RGB + HSI + escala cinza	66,67
<b>10,0x</b>	RGB	94,44
	Escala cinza	88,89
	RGB + escala cinza	77,78
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

já que essas constituem da minoria das amostras utilizadas, Tabela 2. Essas amostras por suas vezes apresentaram altos valores de *leverage*, ou seja, apresentam alta influência nos modelos, o que, de modo geral, é desfavorável a calibração (FERREIRA et al., 1999). Dessa maneira, realizou-se a remoção dessas amostras, e em seguida, construíram-se os modelos de calibração novamente.

Tabela 21 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Índice de acerto(%)
<b>6,3x</b>	RGB	77,78
	Escala cinza	83,33
	RGB + escala cinza	72,22
	HSI	100,00
	HSI + escala cinza	72,22
	RGB + HSI + escala cinza	66,67
<b>10,0x</b>	RGB	94,44
	Escala cinza	88,89
	RGB + escala cinza	77,78
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

#### 6.2.4 Determinação do teor de argila por calibração multivariada em amostras argilosas

Utilizando somente 41 amostras, correspondentes as amostras argilosas com limite máximo de 77 % de argila, construíram-se 6 modelos PLS e 6 modelos SPA-RLM para cada aumento óptico. De modo análogo a seção 6.2.1, realizaram-se a divisão entre os conjuntos de calibração e validação utilizando o algoritmo SPXY e também realizando o particionamento igualmente espaçado.

As características estatísticas desses modelos constam nas Tabelas 22, 23, 24 e 25.

De acordo com as Tabelas 22 e 24, o modelo PLS com melhor desempenho refere-se ao conjunto RGB das imagens de 10x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 20 apresenta a correlação entre os valores de argila previstos por esse modelo e os valores de argila determinados pelo método de referência. As demais correlações constam no Apêndice K.

Segundo as Tabelas 23 e 25, o modelo SPA-RLM com melhor desempenho refere-se ao conjunto RGB das imagens de 10x, assim a Figura 21 apresenta a correlação entre os valores de argila previstos por esse modelo e os valores de argila determinados pelo método de referência. As demais correlações constam no Apêndice L.

De acordo com a Tabela 22, os conjuntos RGB, escala cinza, RGB + escala cinza, HSI e RGB + HSI + escala cinza referente as imagens de 6,3x apresentaram valores de RMSEP estatisticamente iguais, por meio do teste  $F_{(P=0.05,11,11)}$ . Nas imagens de 10x

Tabela 22 – Dados dos modelos PLS na predição dos teores de argila em amostras argilosas, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	n° de LVs	Especificações do modelo PLS					
			RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
<b>6,3x</b>	RGB	1	1,580	0,9200	2,576	0,8744	7,753	2,267
	Escala Cinza	1	1,334	0,9430	4,074	0,6581	5,340	1,491
	RGB + Cinza	1	1,010	0,9669	2,611	0,8401	6,639	2,218
	HSI	1	2,138	0,8610	1,891	0,8647	6,304	2,682
	HSI + Cinza	14	0,968	0,9712	3,644	0,6225	9,208	1,408
	RGB + HSI	7	1,503	0,9310	2,674	0,7670	4,073	1,943
	RGB + HSI + Cinza	1	1,781	0,9002	2,117	0,8998	5,197	3,131
<b>10,0x</b>	RGB	1	1,749	0,9032	1,438	0,9360	3,983	3,747
	Escala Cinza	1	1,006	0,9858	3,524	0,7885	3,048	2,163
	RGB + Cinza	7	1,538	0,9194	1,828	0,9047	3,088	3,238
	HSI	1	1,309	0,9452	2,843	0,8817	3,665	1,990
	HSI + Cinza	5	1,177	0,9549	2,716	0,8957	4,459	2,075
	RGB + HSI	4	1,175	0,9537	2,923	0,8780	3,391	2,010
	RGB + HSI + Cinza	8	0,833	0,9765	2,821	0,8737	3,344	2,146



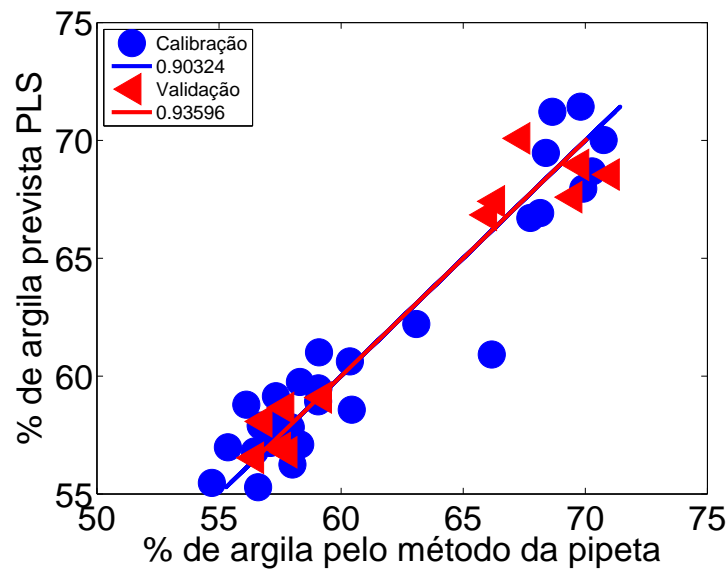


Figura 20 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

somente o conjunto RGB apresentou um valor de RMSEP estatisticamente diferente do restante dos conjuntos.

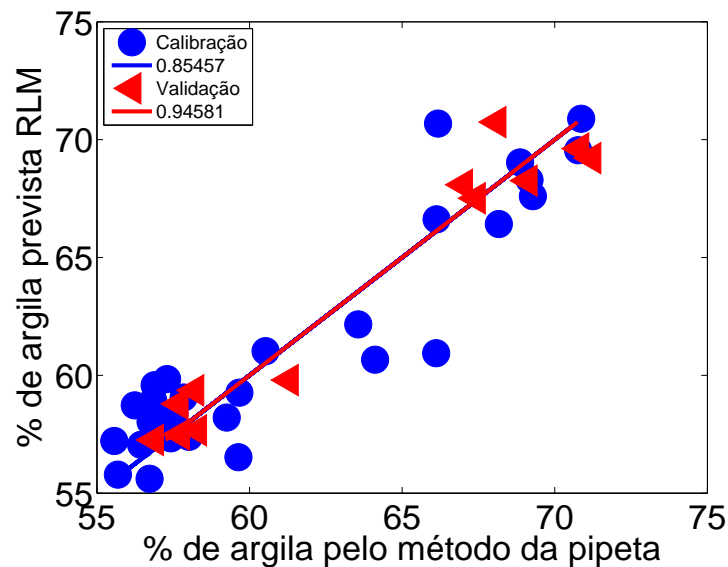


Figura 21 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

Conforme a Tabela 23 somente o modelo HSI referente as imagens de 6, 3x apresentou RMSEP diferente estatisticamente dos demais conjuntos, através do teste  $F_{(P=0.05,11,11)}$ .

Tabela 23 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras argilosas, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RPD
<b>6,3x</b>	RGB	3,198	0,6722	2,581	0,8048	2,181
	Escala Cinza	3,077	0,6965	2,028	0,9105	3,344
	RGB + Cinza	3,327	0,6405	2,467	0,8861	2,311
	HSI	2,090	0,8670	11,510	0,0630	0,470
	HSI + Cinza	3,116	0,7019	2,506	0,7848	2,153
	RGB + HSI	2,861	0,7502	2,686	0,7631	2,004
<b>10,0x</b>	RGB + HSI + Cinza	2,977	0,7214	1,987	0,8856	2,845
	RGB	2,144	0,8546	1,349	0,9458	4,118
	Escala Cinza	2,118	0,8490	1,382	0,9528	4,153
	RGB + Cinza	0,626	0,9866	4,471	0,4545	1,333
	HSI	2,423	0,8125	1,395	0,9435	4,202
	HSI + Cinza	3,274	0,7860	1,731	0,9376	3,274
<b>10,0x</b>	RGB + HSI	2,446	0,7977	2,890	0,8384	2,252
	RGB + HSI + Cinza	3,321	0,7833	2,031	0,8795	2,981

Nas imagens de 10x, somente observou-se diferença estatística no RMSEP referente ao conjunto RGB + escala cinza.

De acordo com a Tabela 24, todos os valores de RMSEP das imagens de 6, 3x apresentaram valores de RMSEP iguais por meio do teste  $F_{(P=0.05,11,11)}$ , exceto o conjunto RGB + HSI + escala cinza. Nas imagens de 10, não se observou diferença estatística nos valores de RMSEP para os conjuntos escala cinza, RGB + escala cinza e HSI.

Comparando os valores de RMSEP para as duas formas empregadas na separação de calibração e validação, Tabela 22 e 24, não se observou diferença estatística somente para o conjunto RGB + HSI + escala cinza referente as imagens de 6, 3x. Nas imagens de 10, não se observou diferença estatística para os conjuntos RGB + escala cinza e HSI + escala cinza, por meio do teste  $F_{(P=0.05,11,11)}$ .

De acordo com a Tabela 25, observou-se diferença estatística nos valores de RMSEP para as imagens de 6, 3x somente para o conjunto RGB + escala cinza. Nas imagens de 10x, somente os conjuntos RGB e RGB + HSI + escala cinza apresentaram valores de RMSEP diferentes, por meio do teste  $F_{(P=0.05,11,11)}$  do restante dos conjuntos.

De acordo com a Tabela 22 todos os valores de RMSEP calculados são estatisticamente iguais ao desvio laboratorial, através do teste  $F_{(P=0.05,11,11)}$ , exceto os valores de RMSEP referente aos modelos de cor na escala cinza e HSI + escala cinza para as imagens de 6, 3x e o modelo RGB + escala cinza para as imagens de 10x. Já de acordo com a Tabela 23, somente o modelo RGB + escala cinza apresentou um valor de RMSEP estatisticamente diferente do desvio laboratorial, por meio do teste  $F_{(P=0.05,11,11)}$ .

Conforme a Tabela 24, todos os valores de RMSEP calculados são estatisticamente diferentes do desvio laboratorial, através do teste  $F_{(P=0.05,11,11)}$ , exceto o valor de RMSEP referente ao modelo RGB + HSI + escala cinza para as imagens de 6, 3x. Nas imagens de 10x, não se observou diferença estatística entre os valores de RMSEP e o desvio laboratorial para os conjuntos escala cinza, HSI e HSI + escala cinza.

Segundo a Tabela 25, nas imagens de 6, 3x, observaram-se diferenças estatísticas através do teste  $F_{(P=0.05,11,11)}$  entre os valores de RMSEP e o desvio laboratorial somente para os conjuntos RGB, RGB + escala cinza, HSI + escala cinza e RGB + HSI + escala cinza. Nas imagens de 10x observaram-se diferenças estatísticas por meio do teste  $F_{(P=0.05,11,11)}$  entre os RMSEPs e o desvio laboratorial somente para os conjuntos RGB e RGB + HSI + escala cinza.

Tabela 24 – Dados dos modelos PLS na predição dos teores de argila em amostras argilosas, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo PLS						
		nº de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
6,3x	RGB	2	1,453	0,9361	5,436	0,5136	6,098	1,100
	Escala Cinza	1	1,406	0,9402	4,741	0,4272	6,659	1,208
	RGB + Cinza	1	1,187	0,9573	7,225	0,4844	5,316	0,741
	HSI	1	2,413	0,8239	4,974	0,3875	7,215	1,072
	HSI + Cinza	1	2,188	0,8552	4,526	0,3671	7,047	1,159
	RGB + HSI	1	1,790	0,9031	2,298	0,8114	4,803	2,302
	RGB + HSI + Cinza	5	1,699	0,9127	2,224	0,8399	5,232	2,487
10,0x	RGB	1	0,925	0,9741	4,112	0,7156	4,977	1,345
	Escala Cinza	1	0,609	0,9888	2,965	0,8795	3,735	1,971
	RGB + Cinza	1	0,549	0,9908	3,805	0,7790	3,882	1,406
	HSI	3	1,545	0,9278	2,427	0,8022	4,052	2,169
	HSI + Cinza	1	1,314	0,9477	2,401	0,8236	3,873	2,181
	RGB + HSI	1	1,478	0,9340	2,851	0,7685	4,476	1,902
	RGB + HSI + Cinza	2	1,222	0,9548	2,515	0,8007	4,232	2,088

Tabela 25 – Dados dos modelos SPA-RLM na predição dos teores de argila em amostras argilosas, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	$r^2_{cal}$	RMSEP	$r^2_{val}$	RPD
<b>6,3x</b>	RGB	2,956	0,7358	3,047	0,6888	1,757
	Escala Cinza	3,013	0,7255	2,783	0,7509	1,907
	RGB + Cinza	3,551	0,6187	4,821	0,1891	1,110
	HSI	3,237	0,6831	2,019	0,8527	2,595
	HSI + Cinza	2,921	0,7419	3,091	0,6787	1,695
	RGB + HSI	2,451	0,8208	3,113	0,6805	1,752
	RGB + HSI + Cinza	2,434	0,8208	3,107	0,6805	1,753
<b>10,0x</b>	RGB	1,448	0,9366	3,228	0,6672	1,622
	Escala Cinza	1,519	0,9302	2,118	0,8497	2,512
	RGB + Cinza	1,816	0,8952	1,910	0,8880	2,744
	HSI	2,280	0,8427	2,113	0,8386	2,484
	HSI + Cinza	2,267	0,8446	2,060	0,8474	2,555
	RGB + HSI	2,009	0,8779	2,207	0,8234	2,371
	RGB + HSI + Cinza	3,318	0,6670	3,230	0,6239	1,627

Além disso, observou-se que a remoção das amostras arenosas e de altos teores de argila ( $> 77\%$ ), possibilitou uma melhora significativa nos modelos, já que, de modo geral, houve uma diminuição dos valores de RMSEP. Comparando as Tabelas 3 e 22, primeiramente para as imagens de 6,3x, observaram-se diferenças estatísticas para todos os valores de RMSEP calculados para os modelos PLS antes e depois da remoção das amostras mencionadas, por meio do teste  $F_{(P=0.05,17,11)}$ . Já para as imagens de 10x, somente houve diferença estatística para os conjuntos escala cinza, RGB + escala cinza e RGB + HSI + escala cinza.

De modo análogo, comparando as Tabelas 4 e 23, observaram-se diferença estatística entre os RMSEPs, através do teste  $F_{(P=0.05,17,11)}$ , somente para o modelo HSI das imagens de 6,3x e para o modelo RGB + escala cinza das imagens de 10x.

A mesma comparação foi realizada entre as Tabelas 5 e 24, assim, observaram-se diferenças estatísticas para os modelos escala cinza e RGB + HSI + escala cinza, ambas pertencentes as imagens de 6,3x, e os modelos escala cinza, HSI e RGB + HSI + escala cinza, pertencentes as imagens de 10x.

Da mesma maneira, comparando as Tabelas 6 e 25, não se observou diferença estatística através do teste  $F_{(P=0.05,17,11)}$  entre os valores de RMSEP somente para o modelo RGB + escala cinza, pertencente as imagens de 10x. Esses resultados indicam que a remoção das amostras arenosas, aprimoraram de modo geral, as previsões dos modelos de calibração.

Do mesmo modo que na seção 6.2.1, modelos PLS empregando algoritmo genético (GA) para seleção de variáveis também foram testados empregando os mesmos parâmetros citados anteriormente, Tabela 7.

Os parâmetros estatísticos dos modelos GA-PLS constam nas Tabelas 26 e 27.

De acordo com as Tabelas 26 e 27, o modelo GA-PLS com melhor desempenho refere-se ao conjunto RGB + HSI + escala cinza das imagens de 10x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 22 apresenta a correlação entre os valores de argila previstos por esse modelo e os valores de argila determinados pelo método de referência. As demais correlações constam no Apêndice M.

Segundo a Tabela 26, todos os modelos dos dois conjuntos de imagens (6,3x e 10x) apresentaram valores de RMSEP iguais estatisticamente, por meio do teste  $F_{(P=0.05,11,11)}$ .

Conforme as Tabelas 22 e 26, a aplicação do GA não ocasionou reduções significativas nos erros de predição no conjunto de imagens de 6,3x, pois não se observou nenhuma

Tabela 26 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras argilosas, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	nº de LVs	Especificações do modelo GA-PLS				
			RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV
6,3x	RGB	9	1,941	0,8817	2,453	0,8902	3,341
	Escala Cinza	7	2,258	0,8349	2,659	0,8034	3,289
	RGB + Cinza	7	2,017	0,8675	1,938	0,8980	3,274
	HSI	7	2,368	0,8230	2,353	0,8348	3,103
	HSI + Cinza	6	2,457	0,8155	2,579	0,8088	3,105
	RGB + HSI	4	2,709	0,7781	2,872	0,7492	3,385
10,0x	RGB + HSI + Cinza	6	1,953	0,8578	2,667	0,8051	3,043
	RGB	8	1,906	0,8414	2,445	0,8558	3,381
	Escala Cinza	8	1,814	0,8443	2,550	0,8858	3,362
	RGB + Cinza	7	1,635	0,8762	2,141	0,8848	3,014
	HSI	12	1,188	0,9424	1,835	0,9401	3,774
	HSI + Cinza	11	1,170	0,9412	2,052	0,9362	3,617
	RGB + HSI + Cinza	9	1,621	0,8860	1,543	0,9566	3,814
							2,437
							2,586
							2,811
							3,515
							3,377
							4,694

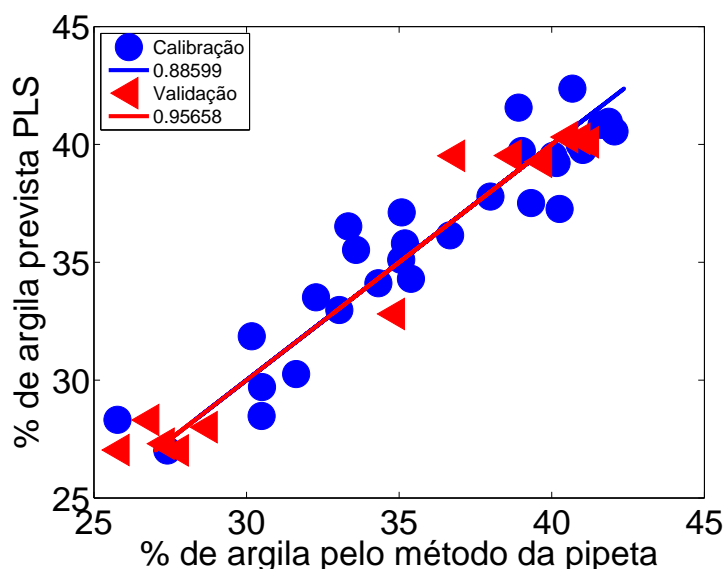


Figura 22 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

diferença estatística entre os valores de RMSEP anteriormente e após a aplicação de GA. Já nas imagens de 10x, a aplicação de GA, resultou em valores de RMSEP diferentes estatisticamente para os modelos RGB, RGB + escala cinza, RGB + HSI + escala cinza.

Segundo a Tabela 27, não se observou diferença estatística para os valores de RMSEP em nenhum modelo dos dois conjuntos de imagens (6, 3x e 10x), por meio do teste  $F_{(P=0.05,11,11)}$ .

Conforme as Tabelas 24 e 27, a aplicação do GA resultou em alterações significativas nos valores de RMSEP, por meio do teste  $F_{(P=0.05,11,11)}$ , para os modelos RGB, RGB + escala cinza, HSI e HSI + escala cinza, pertencentes as imagens de 6, 3x. Nas imagens de 10x, a utilização de GA não alterou de forma significativa os valores de RMSEP, através do teste  $F_{(P=0.05,11,11)}$  para nenhum conjunto.

Comparando os valores de RMSEP determinados pelos modelos PLS utilizando os dois modos de subdivisão de amostras entre calibração e validação, Tabelas 26 e 27, observou-se diferença estatística, somente para o modelo HSI + escala cinza, pertencente as imagens de 10x. Tal resultado, indica que, de modo geral, a potencialidade de previsão dos modelos de calibração não é afetada pela forma de divisão entre calibração e validação.

De acordo com as Tabelas 22, 23, 24, 25, 26 e 27 o modelo de calibração multivariada com o menor valor de RMSEP, aplicado nas predições dos teores de argila em amostras



Tabela 27 – Dados dos modelos GA-PLS na predição dos teores de argila em amostras argilosas, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo GA-PLS							
		nº de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD	
6,3x	RGB	2	2,896	0,7464	2,718	0,7631	3,501	1,955	
	Escala Cinza	3	2,631	0,7907	2,870	0,7420	3,116	1,889	
	RGB + Cinza	3	2,729	0,7747	2,547	0,7783	3,206	2,077	
	HSI	2	3,025	0,7234	2,512	0,7854	3,091	2,097	
	HSI + Cinza	2	3,028	0,7227	2,515	0,7845	3,210	2,094	
	RGB + HSI	6	2,457	0,8155	2,579	0,8088	3,105	2,092	
	RGB + HSI + Cinza	2	2,784	0,7657	2,672	0,7671	3,165	1,963	
10,0x	RGB	3	2,150	0,8602	2,587	0,8014	2,769	2,044	
	Escala Cinza	3	2,157	0,8593	2,284	0,8257	2,813	2,294	
	RGB + Cinza	1	2,644	0,7886	2,728	0,7302	2,933	1,921	
	HSI	5	2,051	0,8727	2,169	0,8428	3,170	2,423	
	HSI + Cinza	3	2,011	0,8777	3,311	0,6795	3,204	1,582	
	RGB + HSI	1	3,350	0,6064	4,115	0,3666	3,229	1,181	
		RGB + HSI + Cinza	7	1,819	0,9000	2,375	0,8435	3,307	2,271

argilosas, corresponde ao modelo SPA-RLM referente ao espaço de cor RGB das imagens de 10x.

### 6.2.5 Determinação do teor de areia por calibração multivariada em amostras argilosas

A determinação dos teores de areia foi realizado de forma semelhante a utilizada na seção 6.2.4, entretanto, utilizou-se o vetor  $y$ , correspondendo as porcentagens de areia. Dessa maneira, construíram-se 6 modelos PLS e 6 modelos SPA-RLM para cada aumento óptico. As características estatísticas desses modelos constam nas Tabelas 28, 29, 30 e 31.

De acordo com as Tabelas 28 e 30, o modelo PLS com melhor desempenho refere-se ao conjunto RGB + escala cinza das imagens de 6,3x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 23 apresenta a correlação entre os valores de areia previstos por esse modelo e os valores de areia determinados pelo método de referência. As demais correlações constam no Apêndice N.

Segundo as Tabelas 29 e 31, o modelo SPA-RLM com melhor desempenho refere-se ao conjunto escala cinza das imagens de 6,3x, assim a Figura 24 apresenta a correlação entre os valores de areia previstos por esse modelo e os valores de areia determinados pelo método de referência. As demais correlações constam no Apêndice O.

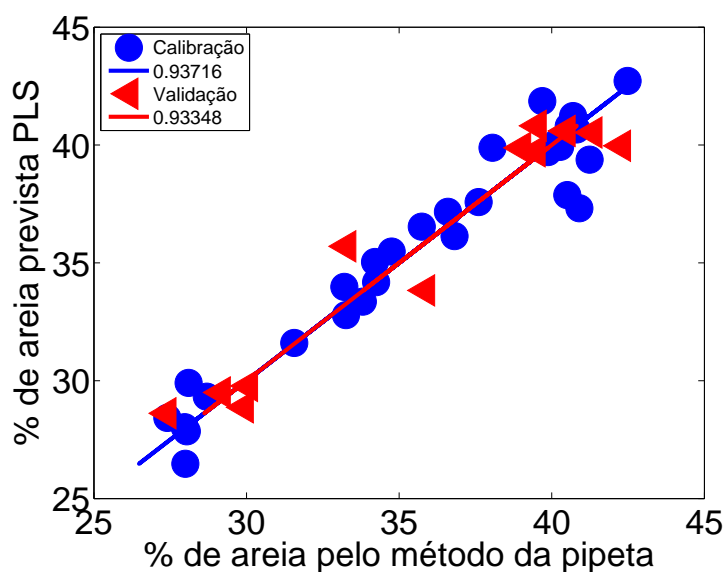


Figura 23 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

Tabela 28 – Dados dos modelos PLS na predição dos teores de areia em amostras argilosas, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo PLS						
		nº de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
6,3x	RGB	8	1,197	0,9432	2,099	0,8750	3,807	2,603
	Escala Cinza	1	1,495	0,9114	2,883	0,7592	4,640	2,038
	RGB + Cinza	1	1,253	0,9371	1,443	0,9334	5,794	3,845
	HSI	8	1,666	0,8919	2,398	0,8799	5,301	2,276
	HSI + Cinza	9	1,776	0,8771	3,706	0,6414	5,414	1,578
	RGB + HSI	1	1,838	0,8683	1,907	0,9377	3,745	2,768
	RGB + HSI + Cinza	8	1,573	0,9067	2,929	0,7621	4,658	1,931
10,0x	RGB	1	1,676	0,8667	2,064	0,9450	4,298	3,325
	Escala Cinza	8	0,287	0,9961	2,421	0,9257	3,077	3,592
	RGB + Cinza	14	0,032	0,9999	2,536	0,8658	5,113	2,625
	HSI	1	2,114	0,8030	1,461	0,9532	6,154	4,283
	HSI + Cinza	14	0,277	0,9967	3,413	0,7624	5,197	1,726
	RGB + HSI	4	1,666	0,8802	2,597	0,8404	4,506	2,329
	RGB + HSI + Cinza	4	1,568	0,8973	2,501	0,8363	4,701	2,364

De acordo com a Tabela 28, os modelos RGB e RGB + escala cinza das imagens de 6,3x apresentaram valores de RMSEP, estatisticamente iguais, por meio do teste  $F_{(P=0.05,11,11)}$ . Nas imagens de 10x, os modelos RGB e HSI apresentaram valores de RMSEP sem diferença estatística.

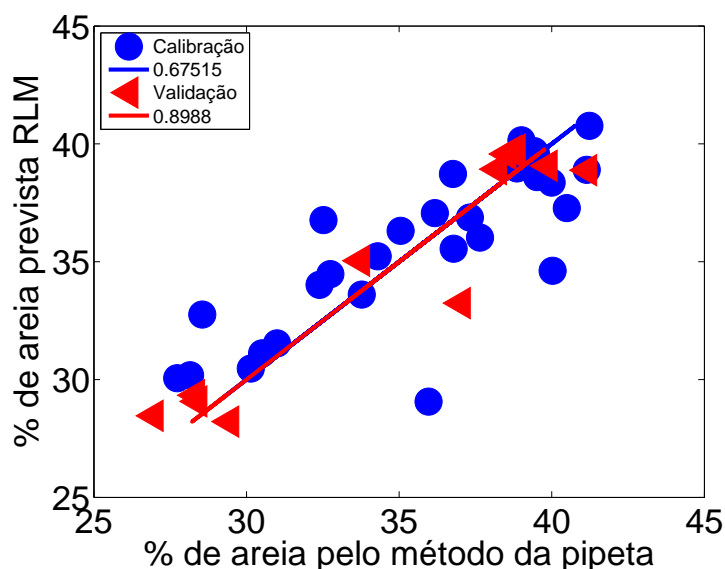


Figura 24 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

Conforme a Tabela 29, todos os valores de RMSEP calculados para as imagens 6,3x são iguais por meio do teste  $F_{(P=0.05,11,11)}$ , exceto os modelos HSI e HSI + escala cinza. Nas imagens de 10x, somente houve diferença estatística entre os valores de RMSEP para os conjuntos RGB e HSI.

De acordo com a Tabela 30, todos os modelos das imagens 6,3x apresentaram RMSEP estatisticamente iguais por meio do teste  $F_{(P=0.05,11,11)}$ . Os modelos escala cinza, HSI, HSI + escala cinza e RGB+ HSI + escala cinza das imagens de 10x apresentaram RMSEP estatisticamente iguais.

De acordo com a Tabela 31, os modelos escala cinza, RGB + escala cinza e HSI apresentaram RMSEP estatisticamente iguais por meio do teste  $F_{(P=0.05,11,11)}$ . Nas imagens de 10x, não houve diferença estatística entre os valores de RMSEP para os modelos escala cinza, RGB + escala cinza, HSI e HSI + escala cinza.

Comparando as Tabelas 28 e 30, especialmente os erros de previsão para os dois modos de subdivisão de amostras, observaram-se diferenças estatísticas, por meio do teste  $F_{(P=0.05,11,11)}$ , entre os valores de RMSEP somente para os conjuntos RGB + escala cinza

Tabela 29 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras argilosas, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RPD
6,3x	RGB	2,759	0,6988	2,821	0,7337	1,935
	Escala Cinza	2,864	0,6751	1,894	0,8988	3,127
	RGB + Cinza	2,787	0,6889	2,658	0,8491	2,199
	HSI	2,322	0,7900	3,179	0,7432	1,884
	HSI + Cinza	1,629	0,8966	3,331	0,6150	1,599
	RGB + HSI	2,342	0,7864	8,799	0,1438	3,717
	RGB + HSI + Cinza	2,429	0,7777	2,621	0,8585	2,177
10,0x	RGB	3,115	0,5397	4,193	0,7150	1,513
	Escala Cinza	2,718	0,6495	2,226	0,9519	2,734
	RGB + Cinza	3,104	0,5396	4,181	0,7150	1,512
	HSI	2,771	0,6615	2,200	0,9211	3,467
	HSI + Cinza	3,118	0,5804	2,660	0,8989	2,316
	RGB + HSI	2,629	0,7018	2,378	0,9377	1,856
	RGB + HSI + Cinza	2,041	0,8973	2,874	0,8363	2,003

Tabela 30 – Dados dos modelos PLS na predição dos teores de areia em amostras argilosas, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	n° de LVs	Especificações do modelo PLS					
			RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
<b>6,3x</b>	RGB	1	1,151	0,9535	2,851	0,6735	3,875	1,680
	Escala Cinza	1	1,369	0,9342	3,241	0,5848	4,984	1,463
	RGB + Cinza	1	1,016	0,9638	4,283	0,5332	3,374	1,110
	HSI	8	1,724	0,8956	4,586	0,4147	4,881	1,087
	HSI + Cinza	10	1,331	0,9378	3,292	0,6164	4,628	1,446
	RGB + HSI	4	1,335	0,9374	3,325	0,6515	4,327	1,514
<b>10,0x</b>	RGB + HSI + Cinza	1	1,224	0,9474	3,546	0,5811	4,364	1,390
	RGB	1	1,364	0,9348	6,679	0,4821	4,497	0,755
	Escala Cinza	7	0,265	0,9975	3,844	0,6356	4,696	1,311
	RGB + Cinza	1	0,4339	0,9933	4,748	0,6067	5,004	1,031
	HSI	1	1,279	0,9426	2,916	0,6520	4,673	1,639
	HSI + Cinza	3	1,198	0,9496	2,460	0,7555	4,679	1,928
	RGB + HSI	1	1,340	0,9370	2,907	0,6476	5,022	1,665
	RGB + HSI + Cinza	3	1,039	0,9621	2,539	0,7205	5,122	1,887

Tabela 31 – Dados dos modelos SPA-RLM na predição dos teores de areia em amostras argilosas, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo SPA-RLM				
		RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RPD
6,3x	RGB	2,172	0,8345	7,047	0,2914	0,680
	Escala Cinza	2,982	0,6881	3,625	0,4858	1,310
	RGB + Cinza	2,398	0,7823	3,619	0,5312	1,357
	HSI	1,835	0,8819	4,434	0,4331	1,189
	HSI + Cinza	1,132	0,9553	16,949	0,0002	0,2982
	RGB + HSI	1,129	0,9553	16,949	0,0002	0,2983
10,0x	RGB + HSI + Cinza	1,126	0,9553	17,041	0,0002	0,2694
	RGB	1,725	0,8956	5,527	0,0845	0,874
	Escala Cinza	2,564	0,7693	3,220	0,6196	1,558
	RGB + Cinza	2,242	0,8237	3,881	0,4279	1,227
	HSI	1,737	0,8941	3,023	0,6583	1,578
	HSI + Cinza	1,871	0,8771	3,382	0,6099	1,450
	RGB + HSI	1,871	0,8771	3,382	0,6099	1,450
	RGB + HSI + Cinza	0,837	0,9754	5,907	0,4719	0,833

e HSI, pertencentes as imagens de 6,3x. Já em relação as imagens de 10x, observaram-se diferença estatística somente para os modelos RGB, RGB + escala cinza e HSI.

Do mesmo modo, comparando as Tabelas 29 e 31, não foi observado diferença estatística entre os valores de RMSEP somente para os modelos RGB + escala cinza e HSI das imagens de 6,3x. Nas imagens de 10x, somente houve diferença estatística entre os RMSEP dos modos de divisão de amostras para os modelos RGB + escala cinza e RGB + HSI + escala cinza.

Do mesmo modo que na seção 6.2.4 utilizou-se o GA para seleção de variáveis nos modelos PLS, utilizando os mesmos parâmetros, Tabela 7.

As características estatísticas dos modelos GA-PLS constam nas Tabelas 32 e 33.

De acordo com as Tabelas 32 e 33, o modelo GA-PLS com melhor desempenho refere-se ao conjunto RGB + HSI + escala cinza das imagens de 10x, no qual foi empregado a subdivisão de amostras pelo algoritmo SPXY, assim a Figura 25 apresenta a correlação entre os valores de areia previstos por esse modelo e os valores de areia determinados pelo método de referência. As demais correlações constam no Apêndice P.

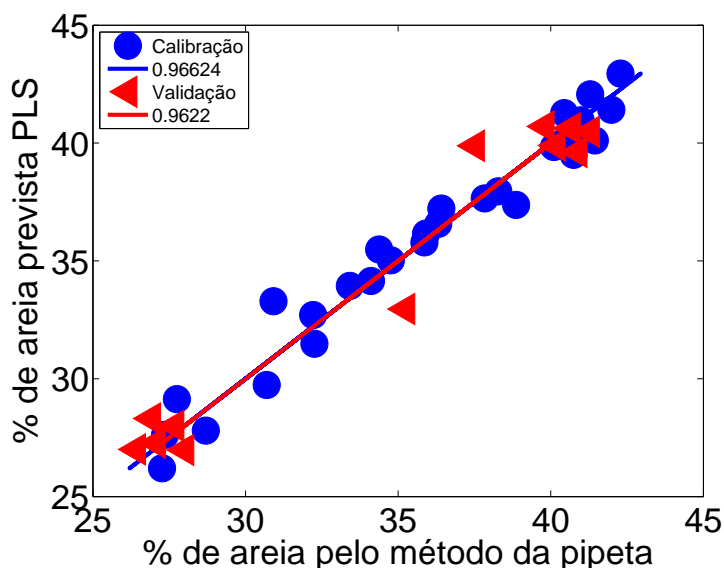


Figura 25 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10x.

Segundo a Tabela 32, nos dois conjuntos de imagens (6,3x e 10x), observou-se diferença estatística somente para os modelos RGB + escala cinza, por meio do teste

$$F_{(P=0.05,11,11)}.$$



Tabela 32 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras argilosas, utilizando o algoritmo SPXY

Aumento óptico	Modelo de cor	Especificações do modelo GA-PLS					
		nº de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RPD
<b>6,3x</b>	RGB	10	1,675	0,8688	2,112	0,8930	3,618
	Escala Cinza	8	1,813	0,8719	2,692	0,8383	3,790
	RGB + Cinza	6	1,930	0,8231	2,375	0,9019	3,385
	HSI	8	1,995	0,8592	2,335	0,7523	3,330
	HSI + Cinza	10	2,025	0,8408	2,761	0,8086	3,574
<b>10,0x</b>	RGB + HSI + Cinza	9	1,572	0,8979	1,900	0,8917	3,327
	RGB	9	1,562	0,8934	2,250	0,8586	3,381
	Escala Cinza	9	1,543	0,8873	2,330	0,8728	3,362
	RGB + Cinza	8	1,272	0,9250	2,487	0,8498	3,014
	HSI	10	1,519	0,9058	1,617	0,9336	3,774
<b>10,0x</b>	HSI + Cinza	10	1,353	0,9213	2,042	0,9474	3,617
	RGB + HSI + Cinza	11	0,882	0,9662	1,409	0,9622	3,814
							4,830

Conforme as Tabelas 28 e 32, a aplicação do GA não ocasionou reduções significativas nos erros de predição no conjunto de imagens de 6, 3x, pois não se observou nenhuma diferença estatística entre os valores de RMSEP anteriormente e após a aplicação de GA. Já nas imagens de 10x, a aplicação de GA, proporcionou uma redução significativa dos valores de RMSEP para o modelo RGB + HSI + escala cinza.

Segundo a Tabela 33, não se observou diferença estatística em relação aos dados de RMSEP em nenhum dos modelos dos dois conjuntos de imagens (6, 3x e 10x), por meio do teste  $F_{(P=0.05,11,11)}$ .

Conforme as Tabelas 28 e 33 a aplicação do GA não ocasionou reduções significativas nos erros de predição no conjunto de imagens de 6, 3x, pois não se observou nenhuma diferença estatística entre os valores de RMSEP anteriormente e após a aplicação de GA. Já nas imagens de 10x, a aplicação de GA, proporcionou uma redução significativa dos valores de RMSEP para o modelo RGB, através do teste  $F_{(P=0.05,11,11)}$ .

Comparando os valores de RMSEP determinados pelos modelos PLS utilizando os dois modos de subdivisão de amostras entre calibração e validação, Tabelas 32 e 33, observaram-se diferença estatística, por meio do teste  $F_{(P=0.05,11,11)}$ , somente para os modelos RGB, RGB + escala cinza, RGB + HSI + escala cinza, pertencente as imagens de 6, 3x, e nos modelos HSI e RGB + HSI + escala cinza, ambas pertencentes as imagens de 10x.

De acordo com as Tabelas 28, 29, 30, 31, 32 e 33 o melhor modelo de calibração multivariada com o menor valor de RMSEP, aplicado nas predições dos teores de areia em amostras argilosas, corresponde ao modelo GA-PLS referente aos espaços de cores RGB + HSI + escala cinza das imagens de 10x. Esse modelo GA-PLS juntamente com o modelo PLS construído com a combinação dos histogramas de frequência dos espaços de cores RGB e escala cinza (RGB + escala cinza) referente as imagens de 6, 3x foram os únicos modelos que apresentaram valores de RMSEP iguais estatisticamente ao desvio laboratorial, por meio do teste  $F_{(P=0.05,11,17)}$ .

### 6.2.6 Classificação textural das amostras de solo por análise de imagens

Posteriormente a remoção das amostras arenosas e amostras com altos teores de argila, construíram-se modelos SPA-RLM e PLS para estimação da areia e da argila.

Tabela 33 – Dados dos modelos GA-PLS na predição dos teores de areia em amostras argilosas, utilizando o particionamento igualmente espaçado

Aumento óptico	Modelo de cor	Especificações do modelo GA-PLS						
		nº de LVs	RMSEC	r <sup>2</sup> <sub>cal</sub>	RMSEP	r <sup>2</sup> <sub>val</sub>	RMSECV	RPD
6,3x	RGB	2	2,629	0,7575	4,067	0,4191	3,067	1,200
	Escala Cinza	1	2,723	0,7398	4,305	0,3730	2,910	1,113
	RGB + Cinza	1	2,738	0,7369	4,182	0,3791	2,933	1,146
	HSI	4	2,494	0,7819	3,629	0,5048	2,912	1,315
	HSI + Cinza	4	2,444	0,7905	3,603	0,5246	2,944	1,329
	RGB + HSI	1	3,318	0,6138	4,846	0,2554	3,229	1,101
10,0x	RGB + HSI + Cinza	4	2,464	0,7870	4,079	0,4143	3,102	1,169
	RGB	3	2,313	0,8123	3,331	0,6408	2,994	1,494
	Escala Cinza	5	2,163	0,8359	3,464	0,6041	3,035	1,385
	RGB + Cinza	3	2,454	0,7887	3,381	0,5458	3,141	1,412
	HSI	8	1,895	0,8740	2,772	0,6794	3,121	1,713
	HSI + Cinza	5	2,144	0,8387	3,063	0,6380	3,365	1,558
	RGB + HSI	9	2,398	0,7982	3,487	0,5260	4,106	1,367
	RGB + HSI + Cinza	8	1,419	0,9294	2,952	0,6906	3,347	1,607

Nessa direção, 12 amostras do conjunto de validação foram classificadas quanto à textura (Figura 1) e comparadas com a classificação textural determinada com base nos resultados do método da pipeta. Assim, as Tabelas 34, 35, 36, 37, 38 e 39 apresentam os índices de acertos de cada modelo comparada com a classificação em base na metodologia de referência.

Tabela 34 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o algoritmo SPXY

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	91,67
	Escala cinza	91,67
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	91,67
	RGB + HSI + escala cinza	91,67
<b>10,0x</b>	RGB	100,00
	Escala cinza	100,00
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

Tabela 35 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o algoritmo SPXY

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	91,67
	Escala cinza	91,67
	RGB + escala cinza	83,33
	HSI	83,33
	HSI + escala cinza	83,33
	RGB + HSI + escala cinza	100,00
<b>10,0x</b>	RGB	91,67
	Escala cinza	100,00
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

De acordo com as Tabelas 34, 35, 36, 37, 38 e 39, os índices de acertos foram satisfatórios, com todos modelos apresentando índices maiores que 65 %. Diante disso, observa-se que a remoção das amostras pouco representativas possibilitou um aumento do índice de acerto na classificação textural, indicando a potencialidade desse método à

Tabela 36 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o algoritmo SPXY

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	91,67
	Escala cinza	91,67
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	91,67
	RGB + HSI + escala cinza	91,67
<b>10,0x</b>	RGB	100,00
	Escala cinza	100,00
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

Tabela 37 – Índices de acertos dos modelos PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	91,67
	Escala cinza	91,67
	RGB + escala cinza	91,67
	HSI	83,33
	HSI + escala cinza	91,67
	RGB + HSI + escala cinza	91,67
<b>10,0x</b>	RGB	83,33
	Escala cinza	83,33
	RGB + escala cinza	91,67
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

rotina laboratorial, o que possibilitaria um aumento na produtividade e uma diminuição significativa do custo por análise, assim como a minimização dos resíduos gerados.

Embora, a metodologia proposta tenha apresentado resultados satisfatórios, nota-se uma deficiência nas previsões de amostras com caráter arenoso e com altos teores de argila, devido principalmente a pouca representabilidade dessas amostras no conjunto total de amostras. Em contra partida, ressalta-se que, como a maioria das amostras de solo requisitadas à EMBRAPA para análise de textura é de caráter argiloso, normalmente pertencente a faixa de maior grupamento de amostras (41) utilizadas neste trabalho, tais modelos podem apresentar utilização prática.

Tabela 38 – Índices de acertos dos modelos SPA-RLM na classificação textural do solo, utilizando o particionamento igualmente espaçado

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	83,33
	Escala cinza	91,67
	RGB + escala cinza	83,33
	HSI	66,66
	HSI + escala cinza	66,66
	RGB + HSI + escala cinza	100,00
<b>10,0x</b>	RGB	91,67
	Escala cinza	100,00
	RGB + escala cinza	100,00
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	91,67

Tabela 39 – Índices de acertos dos modelos GA-PLS na classificação textural do solo, utilizando o particionamento igualmente espaçado

<b>Aumento óptico</b>	<b>Modelo de cor</b>	<b>Índice de acerto(%)</b>
<b>6,3x</b>	RGB	91,67
	Escala cinza	91,67
	RGB + escala cinza	91,67
	HSI	83,33
	HSI + escala cinza	91,67
	RGB + HSI + escala cinza	91,67
<b>10,0x</b>	RGB	91,67
	Escala cinza	91,67
	RGB + escala cinza	91,67
	HSI	100,00
	HSI + escala cinza	100,00
	RGB + HSI + escala cinza	100,00

## 7 Conclusão

Modelos de calibração PLS, GA-PLS e SPA-RLM foram construídos com o intuito de estimar os teores de argila e areia em amostras de solo. Como etapa de validação, esses modelos foram utilizados na estimação da constituição textural de 18 amostras com a finalidade de mensurar a qualidade de cada modelo. O modelo de calibração GA-PLS, referente aos espaços de cores RGB + escala cinza das imagens de 10x, apresentou os menores valores de RMSEP nas predições de argila e areia. Entretanto, esses valores de RMSEP diferem estatisticamente dos desvios laboratoriais através do teste  $F_{(P=0.05,17,17)}$ , o que indica que a metodologia desenvolvida não tem a mesma precisão do que a metodologia de referência. No entanto, com base nos valores preditos por esse modelo, as amostras de validação foram classificadas quanto à textura e comparadas com a classificação textural obtida segundo os resultados do método da pipeta, obtendo-se um índice de acerto de 100 %.

Após a remoção das amostras não representativas do conjunto total, constituídas por solos arenosos e de altos teores de argila e que também apresentavam altos resíduos, realizaram-se novas estimações através dos modelos PLS, GA-PLS e SPA-RLM. Assim, estimaram-se os teores de argila e areia de um conjunto de 12 amostras pertencentes ao grupamento de validação. Os modelos de calibração SPA-RLM, referente ao espaço de cor RGB das imagens de 10x, e o modelo GA-PLS, referente aos espaços de cores RGB + HSI + escala cinza, apresentaram os menores valores de RMSEP respectivamente para as predições de argila e areia. Tais valores de RMSEP são estatisticamente iguais aos desvios laboratoriais através do teste  $F_{(P=0.05,11,17)}$ , o que indica que a exatidão do método proposto é igual a da metodologia de referência. Assim, com base nos valores preditos por esses modelos, as amostras de validação foram classificadas quanto à textura e comparadas com a classificação textural obtida segundo os resultados do método da pipeta, obtendo-se um índice de acerto de 100 %.

Esses resultados, indicam a potencialibilidade da metodologia desenvolvida e sua possível aplicação na rotina laboratorial, visto que, ela é de baixo custo, já que não utiliza nenhum tipo de reagente, nem de equipamentos de alto custo de aquisição, já o método da pipeta custa, por amostra, em média R\$ 10 reais, e ainda necessita de um conjunto de vidrarias e reagentes. Em segundo lugar, o método proposto não gera resíduo, diferente da metodologia de referência, em que cada análise se produz cerca de 1 L de resíduo básico, que

necessita de tratamento antes de descartado. Mais ainda, a nova técnica possibilitará um aumento significativo da produtividade, pois será possível analisar cerca de 120 amostras por dia, ou seja, praticamente 4 vezes mais do que a metodologia usual.

Apesar dos bons resultados obtidos, a otimização dos modelos foi realizada removendo as amostras arenosas e de altos teores de argila, devido as suas baixas representabilidades no conjunto de amostras. Desse modo, com a finalidade de aperfeiçoar os modelos, faz-se necessário ampliar o número de amostras de natureza arenosa e de alto teor de argila ao conjunto de dados e, em seguida, realizar novas validações. Entretanto ressalta-se que, como a grande maioria das amostras de solo requisitadas à EMBRAPA para análise de textura é de caráter argiloso, pertencente a faixa de maior concentração de amostras utilizadas nesse trabalho, tais modelos podem apresentar utilização prática.



## Referências<sup>1</sup>

- ALBUQUERGUE, M. P. et al. *Análise de Imagens e Visão Computacional*. Rio de Janeiro: [s.n.], 2004. 145–176 p. Disponível em: <<http://www.cbpf.br/~mpa/G7-marcio.pdf>>. Citado na página 40.
- ALCIATURI, C. et al. Aplicación de métodos matemáticos y estadísticos en la industria. *Revista Técnica de la Facultad de Ingeniería Universidad del Zulia*, v. 33, n. 2, p. 164–168, 2010. ISSN 02540770. Citado na página 50.
- ALEXANDRE, C.; SILVA, J. R. M. D.; FERREIRA, A. G. Comparação de dois métodos de determinação da textura do solo: sedimentometria por raios X vs. método da pipeta. *Revista de Ciências Agrárias*, v. 24, n. 3 e 4, p. 99–110, 2001. Citado 2 vezes nas páginas 39 e 40.
- ALMEIDA, V. E. de et al. Using color histograms and spa-lda to classify bacteria. *Analytical and bioanalytical chemistry*, Springer, v. 406, n. 24, p. 5989–5995, 2014. Citado na página 35.
- ALVES, G. M. *Método fundamentado em Processamento Digital de Imagens para Contagem Automática de Unidades Formadoras de Colônias*. 125 p. Tese (Doutorado) — Universidade Federal de São Carlos, 2006. Citado na página 47.
- ANDERSEN, C. M.; BRO, R. Variable selection in regression-a tutorial. *Journal of Chemometrics*, v. 24, n. 11-12, p. 728–737, 2010. ISSN 08869383. Citado 2 vezes nas páginas 59 e 60.
- BARBOSA, J. M. G. *Paralelismo em Processamento e Análise de Imagens Médicas*. 224 p. Tese (Tese de doutorado) — Universidade do Porto, 2000. Citado 2 vezes nas páginas 31 e 46.
- BARROS-NETO, B. de; SCARMINIO, I. S.; BRUNS, R. E. *Como fazer experimentos, Pesquisa e Desenvolvimento na Ciência e na Indústria*. 2. ed. Campinas: Unicamp, 2001. 412 p. ISBN 8526805444. Citado na página 49.
- BILLINGSLEY, F. C. Digital image processing for information extraction. *International Journal of Man-Machine Studies*, Elsevier, v. 5, n. 2, p. 203–214, 1973. Citado na página 35.
- BORIN, A. et al. Quantification of lactobacillus in fermented milk by multivariate image analysis with least-squares support-vector machines. *Analytical and bioanalytical chemistry*, Springer, v. 387, n. 3, p. 1105–1112, 2007. Citado na página 35.
- BOTVINICK, E. H. et al. Noninvasive diagnosis of a false left ventricular aneurysm with radioisotope gated cardiac blood pool imaging: differentiation from true aneurysm. *The American journal of cardiology*, Elsevier, v. 37, n. 7, p. 1089–1093, 1976. Citado na página 34.
- BREITKREITZ, M. C.; SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para planejamento de experimentos: avaliação das condições

<sup>1</sup> De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

experimentais na determinação espectrofotométrica de Ferro II com o-fenantrolina. Um tutorial, parte III. *Quim. Nova*, v. 37, n. 3, p. 564–573, 2014. Disponível em: <<http://dx.doi.org/10.5935/0100-4042.20140092>>. Citado na página 49.

BRERETON, R. G. *Applied chemometrics for scientists*. Chichester: John Wiley & Sons, 2007. 1–379 p. ISBN 9780470016862. Citado na página 56.

CONAB. *Acompanhamento da Safra Brasileira de Grãos 2015/16*. 2015. 152 p. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 23 dez. 2015. Citado na página 30.

CORÁ, J. E. et al. Adição de areia para dispersão de solos na análise granulométrica. *R. Bras. Ci. Solo*, v. 33, n. 1, p. 255–262, 2009. Citado 4 vezes nas páginas 30, 31, 38 e 40.

CORREIA, A. *Liderança agrícola deve aumentar influência internacional do Brasil*. 2009. 3 p. Disponível em: <[http://www.bbc.co.uk/portuguese/noticias/2009/03-/090331\\_brasil\\_agricultura\\_ac.shtml](http://www.bbc.co.uk/portuguese/noticias/2009/03-/090331_brasil_agricultura_ac.shtml)>. Acesso em: 22 jun. 2015. Citado na página 30.

COSTA, C. *Educação, imagem e mídias*. [S.l.]: Cortez, 2005. 198 p. Citado na página 34.

COSTA FILHO, P. A.; POPPI, R. J. Algoritmo Genético em Química. *Quim. Nova*, v. 22, n. 3, p. 405–411, 1999. Citado 2 vezes nas páginas 59 e 60.

CRUZ, T. G. da. *Monitoramento por microscopia óptica e processamento digital de imagens do processo de conformação cerâmica por conformação com amidos comerciais*. 1–128 p. Tese (Doutorado) — Universidade Estadual Paulista, 2007. Citado na página 47.

DAMASCENO, D. et al. Análise multivariada de imagens na química: Um experimento para determinação do pH de águas potáveis. *Quim. Nova*, v. 38, n. 6, p. 836–841, 2015. Citado na página 36.

DEMATTE, J. A. et al. Quantificação de matéria orgânica do solo através de modelos matemáticos utilizando colorimetria no sistema munsell de cores. *Bragantia*, SciELO Brasil, v. 70, n. 3, p. 590–597, 2011. Citado na página 36.

DINIZ, M. R. B. R. S. et al. Medida da cor em solos do rio grande do sul com a carta de munsell e por colorimetria. *Ciência Rural*, SciELO Brasil, v. 36, n. 4, 2006. Citado na página 36.

DINIZ, P. H. et al. Using a simple digital camera and spa-lda modeling to screen teas. *Analytical Methods*, Royal Society of Chemistry, v. 4, n. 9, p. 2648–2652, 2012. Citado na página 35.

DOURADO, T. A.; SILVA, L. F. S. d.; MARINHO, M. d. A. Performance of a reciprocal shaker in mechanical dispersion of soil samples for particle-size analysis. *Revista Brasileira de Ciência do Solo*, SciELO Brasil, v. 36, n. 4, p. 1131–1148, 2012. Citado na página 39.

ERIKSSON, L. et al. Multivariate analysis of aquatic toxicity data with pls. *Aquatic Sciences*, Springer, v. 57, n. 3, p. 217–241, 1995. Citado na página 54.

Estado de Minas. *Brasil deve assumir liderança na exportação de produtos agrícolas a partir de 2024*. 2015. 1 p. Disponível em: <[http://www.em.com.br/app/noticia/economia-/2015/11/09/internas\\_economia,705850/oferta-mundial-nas-maos-do-brasil.shtml](http://www.em.com.br/app/noticia/economia-/2015/11/09/internas_economia,705850/oferta-mundial-nas-maos-do-brasil.shtml)>. Acesso em: 23 dez. 2015. Citado na página 30.

FEARN, T. Assessing calibrations: Sep, rpd, rer and r2. *NIR news*, v. 13, n. 6, p. 12–14, 2002. Citado na página 58.

FERREIRA, M. M. C. et al. Quimiometria I: Calibração Multivariada, Um tutorial. *Quim. Nova*, v. 22, n. 5, p. 724–731, 1999. Citado 7 vezes nas páginas 49, 50, 51, 52, 56, 57 e 91.

FIGUEIREDO, D. *Conceitos Básicos de Sensoriamento Remoto*. 2005. 32 p. Disponível em: <[http://www.conab.gov.br/conabweb/download/SIGABRASIL/manuais/conceitos\\_sm.pdf](http://www.conab.gov.br/conabweb/download/SIGABRASIL/manuais/conceitos_sm.pdf)>. Citado 2 vezes nas páginas 44 e 47.

FLOR, E. P. O. et al. Avaliação de danos mecânicos em sementes de soja por meio da análise de imagens. *Revista Brasileira de Sementes*, v. 26, n. 1, p. 68–76, 2004. Citado na página 35.

FREITAS, R. E.; MENDONÇA, M. A. A. d.; LOPES, G. d. O. *Expansão de área agrícola: perfil e desigualdade entre as mesorregiões brasileiras*. Instituto de Pesquisa Econômica Aplicada (Ipea), 2014. Disponível em: <[http://ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_1926.pdf](http://ipea.gov.br/portal/images/stories/PDFs/TDs/td_1926.pdf)>. Acesso em: 22 jun. 2015. Citado 2 vezes nas páginas 30 e 31.

GALVÃO FILHO, A. R. *Calibração multivariada empregando o algoritmo das projeções sucessivas com reamostragem e combinação de modelos*. 76 p. Dissertação (Mestrado) — Instituto Tecnológico de Aeronáutica, 2010. Citado na página 61.

GALVÃO, R. K. H. et al. A method for calibration and validation subset partitioning. *Talanta*, v. 67, n. 4, p. 736–740, 2005. ISSN 00399140. Citado 3 vezes nas páginas 52, 53 e 65.

GALVÃO, R. K. H. et al. Cross-Validation for the Selection of Spectral Variables Using the Sucessive Projections Algorithm. *J. Braz. Chem. Soc.*, v. 18, n. 8, p. 1585–1589, 2007. Citado 2 vezes nas páginas 57 e 61.

GARCÍA, J.; MEDINA, H.; NÚÑEZ, D. El método del hidrómetro: influencia de los tiempos de lecturas en el cálculo de la distribución del tamaño de partículas en suelos de la habana. *Cultivos Tropicales*, Ediciones INCA, v. 29, n. 2, p. 21–26, 2008. Citado na página 39.

GASQUES, J. G. et al. Produtividade da agricultura: resultados para o brasil e estados selecionados. *Revista de Política Agrícola*, v. 23, n. 3, p. 87–98, 2014. Citado na página 30.

GELADI, P. et al. Image analysis in chemistry i. properties of images, greylevel operations, the multivariate image. *TRAC Trends In Analytical Chemistry*, Elsevier, v. 11, n. 1, p. 41–53, 1992. Citado na página 34.

GELADI, P.; GRAHN, H. *Multivariate image analysis*. Chichester: John Wiley & Sons, 1996. 316 p. ISBN 0471930016. Citado 7 vezes nas páginas 32, 40, 41, 42, 43, 45 e 47.

GELADI, P.; WOLD, S.; ESBENSEN, K. Image analysis and chemical information in images. *Analytica Chimica Acta*, Elsevier, v. 191, p. 473–480, 1986. Citado na página 35.

GENRICH, D.; BREMNER, J. A reevaluation of the ultrasonic-vibration method of dispersing soils. *Soil Science Society of America Journal*, Soil Science Society of America, v. 36, n. 6, p. 944–947, 1972. Citado na página 39.

GODINHO, M. S. et al. Classificação de refrigerantes através de análise de imagens e análise de componentes principais (PCA). *Quim. Nova*, v. 31, n. 6, p. 1485–1489, 2008. Citado 2 vezes nas páginas 35 e 47.

GOLDSTEIN, M. D.; NAGLER, M. Real time inspection of a large set of surface defects in metal parts. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Robotics and IECON'87 Conferences*. [S.l.], 1988. p. 184–190. Citado na página 35.

GOMES, J. F. S.; LETA, F. R. Applications of computer vision techniques in the agriculture and food industry: a review. *European Food Research and Technology*, Springer, v. 235, n. 6, p. 989–1000, 2012. Citado na página 47.

GOMES, M. S. et al. Uso de scanner em espectrofotometria de absorção molecular: aplicação em experimento didático enfocando a determinação de ácido ascórbico. *Quim. Nova*, v. 31, n. 6, p. 1577–1581, 2008. Citado na página 36.

GOMES, O. F. M. *Processamento e Análise de Imagens Aplicados à Caracterização Automática de Materiais*. 141 p. Dissertação (Dissertação de mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 2001. Citado 2 vezes nas páginas 31 e 32.

GONZÁLEZ, H. M.; CORONADO, J. G.; ACOSTA, D. N. El método del hidrómetro: base teórica para su empleo en la determinación de la distribución del tamaño de partículas de suelo. *Revista Ciencias Técnicas Agropecuarias*, Universidad Agraria de La Habana Fructuoso Rodríguez Pérez, v. 16, n. 3, p. 19–24, 2007. Citado 2 vezes nas páginas 38 e 40.

GONZALEZ, R. C.; WOODS, R. E. *Processamento de Imagens Digitais*. 1. ed. São Paulo: Editora Edgard Blücher Ltda, 2003. 509 p. Citado 9 vezes nas páginas 31, 32, 41, 43, 44, 45, 47, 48 e 49.

GOSSET, W. S. *The Application of the 'Law of Error' to the Work of the Brewery*. *Guinness Laboratory Report 8*. [S.l.]: Arthur Guinness & Son, Ltd., Guinness Archives, 1904. Citado na página 49.

GUOHUA, W. Z. Z.; GUOAN, L. Partial least squares method (pls) and its application in analytical chemistry. *Chinese Journal of Analytical Chemistry*, v. 7, p. 24, 1989. Citado na página 54.

HUGHES, J. F. et al. *Computer Graphics Principles and Practice*. 3. ed. Reading, MA: Addison-Wesley, 2013. 1262 p. ISBN 9780321399526. Citado 2 vezes nas páginas 43 e 44.

IBÁÑEZ-ASENSIO, S. et al. Statistical relationships between soil colour and soil attributes in semiarid areas. *Biosystems Engineering*, Elsevier, v. 116, n. 2, p. 120–129, 2013. Citado na página 36.

IBRAHEEM, N. a. et al. Understanding Color Models : A Review. *ARPJN Journal of Science and Technology*, v. 2, n. 3, p. 265–275, 2012. Citado na página 44.

IGNE, B.; HURBURGH, C. R. Local chemometrics for samples and variables: optimizing calibration and standardization processes. *Journal of Chemometrics*, Wiley Online Library, v. 24, n. 2, p. 75–86, 2010. Citado na página 58.

KENNARD, R. W.; STONE, L. A. Computer Aided Design Experiments. *Technometrics*, v. 11, n. 1, p. 137–148, 1969. Disponível em: <[http://www.jstor.org/stable/10.2307-1266219](http://www.jstor.org/stable/10.2307/1266219)>. Citado na página 52.

KLEIN, V. A. et al. Textura do solo e a estimativa do teor de água no ponto de murcha permanente com psicrômetro. *Ciência Rural*, v. 40, n. 7, p. 1550–1556, 2010. Citado na página 38.

KOURTI, T. Application of latent variable methods to process control and multivariate statistical process control in industry. *International Journal of Adaptive Control and Signal Processing*, Wiley Online Library, v. 19, n. 4, p. 213–246, 2005. Citado na página 54.

KOURTI, T.; LEE, J.; MACGREGOR, J. F. Experiences with industrial applications of projection methods for multivariate statistical process control. *Computers & chemical engineering*, Elsevier, v. 20, p. S745–S750, 1996. Citado na página 54.

KOWALSKI, B. R. Chemometrics: views and propositions. *Journal of Chemical Information and Computer Sciences*, ACS Publications, v. 15, n. 4, p. 201–203, 1975. Citado na página 49.

KOWALSKI, B. R. Chemometrics. *American Chemical Society*, ACS Publications, v. 52, n. 5, p. 112R–122R, 1980. Citado na página 49.

KRAMER, R. *Chemometric techniques for quantitative analysis*. New York: Marcel Dekker, 1998. 1–203 p. ISBN 0824701984. Citado 3 vezes nas páginas 49, 50 e 52.

LANDIM, R. Brasil já é o terceiro maior exportador agrícola do mundo. *Jornal Estadão, Mar*, 2010. Disponível em: <<http://economia.estadao.com.br/noticias/geral,brasil-ja-e-o-terceiro-maior-exportador-agricola-do-mundo,520500>>. Acesso em: 22 jun. 2015. Citado na página 30.

LEDLEY, R. S. High-speed automatic analysis of biomedical pictures. *Science*, American Association for the Advancement of Science, v. 146, n. 3641, p. 216–223, 1964. Citado na página 34.

LENARDÃO, E. J. et al. Green chemistry: the 12 principles of green chemistry and its insertion in the teach and research activities. *Química Nova*, SciELO Brasil, v. 26, n. 1, p. 123–129, 2003. Citado na página 31.

LI, B.; MORRIS, J.; MARTIN, E. B. Model selection for partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, v. 64, p. 79–89, 2002. Citado na página 54.

LU, N. *Fractal Imaging*. San Diego: Academic Press, 1997. 412 p. ISBN 0124580106. Citado 3 vezes nas páginas 42, 43 e 48.

MARQUES FILHO, O.; VIEIRA NETO, H. *Processamento Digital de Imagens*. Rio de Janeiro: Brasport, 1999. 331 p. ISSN 0102261X. ISBN 8574520098. Citado 8 vezes nas páginas 31, 32, 43, 44, 46, 47, 48 e 49.

MARTENS, H.; NÆS, T. *Multivariate Calibration*. Chichester: John Wiley & Sons, 1991. 419 p. ISBN 0471909793. Citado 7 vezes nas páginas 51, 53, 54, 55, 56, 57 e 58.

MAURI, J. et al. Dispersantes químicos na análise granulométrica de Latossolos. *R. Bras. Ci. Solo*, v. 35, n. 1, p. 1277–1284, 2011. Citado 5 vezes nas páginas 30, 31, 38, 39 e 40.

MENESES, F. G. A. de. *Segmentação de imagens de arte rupestre utilizando o sistema de coordenadas estelares*. 1–58 p. Tese (Doutorado) — Universidade Federal Fluminense, 2013. Citado na página 47.

MORITSUKA, N. et al. Soil color analysis for statistically estimating total carbon, total nitrogen and active iron contents in japanese agricultural soils. *Soil Science and Plant Nutrition*, Taylor & Francis, v. 60, n. 4, p. 475–485, 2014. Citado na página 36.

MURTI, G. K.; SATYANARAYANA, K. Influence of chemical characteristics in the development of soil colour. *Geoderma*, Elsevier, v. 5, n. 3, p. 243–248, 1971. Citado na página 36.

NOUWEN, J. et al. Classification of environmentally occurring chemicals using structural fragments and pls discriminant analysis. *Environmental science & technology*, ACS Publications, v. 31, n. 8, p. 2313–2318, 1997. Citado na página 54.

OTTO, M. *Chemometrics: statistics and computer application in analytical chemistry*. Weinheim: Wiley-VCH, 1999. 1–314 p. ISBN 352729628X. Citado 4 vezes nas páginas 50, 51, 55 e 58.

PACIORNIK, S. et al. Scanner image analysis in the quantification of mercury using spot-tests. *Journal of the Brazilian Chemical Society*, SciELO Brasil, v. 17, n. 1, p. 156–161, 2006. Citado na página 36.

PAL, N. R.; PAL, S. K. A review on image segmentation techniques. *Pattern recognition*, Elsevier, v. 26, n. 9, p. 1277–1294, 1993. Citado na página 35.

PANSU, M.; GAUTHEYROU, J. *Handbook of Soil Analysis Mineralogical, organic and inorganic methods*. New York: Springer, 2003. 995 p. ISBN 9783540312109. Citado 3 vezes nas páginas 31, 40 e 64.

PATACA, L. C. M. *Análises de mel e própolis utilizando métodos Quimiométricos de Classificação e Calibração*. 97 p. Tese (Tese de doutorado) — Universidade Estadual de Campinas, 2006. Citado na página 35.

PETROU, M.; PETROU, C. *Image Processing : The Fundamentals*. 2. ed. Chichester: John Wiley & Sons, 2010. 815 p. ISBN 9780470745861. Citado 3 vezes nas páginas 32, 47 e 48.

PIMENTEL, M. F.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. d. Recomendações para calibração em química analítica parte 2: calibração multianalito. *Química Nova*, SciELO Brasil, v. 31, n. 2, p. 462–467, 2008. Citado na página 50.

PINTO, F. et al. Utilização de um sistema de visão artificial para estimativa do teor de umidade de frutos de café. *Revista Brasileira de Armazenamento*, v. 12, n. 6, p. 30–36, 2003. Citado na página 35.

PLATANOTIS, K. N.; VENETSANOPOULOS, A. N. *Color Image Processing and Applications*. Berlin Heidelberg: Springer-Verlag, 2000. 359 p. ISBN 978-3-642-08445-4. Citado na página 44.

PM, N.; CHEZIAN, R. M. VARIOUS COLOUR SPACES AND COLOUR SPACE CONVERSION. *Journal of Global Research in Computer Science*, v. 4, n. 1, p. 44–48, 2013. Citado na página 46.

PORTAL PLANALTO. *Com R\$ 187 bilhões, governo federal lança maior Plano Safra da história do Brasil*. 2015. Disponível em: <<http://www2.planalto.gov.br/noticias/2015/06/com-r-187-bilhoes-governo-federal-lanca-maior-plano-safra-da-historia-do-brasil>>. Acesso em: 22 jun. 2015. Citado na página 31.

PRADO, A. G. S. Química verde, os desafios da química do novo milênio. *Química nova*, v. 26, n. 5, p. 738–744, 2003. Citado na página 31.

PRATS-MONTALBÁN, J.; JUAN, a. de; FERRER, a. Multivariate image analysis: A review with applications. *Chemometrics and Intelligent Laboratory Systems*, Elsevier B.V., v. 107, n. 1, p. 1–23, maio 2011. ISSN 01697439. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169743911000529>>. Citado 2 vezes nas páginas 32 e 34.

RAMOS, M. *Um breve ensaio sobre a fotografia e a leitura crítica do discurso fotográfico*. 2007. 3 p. Disponível em: <<http://www.studium.iar.unicamp.br/23/03.html>>. Acesso em: 19 jan. 2016. Citado na página 34.

RISEMAN, E. M.; ARBIB, M. A. Computational techniques in the visual segmentation of static scenes. *Computer Graphics and Image Processing*, Elsevier, v. 6, n. 3, p. 221–276, 1977. Citado na página 35.

RODIONOVA, O. Y.; POMERANTSEV, a. L. Chemometrics: achievements and prospects. *Russian Chemical Reviews*, v. 75, n. 4, p. 271–287, 2007. ISSN 0036-021X. Citado 2 vezes nas páginas 49 e 50.

ROSSEL, R. V.; FOUAD, Y.; WALTER, C. Using a digital camera to measure soil organic carbon and iron contents. *Biosystems Engineering*, Elsevier, v. 100, n. 2, p. 149–159, 2008. Citado na página 36.

RUIZ, H. A. Incremento da exatidão da análise granulométrica do solo por meio da coleta da suspensão (silte + argila). *R. Bras. Ci. Solo*, v. 29, n. 1, p. 297–300, 2005. Citado 2 vezes nas páginas 30 e 39.

RUSS, J. C. *The Image Processing Handbook*. sixth. Boca Raton: CRC Press, 2011. 853 p. ISBN 9781439840634. Citado 2 vezes nas páginas 40 e 48.

RUSS, J. C.; RUSS, J. C. *Introduction to Image Processing and Analysis*. Boca Raton: CRC Press, 2008. 347 p. ISBN 0849370736. Citado na página 40.

SANTOS, A. D. et al. *Manual de análises químicas de solos, plantas e fertilizantes*. 2. ed. Brasília, DF: Embrapa, 2009. 627 p. ISBN 978-85-7383-430-7 1. Citado na página 30.

SCHOWENGERDT, R. A. *Remote Sensing: Models and Methods for Image Processing*. 2. ed. London: Academic Press, 1997. 522 p. Citado na página 44.

SHIH, T. K. et al. An Intelligent Content-based Image Retrieval System Based on Color, Shape and Spatial Relations. *Proc. Natl. Sci. Counc.*, v. 25, n. 4, p. 232–243, 2001. Citado na página 44.

- SILVA, E. M. et al. Comparação de modelos matemáticos para o traçado de curvas granulométricas. *Pesq. agropec. bras.*, v. 39, n. 1, p. 363–370, 2004. Citado na página 38.
- SILVA, T. C. d. O.; GODINHO, M. S.; OLIVEIRA, A. E. Identification of pale lager beers via image analysis. *Latin American applied research*, SciELO Argentina, v. 41, n. 2, p. 141–145, 2011. Citado na página 35.
- SJÖSTRÖM, M. et al. A multivariate calibration problem in analytical chemistry solved by partial least-squares models in latent variables. *Analytica Chimica Acta*, Elsevier, v. 150, p. 61–70, 1983. Citado na página 54.
- SOUZA, R. V. B.; CELLIGOI, A. Avaliação da condutividade hidráulica do solo em área agrícola e florestada na cidade de Londrina (PR) através do permeâmetro Guelph. *Boletim de Geografia*, v. 29, n. 2, p. 123–133, abr. 2012. ISSN 2176-4786. Citado na página 38.
- STRAUSS, H. W. et al. A scintiphotographic method for measuring left ventricular ejection fraction in man without cardiac catheterization. *The American journal of cardiology*, Elsevier, v. 28, n. 5, p. 575–580, 1971. Citado na página 34.
- TAVARES FILHO, J. A.; MAGALHAES, F. S. Dispersão de amostras de latossolo Vermelho eutroférico influenciadas por pré-tratamento para oxidação da matéria orgânica e pelo tipo de agitação mecânica. *R. Bras. Ci. Solo*, v. 32, n. 3, p. 1429–1435, 2008. Citado 2 vezes nas páginas 31 e 39.
- TKALCIC, M.; TASIC, J. Colour spaces- perceptual, historical and applicational background. In: *The IEEE Region 8 EUROCON 2003. Computer as a Tool*. [S.l.: s.n.], 2003. v. 1, p. 304–308. ISBN 0-7803-7763-X. Citado na página 46.
- VAZ, C.; NAIME, J.; MACEDO, A. Análise da textura de solos por atenuação de raios gama. *SIMPÓSIO NACIONAL DE INSTRUMENTAÇÃO AGROPECUÁRIA*, Embrapa/CNPEDIA São Carlos, 1997. Citado na página 39.
- VERHAAR, H. J. et al. Modelling the toxicity of organophosphates: a comparison of the multiple linear regression and pls regression methods. *Quantitative Structure-Activity Relationships*, Wiley Online Library, v. 13, n. 2, p. 133–143, 1994. Citado na página 54.
- VILAR, W. T. et al. Classification of individual castor seeds using digital imaging and multivariate analysis. *Journal of the Brazilian Chemical Society*, SciELO Brasil, v. 26, n. 1, p. 102–109, 2015. Citado na página 35.
- WILLIAMS, P.; NORRIS, K. et al. *Near-infrared technology in the agricultural and food industries*. [S.l.]: American Association of Cereal Chemists, Inc., 1987. Citado na página 58.
- WOLD, H. Estimation of principal components and related models by iterative least squares. In: KRISHNAIAH, P. R. (Ed.). *Multivariate Analysis*. [S.l.]: Academic Press, 1966. p. 391–420. Citado na página 54.
- WOLD, H. Nonlinear estimation by iterative least square procedures. In: DAVID, F. N. (Ed.). *Research Papers in Statistics*. [S.l.]: Wiley, 1966. p. 411–444. Citado na página 54.
- WOLD, S. et al. Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, v. 44, n. 1-2, p. 175–185, 1998. ISSN 01697439. Citado na página 51.



WOLD, S.; MARTENS, H.; WOLD, H. The multivariate calibration problem in chemistry solved by the pls method. In: LECTURE NOTES IN MATHEMATICS. *Proceedings of the Conference on Matrix Pencils*. [S.l.]: Springer, 1983. p. 286–293. Citado na página 54.

## Apêndice A – Modelos de regressão por mínimos quadrados PLS para predição dos teores de argila em amostras de solo

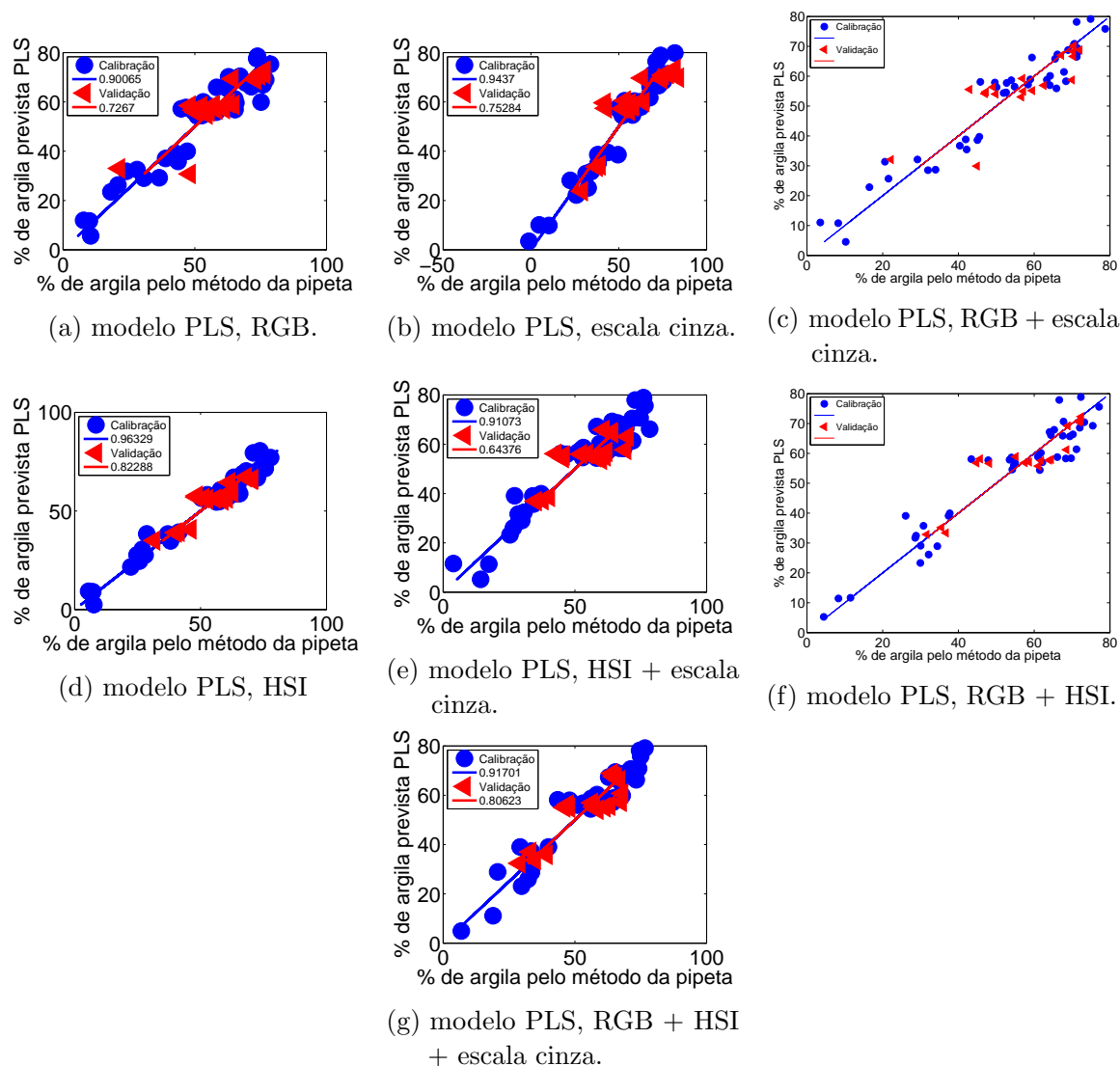


Figura 26 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

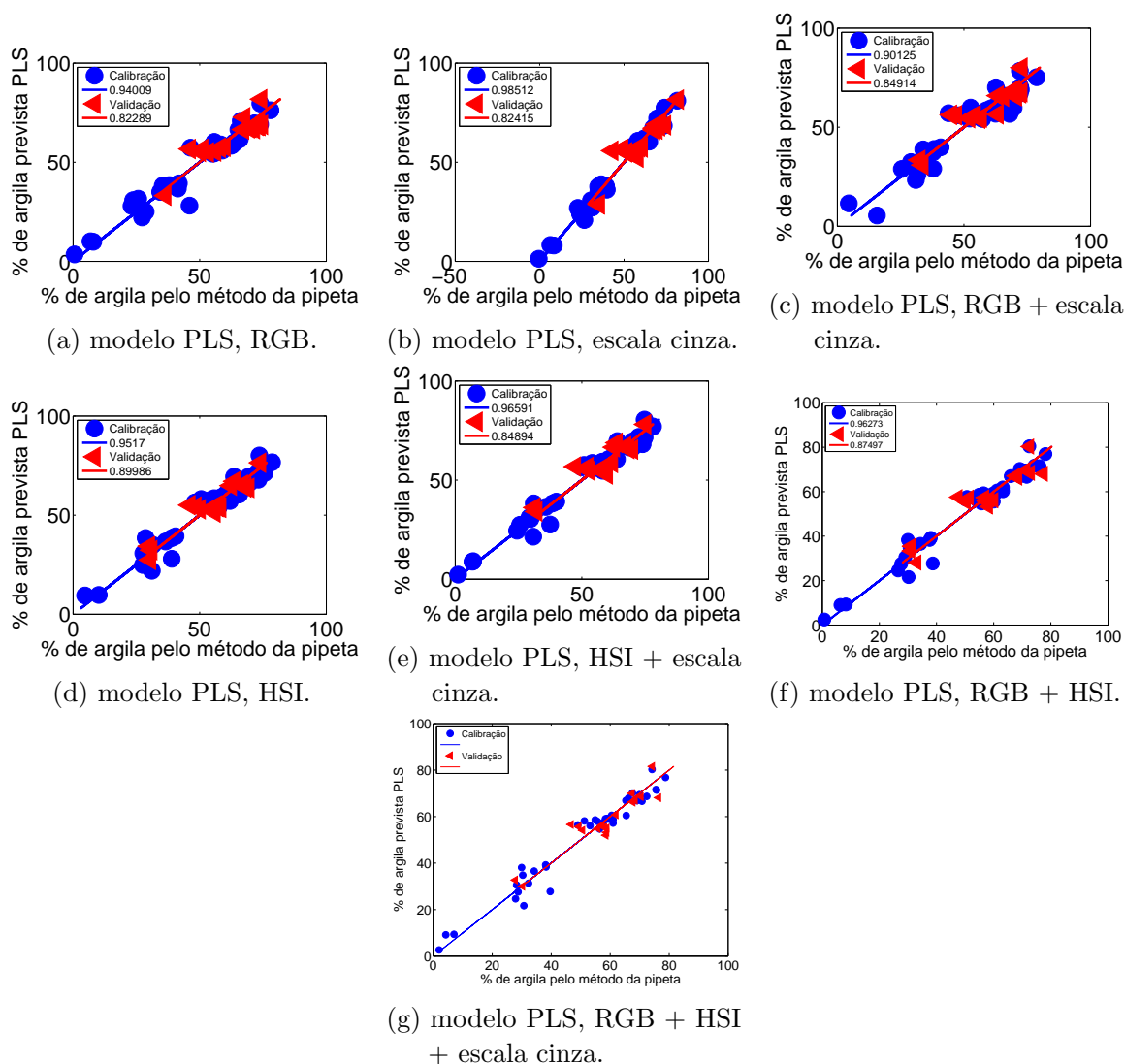


Figura 27 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

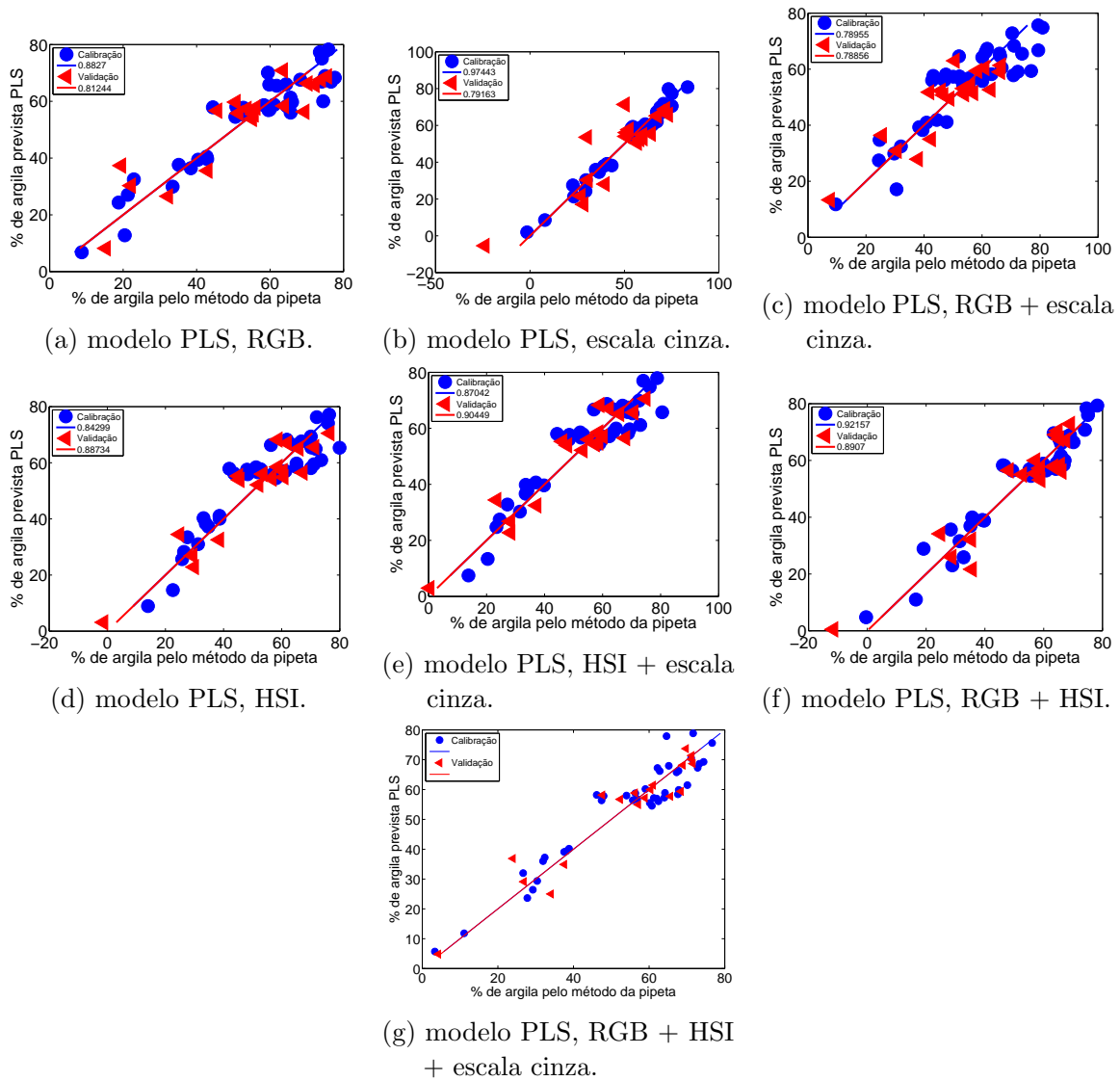


Figura 28 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.

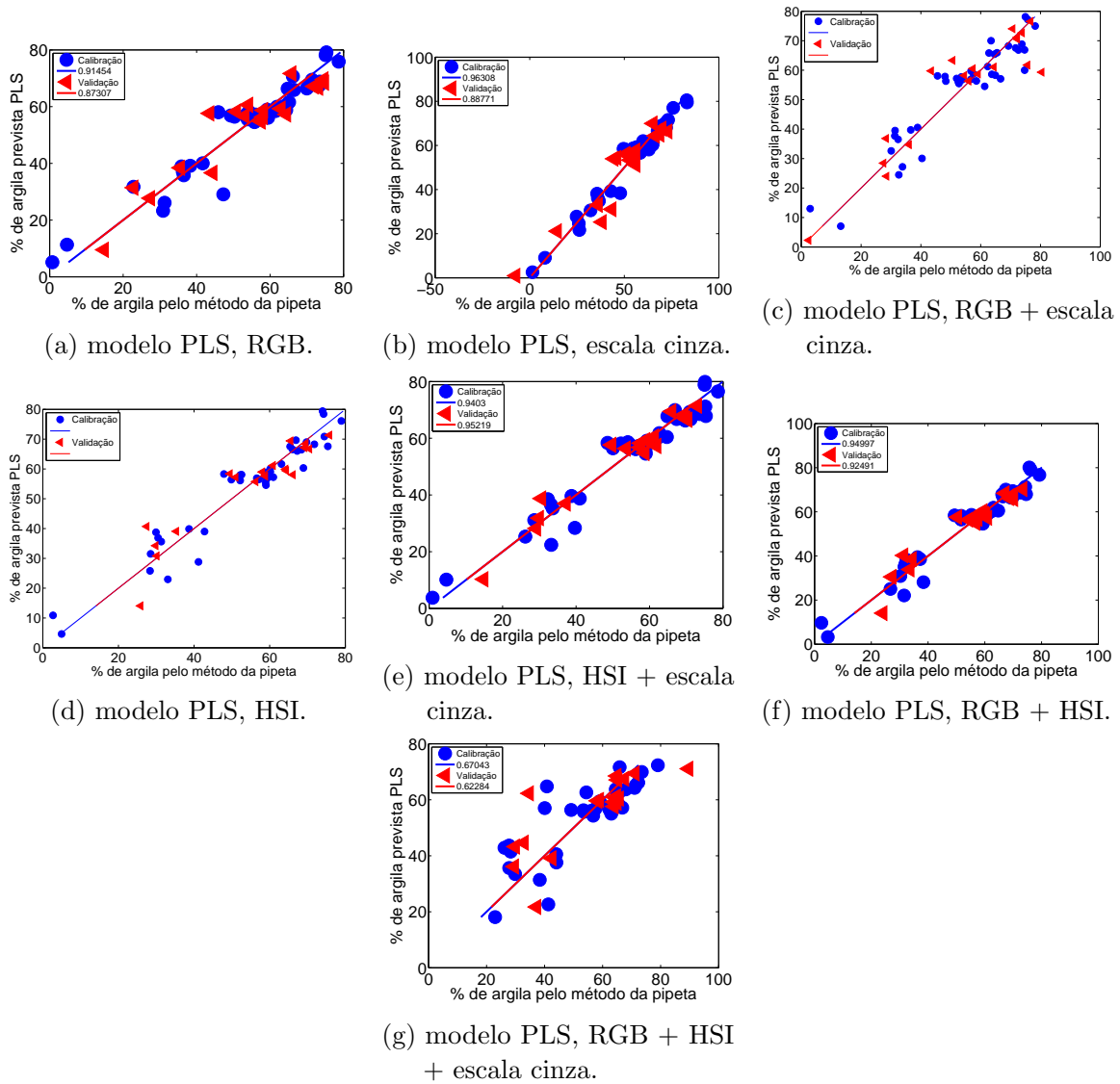


Figura 29 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice B – Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de argila em amostras de solo

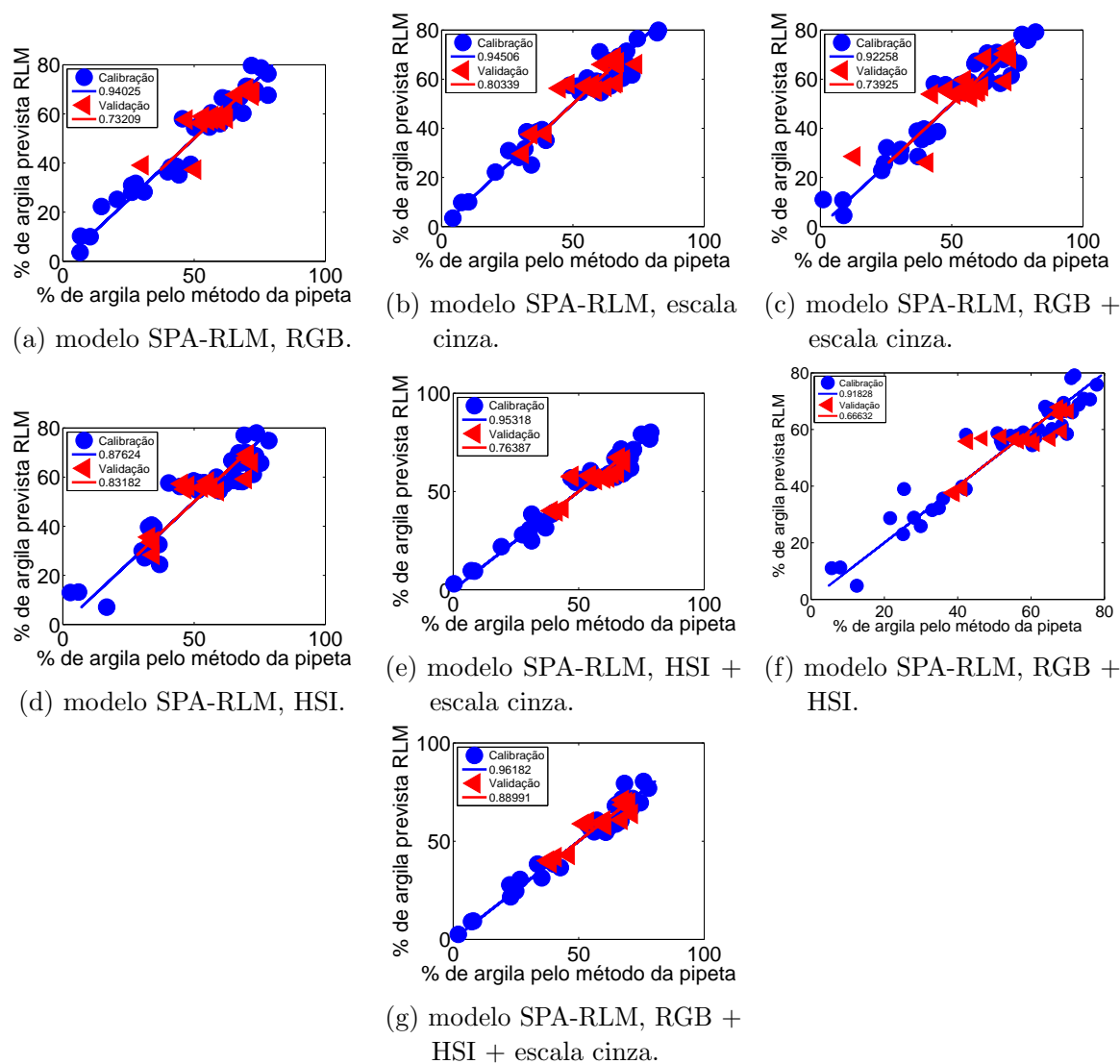


Figura 30 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

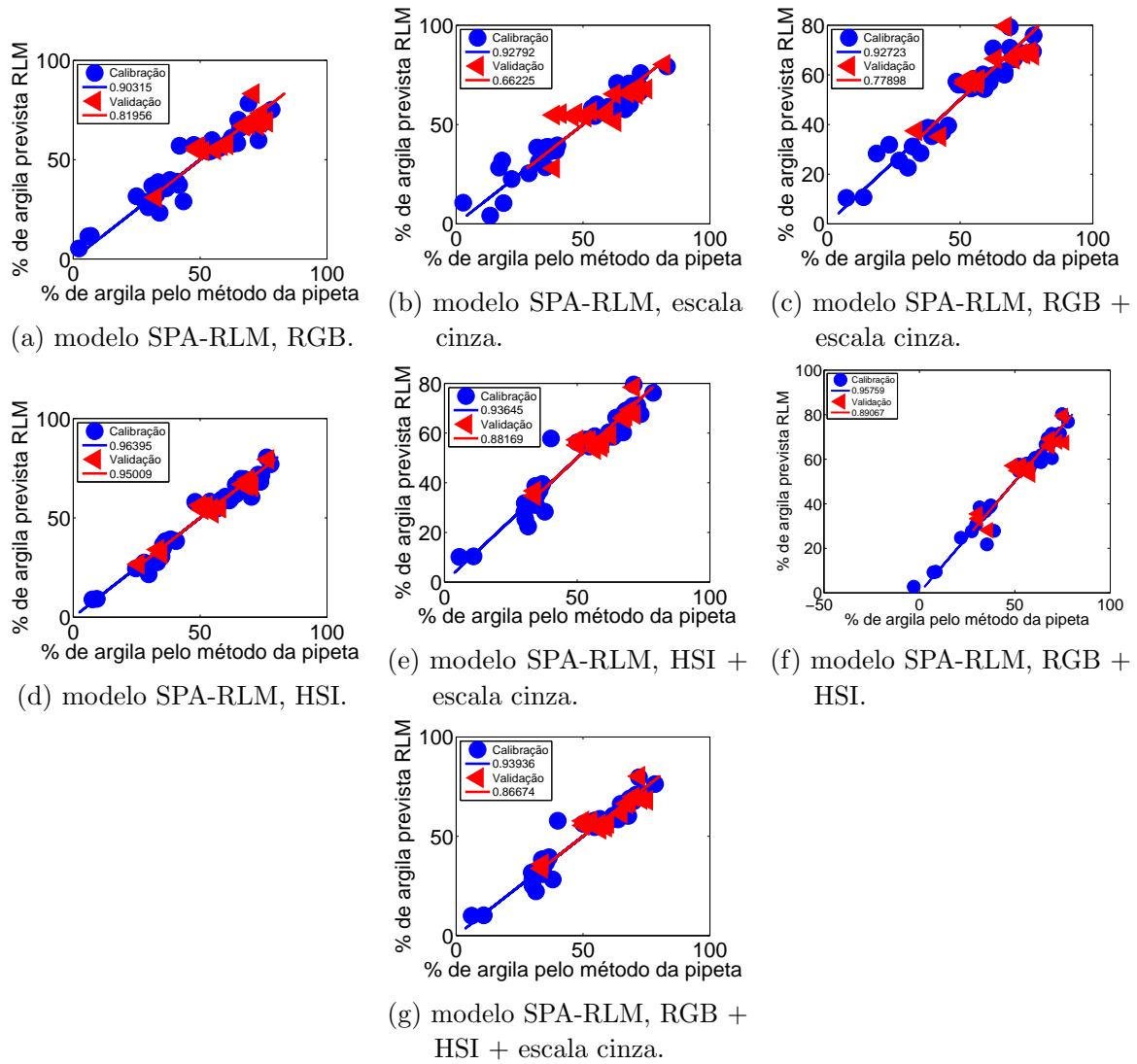


Figura 31 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

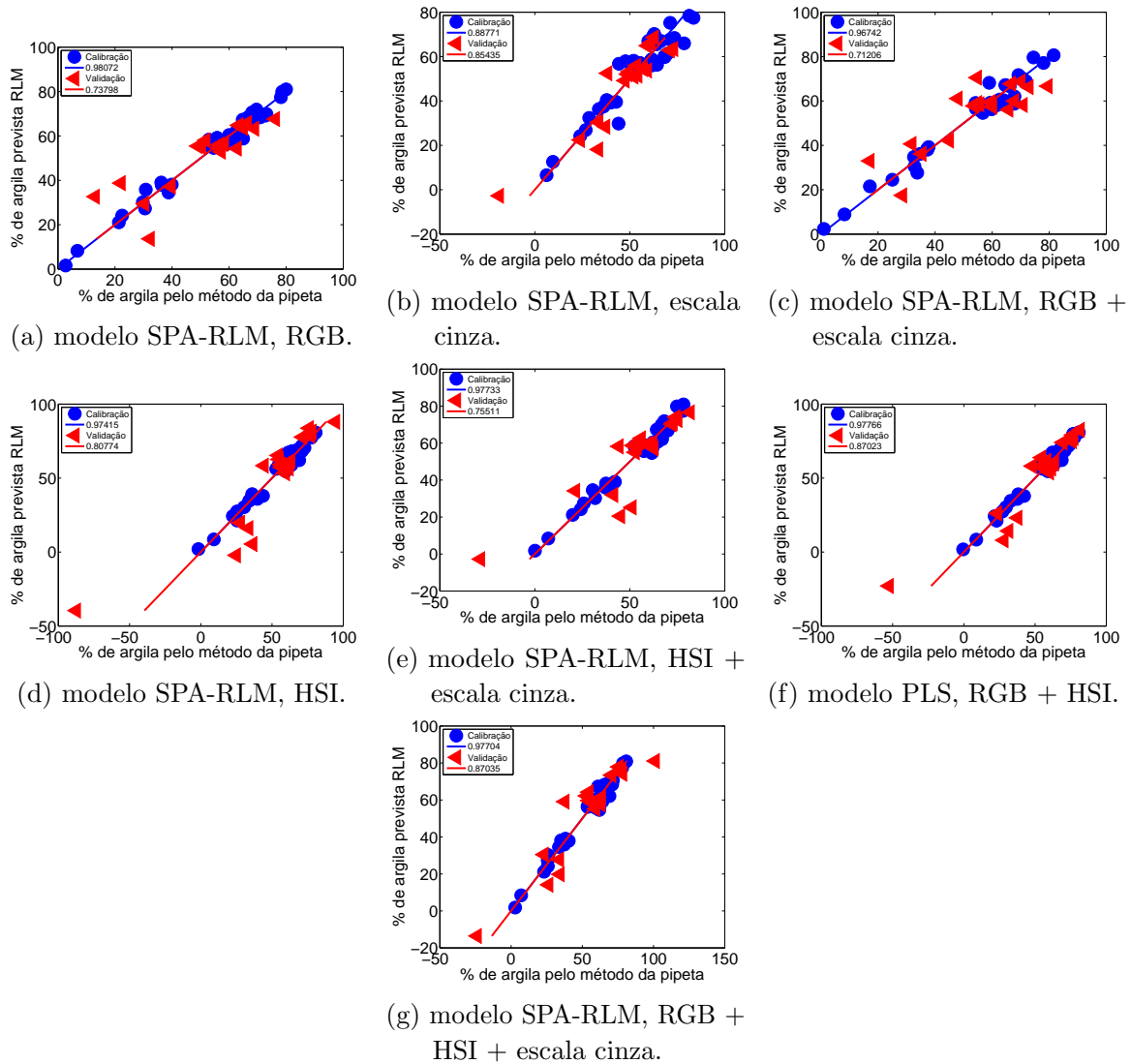


Figura 32 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.



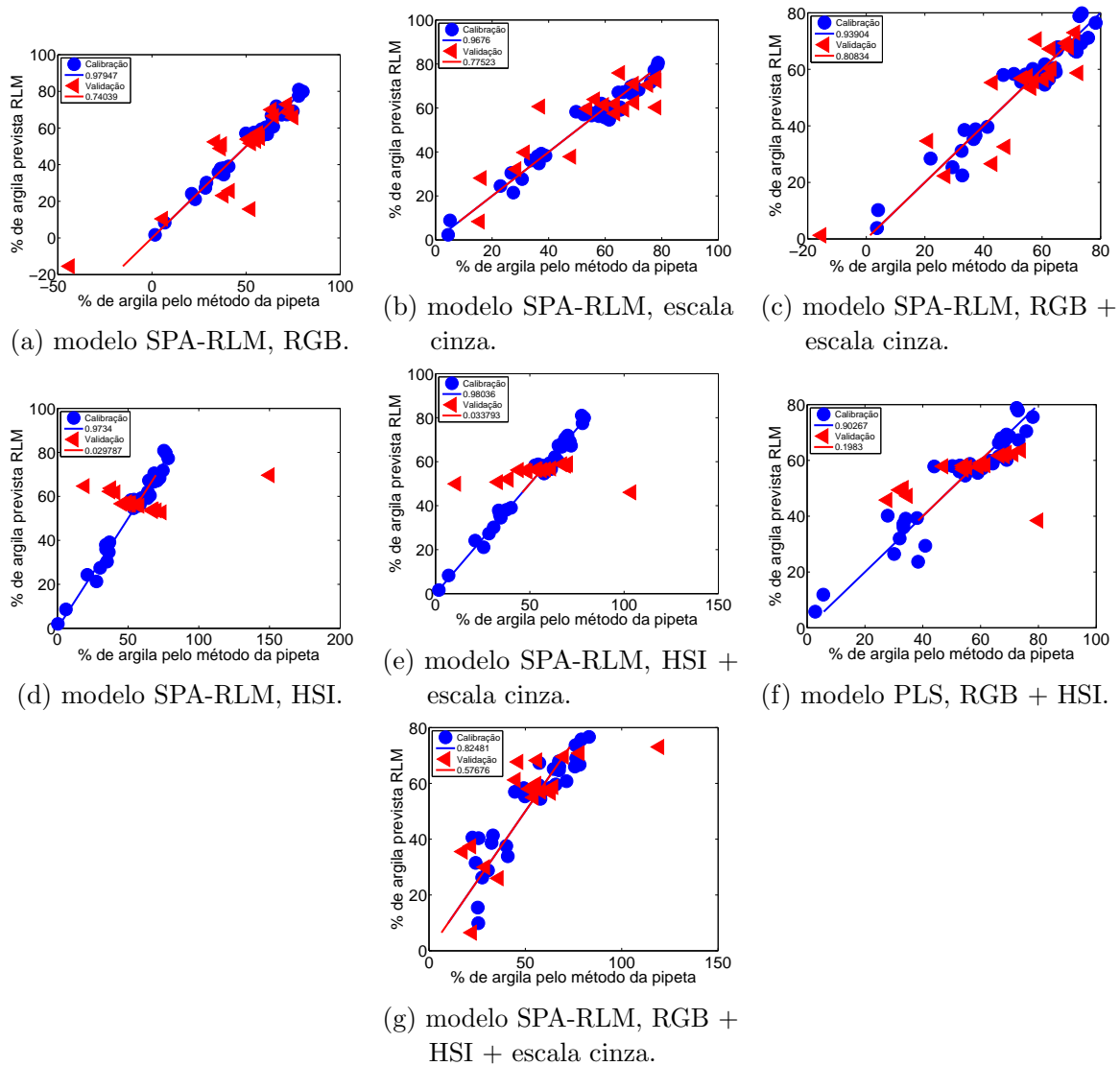


Figura 33 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice C – Relatório de desempenho do algoritmo genético aplicado a modelos PLS referentes as imagens com 6,3x de aumento óptico

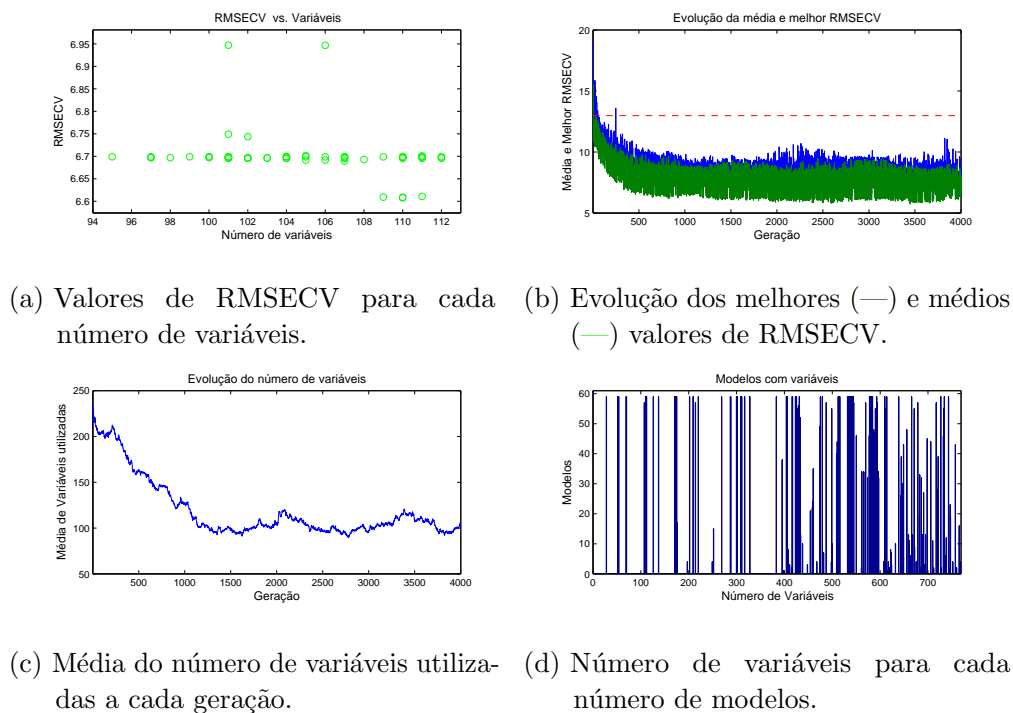


Figura 34 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB.

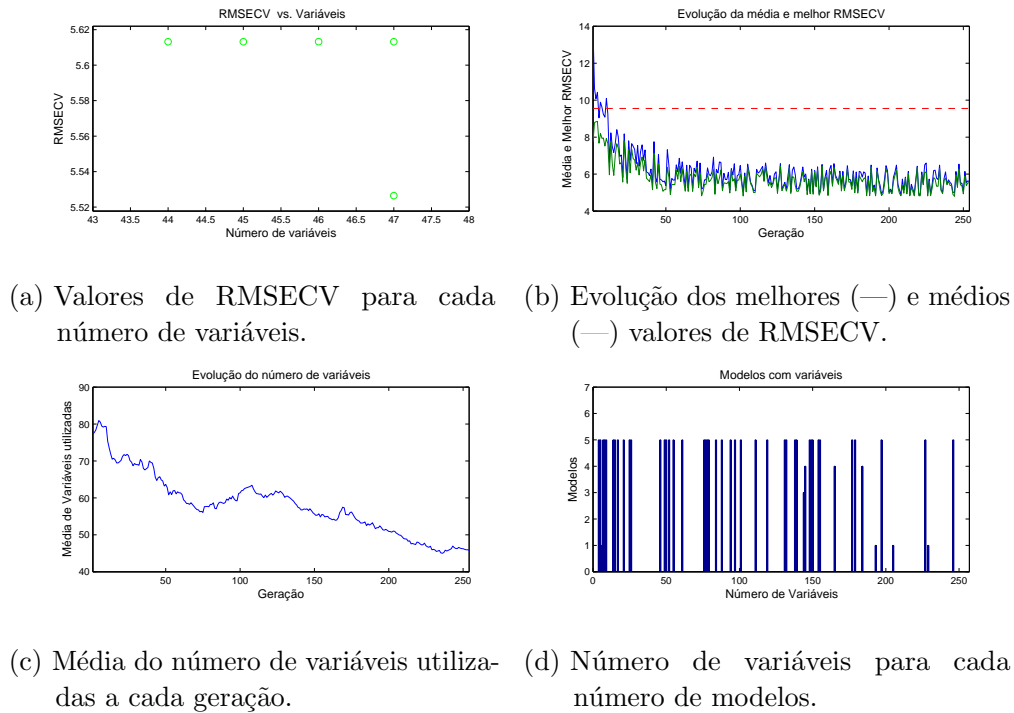


Figura 35 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza.

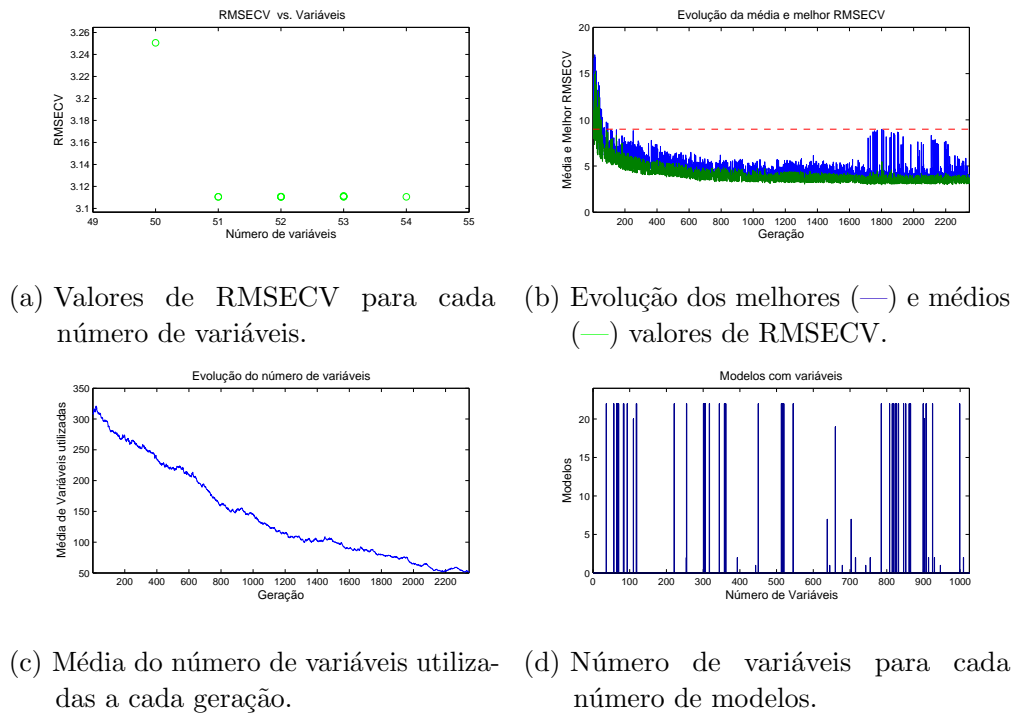


Figura 36 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza.

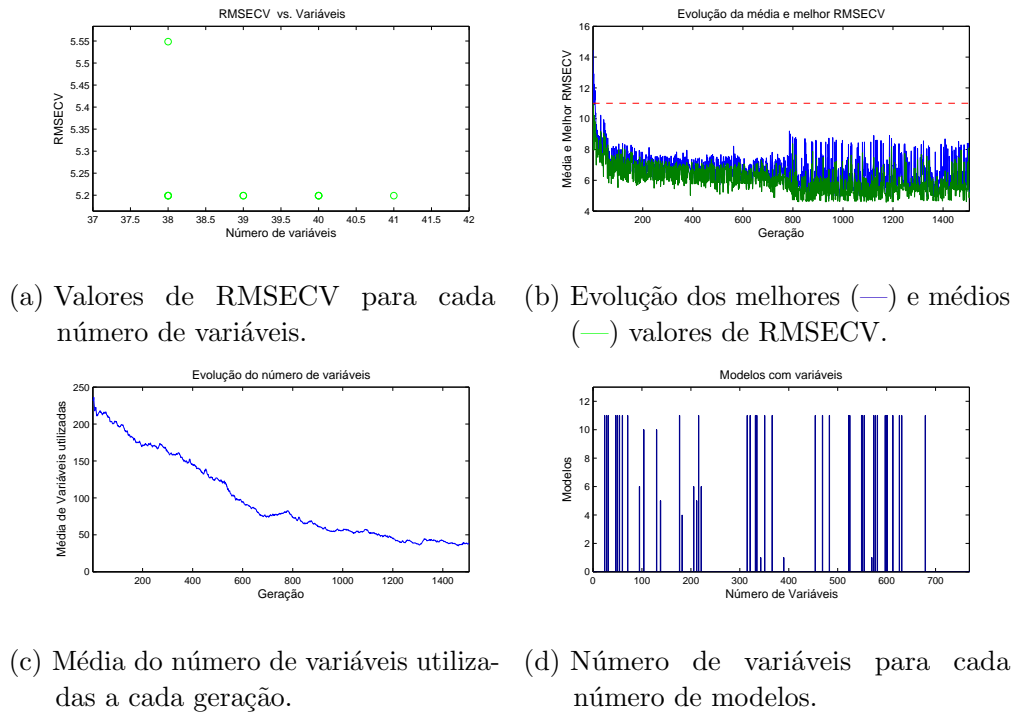


Figura 37 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI.

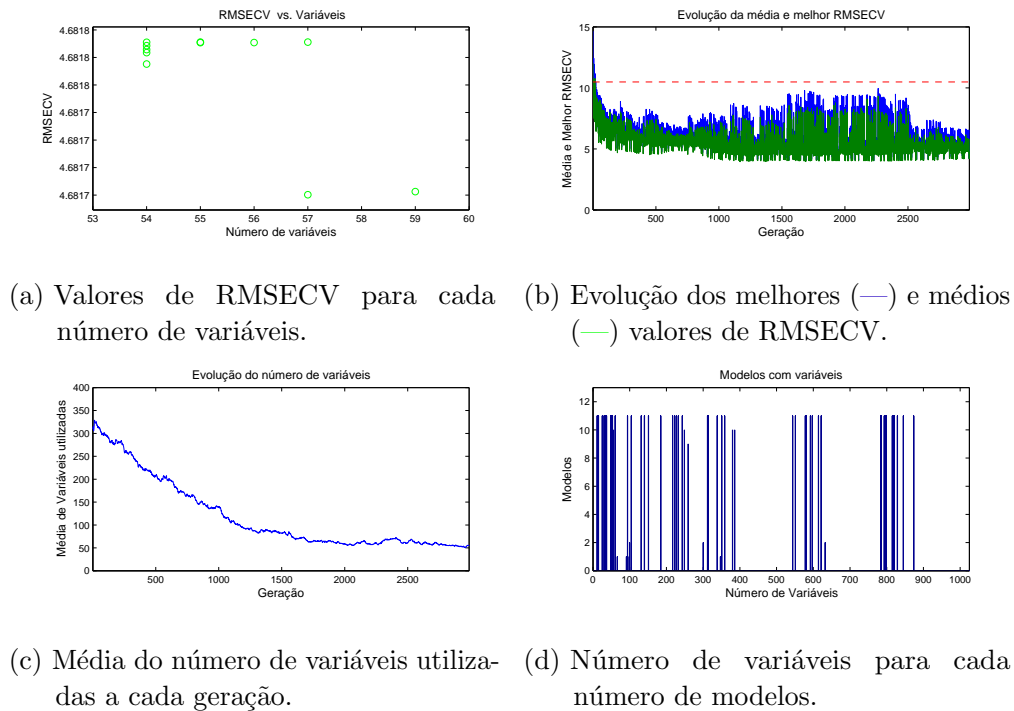


Figura 38 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza.

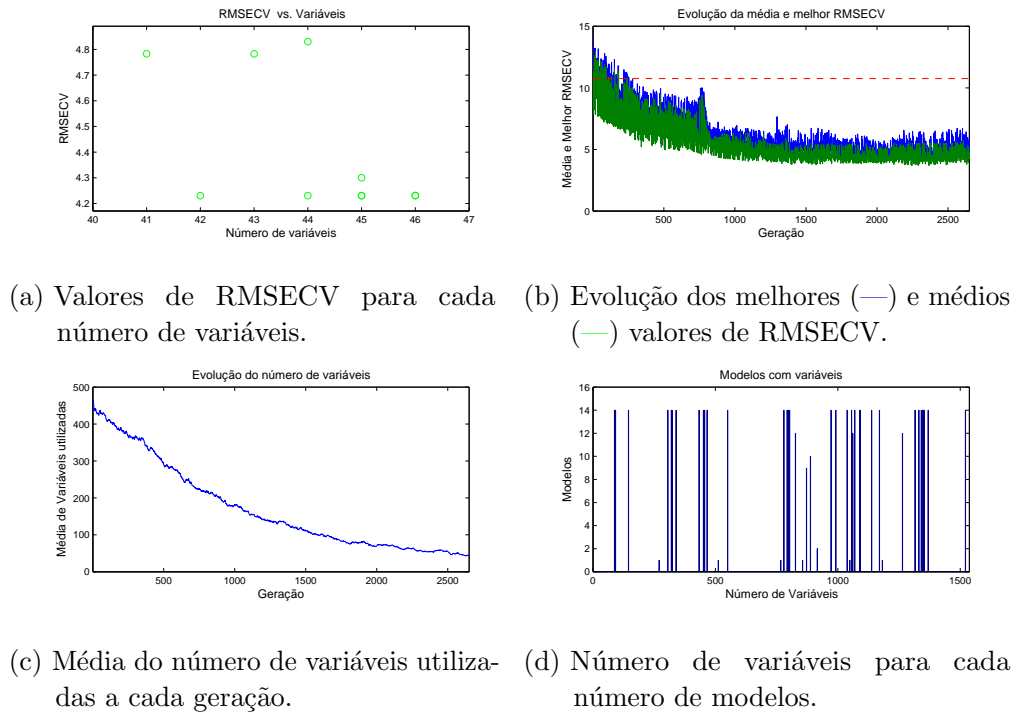


Figura 39 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI.

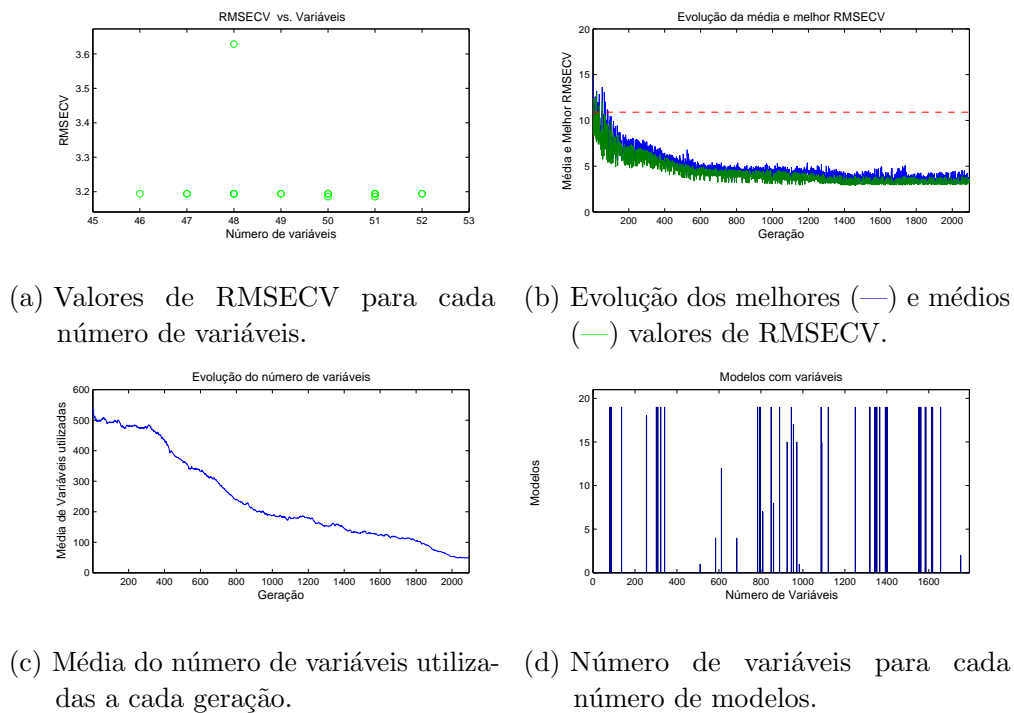


Figura 40 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza.

## Apêndice D – Relatório de desempenho do algoritmo genético aplicado a modelos PLS referentes as imagens com 10x de aumento óptico

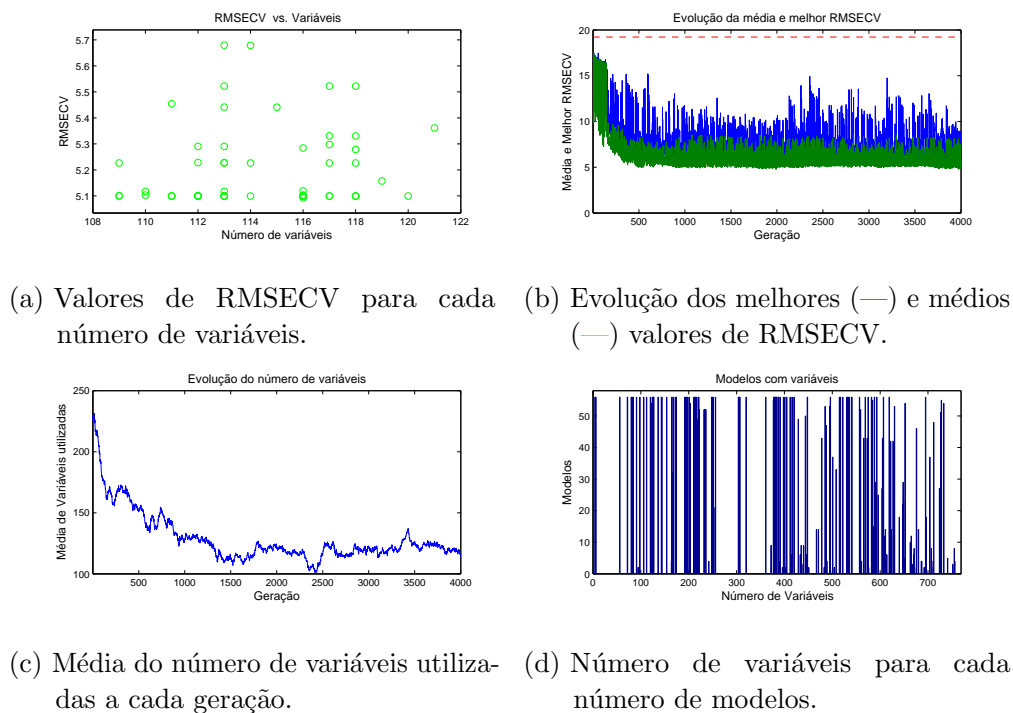


Figura 41 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB.

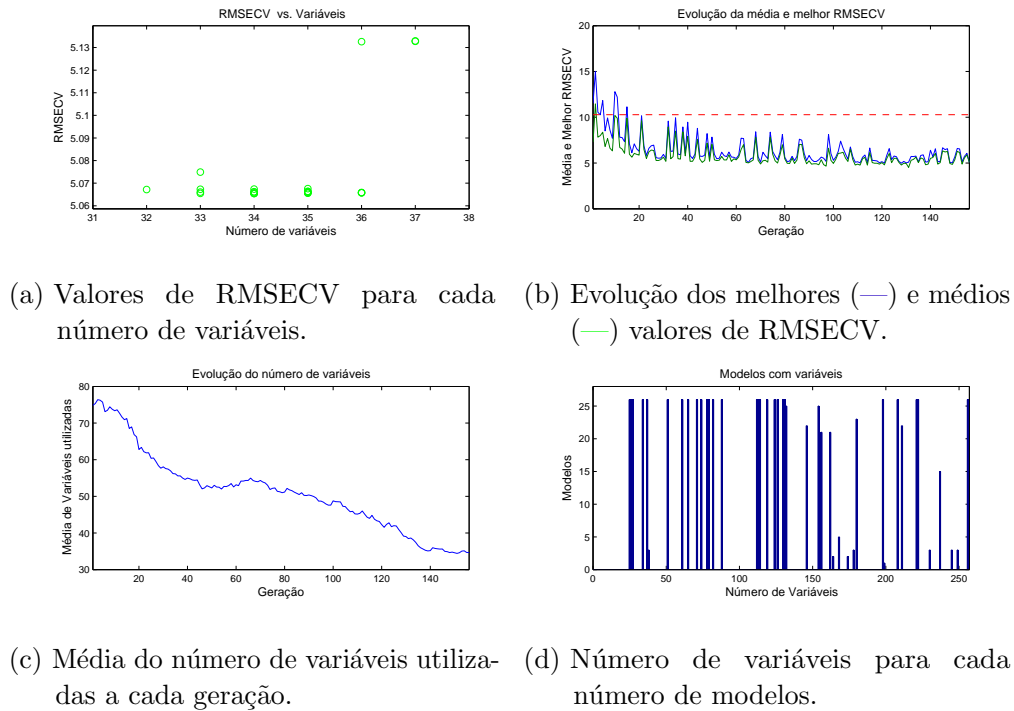


Figura 42 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza.

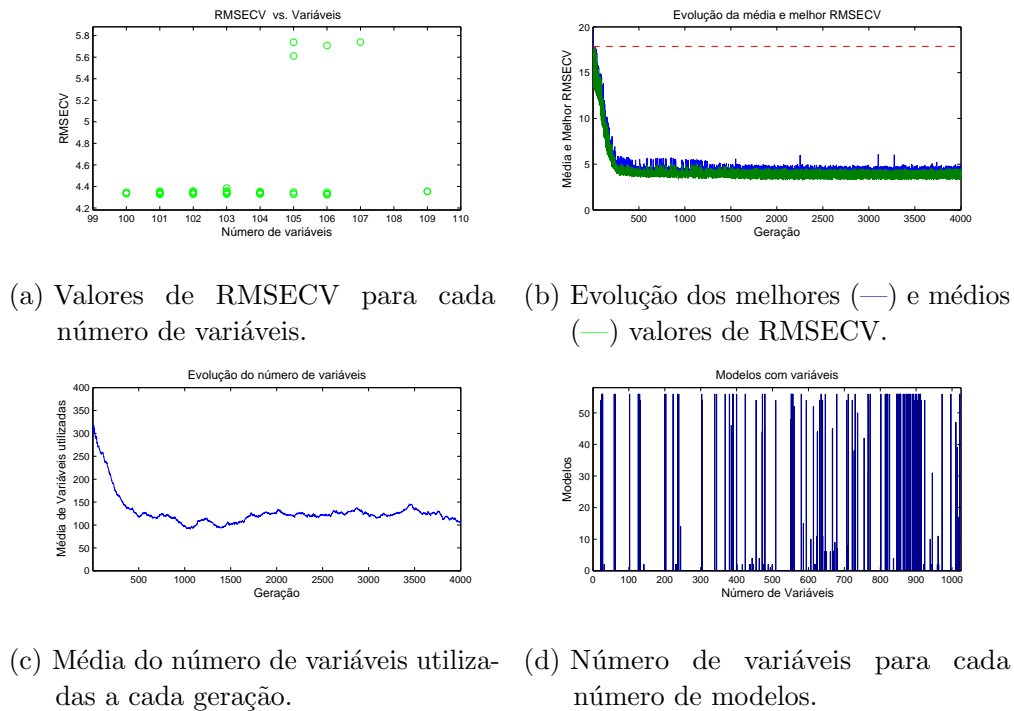
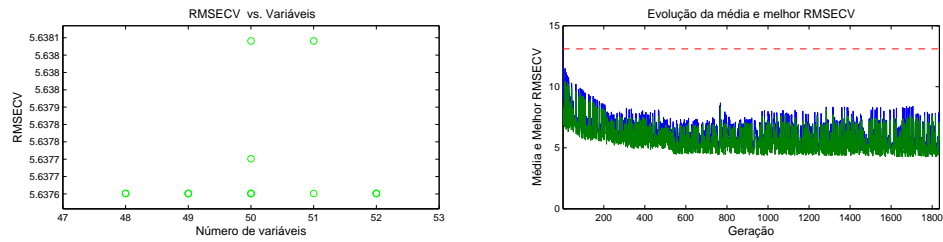
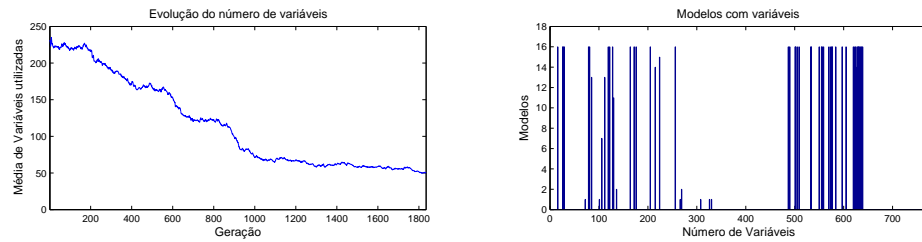


Figura 43 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza.

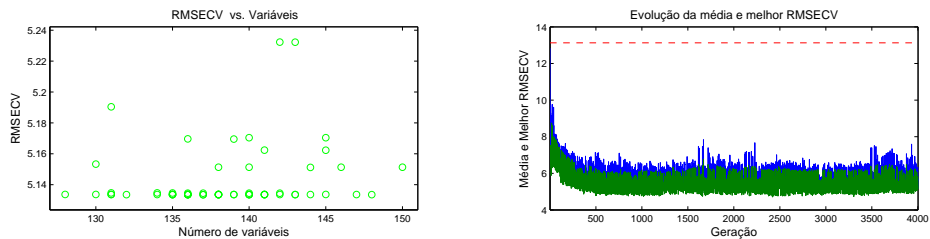


- (a) Valores de RMSECV para cada número de variáveis. (b) Evolução dos melhores (—) e médios (—) valores de RMSECV.

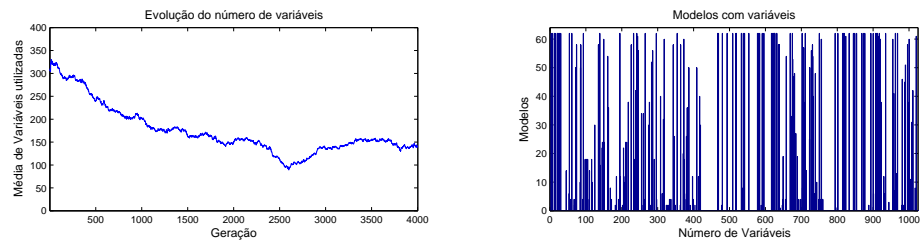


- (c) Média do número de variáveis utilizadas a cada geração. (d) Número de variáveis para cada número de modelos.

Figura 44 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI.



- (a) Valores de RMSECV para cada número de variáveis. (b) Evolução dos melhores (—) e médios (—) valores de RMSECV.



- (c) Média do número de variáveis utilizadas a cada geração. (d) Número de variáveis para cada número de modelos.

Figura 45 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza.



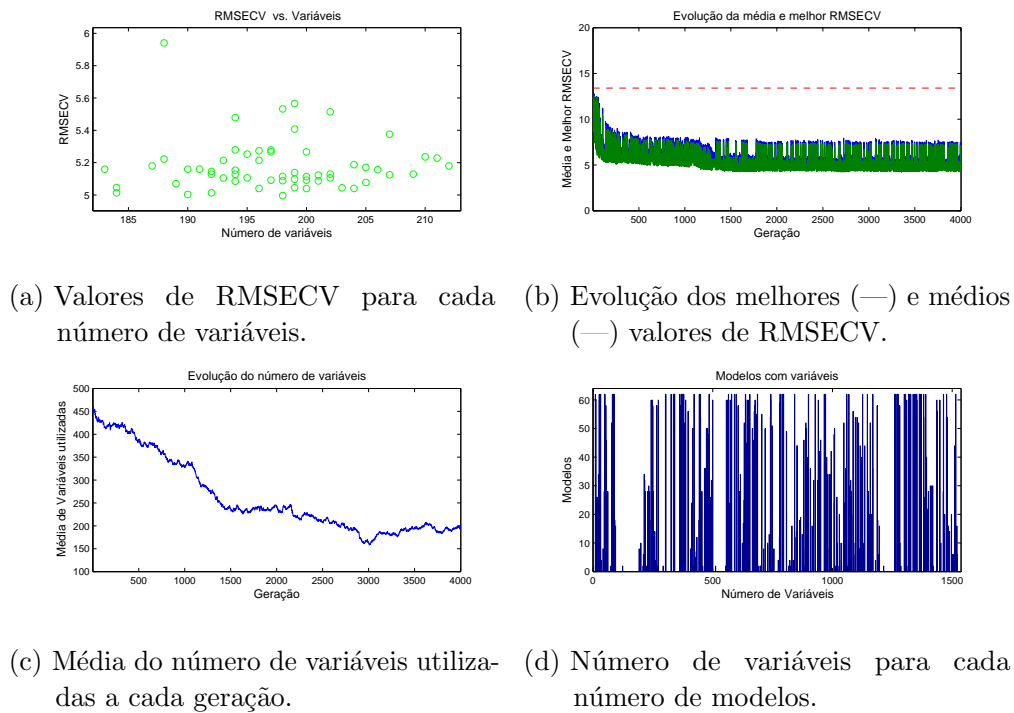


Figura 46 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI.

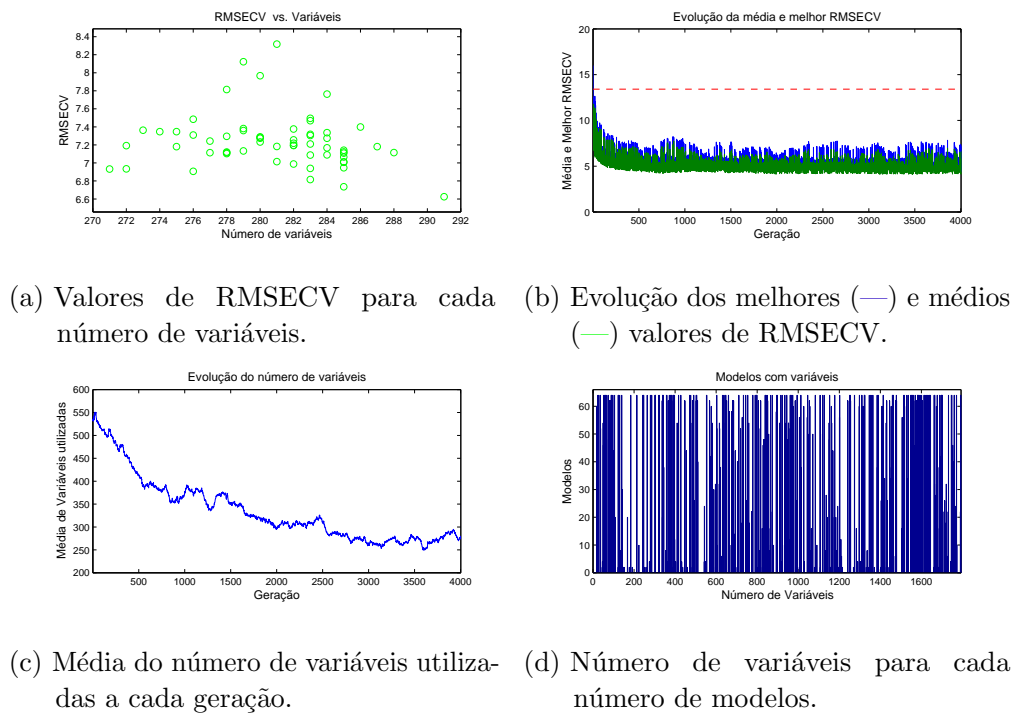


Figura 47 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza.

## Apêndice E – Modelos de regressão GA-PLS para predição dos teores de argila em amostras de solo

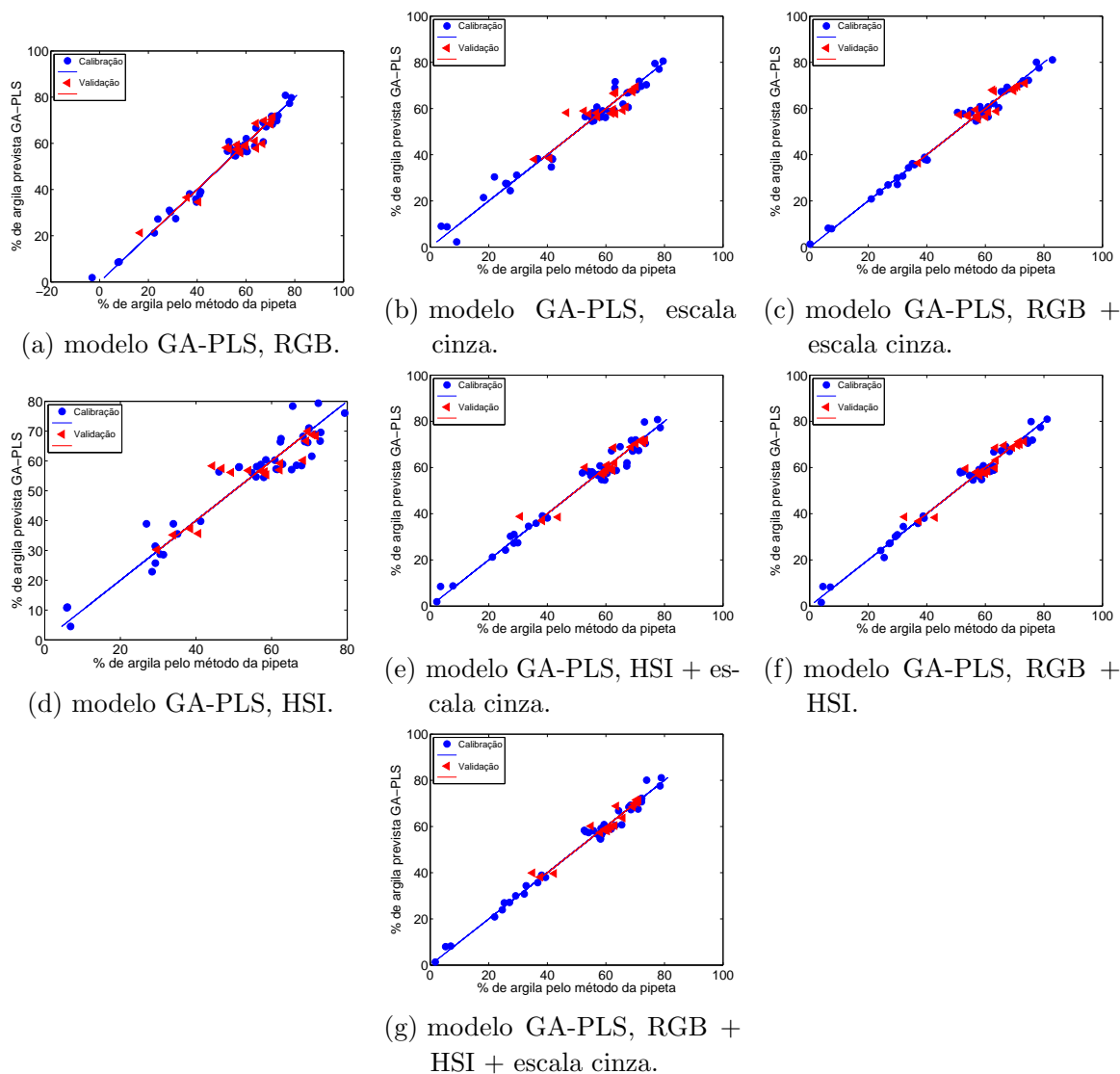


Figura 48 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

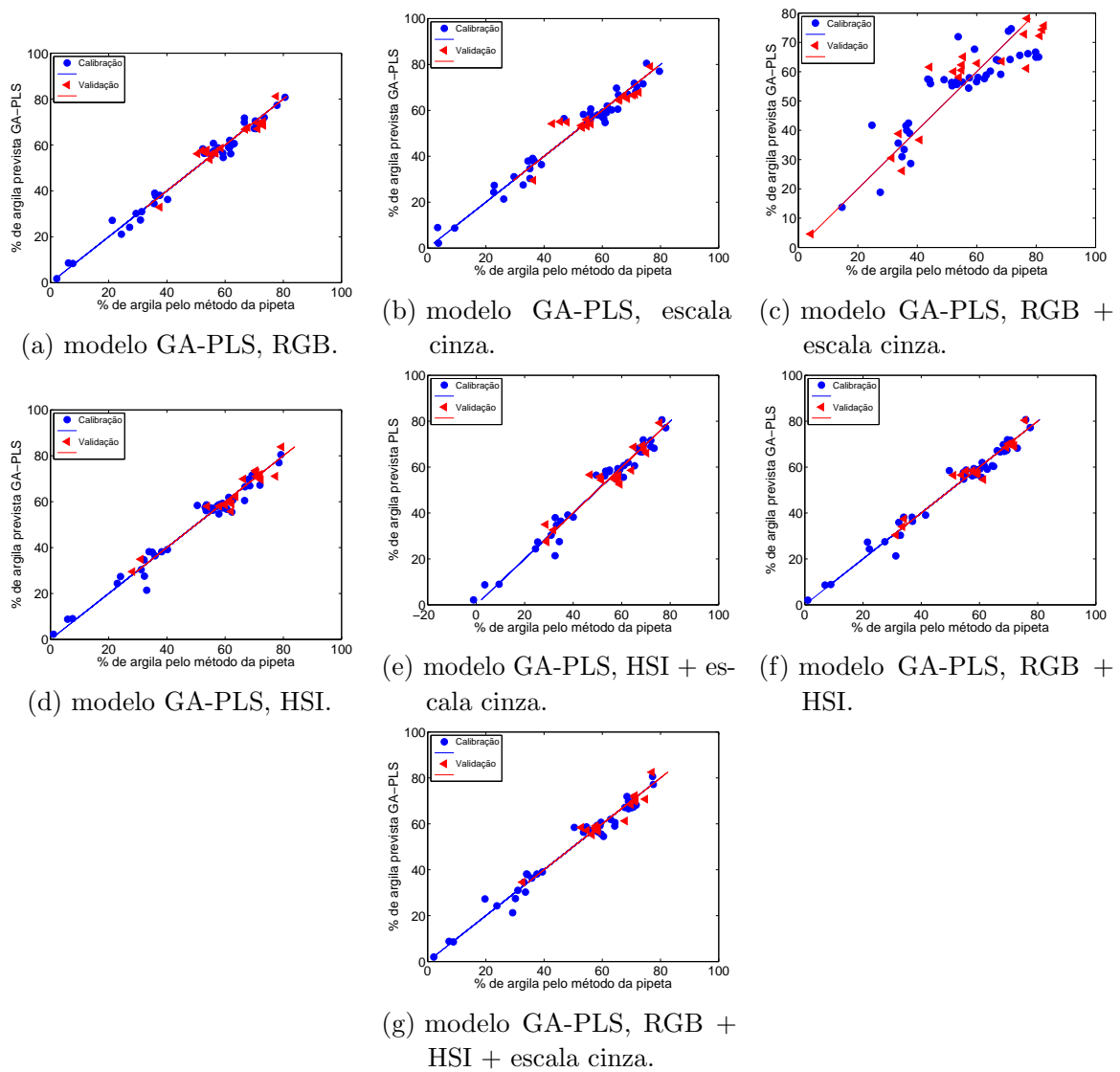


Figura 49 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

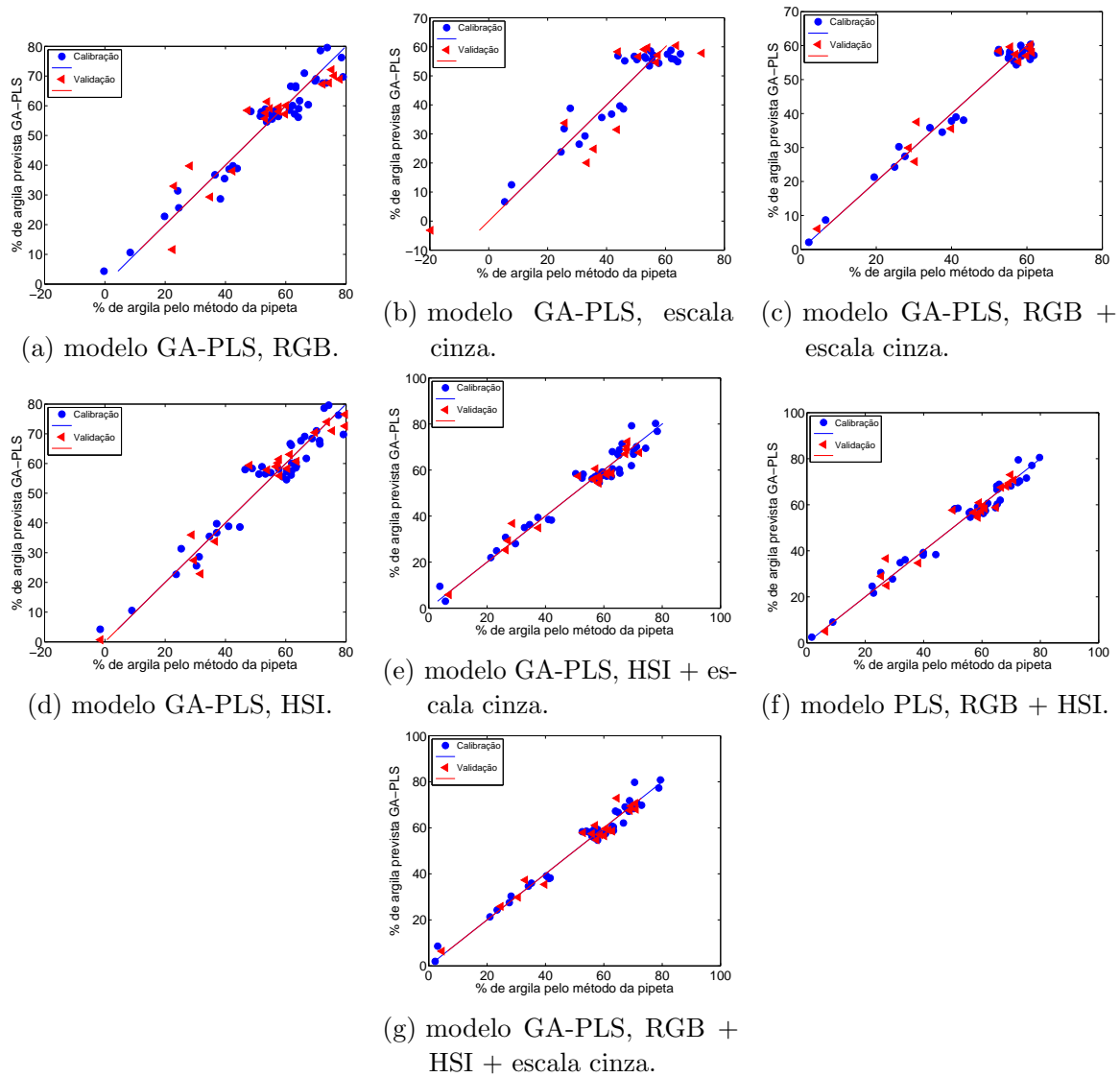


Figura 50 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.

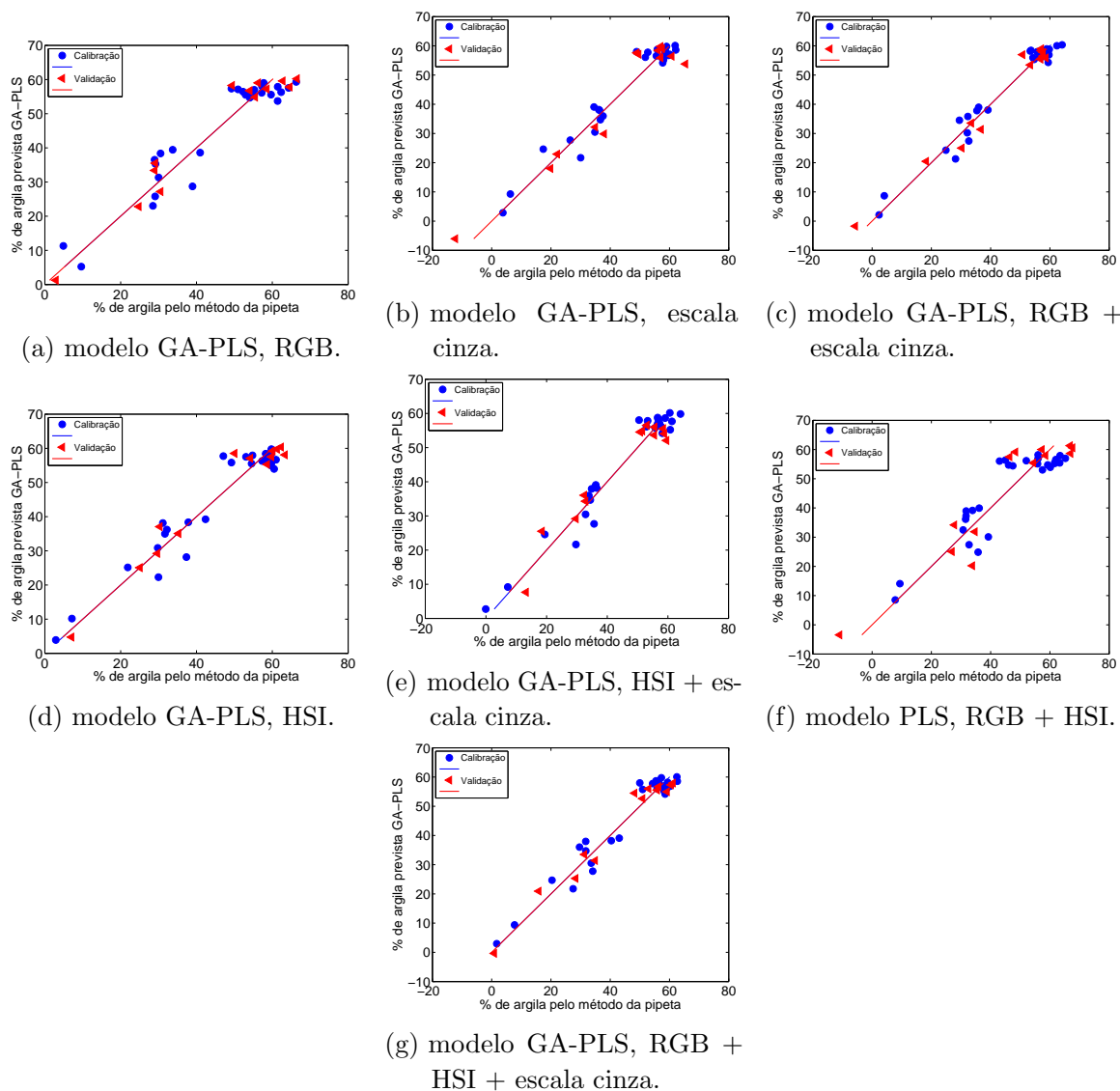


Figura 51 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice F – Modelos de regressão por mínimos quadrados PLS para predição dos teores de areia em amostras de solo

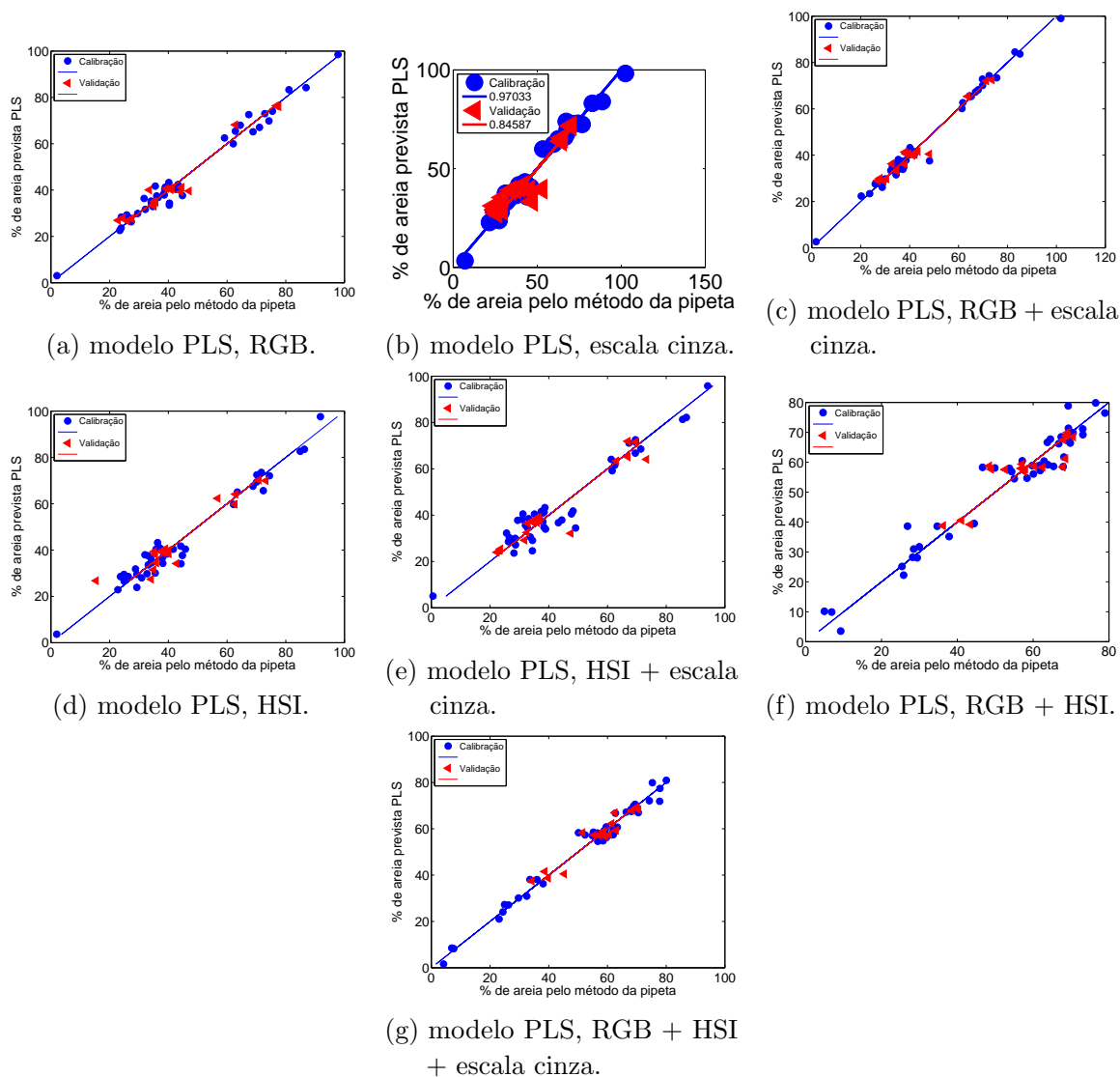


Figura 52 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

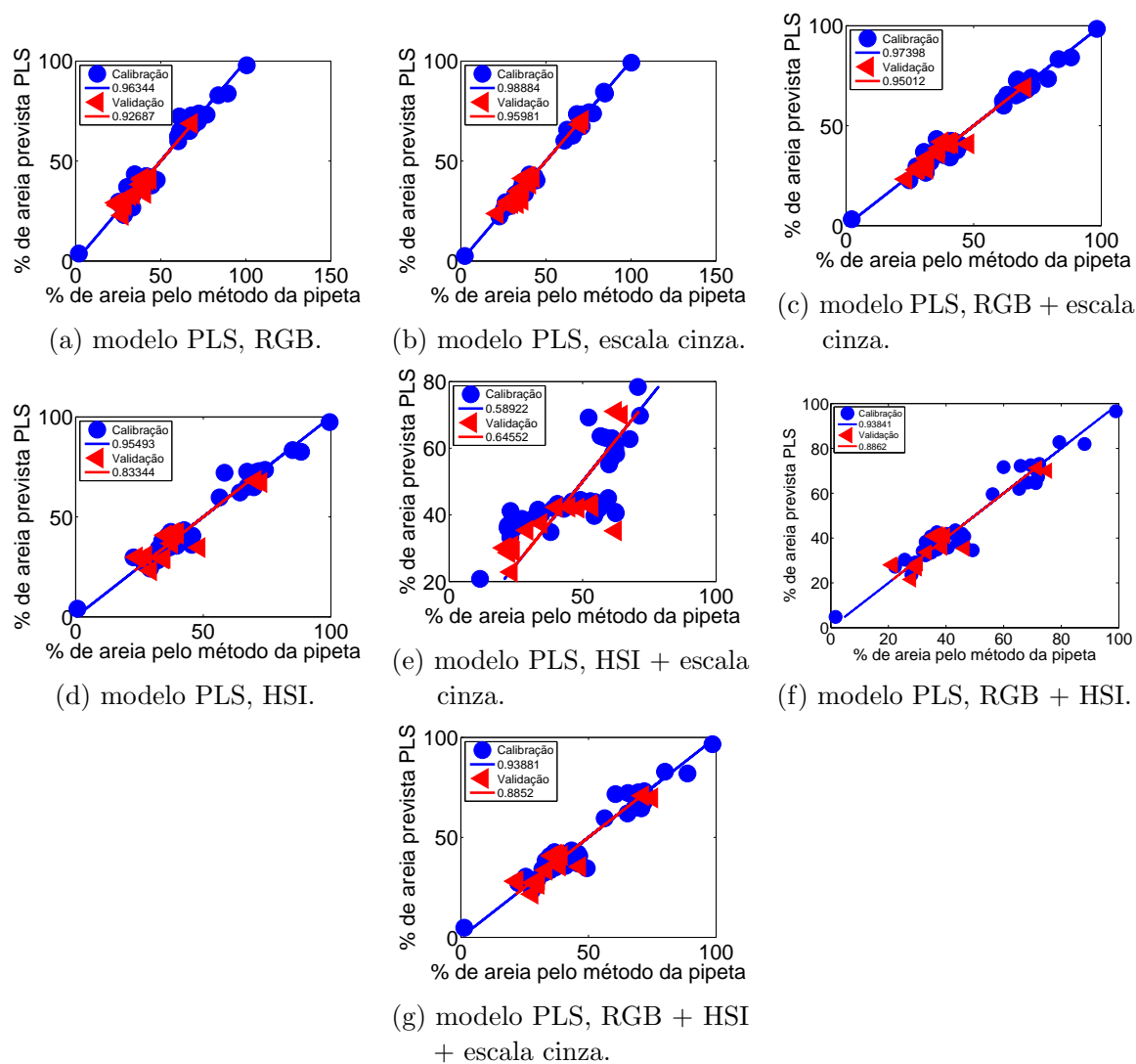


Figura 53 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

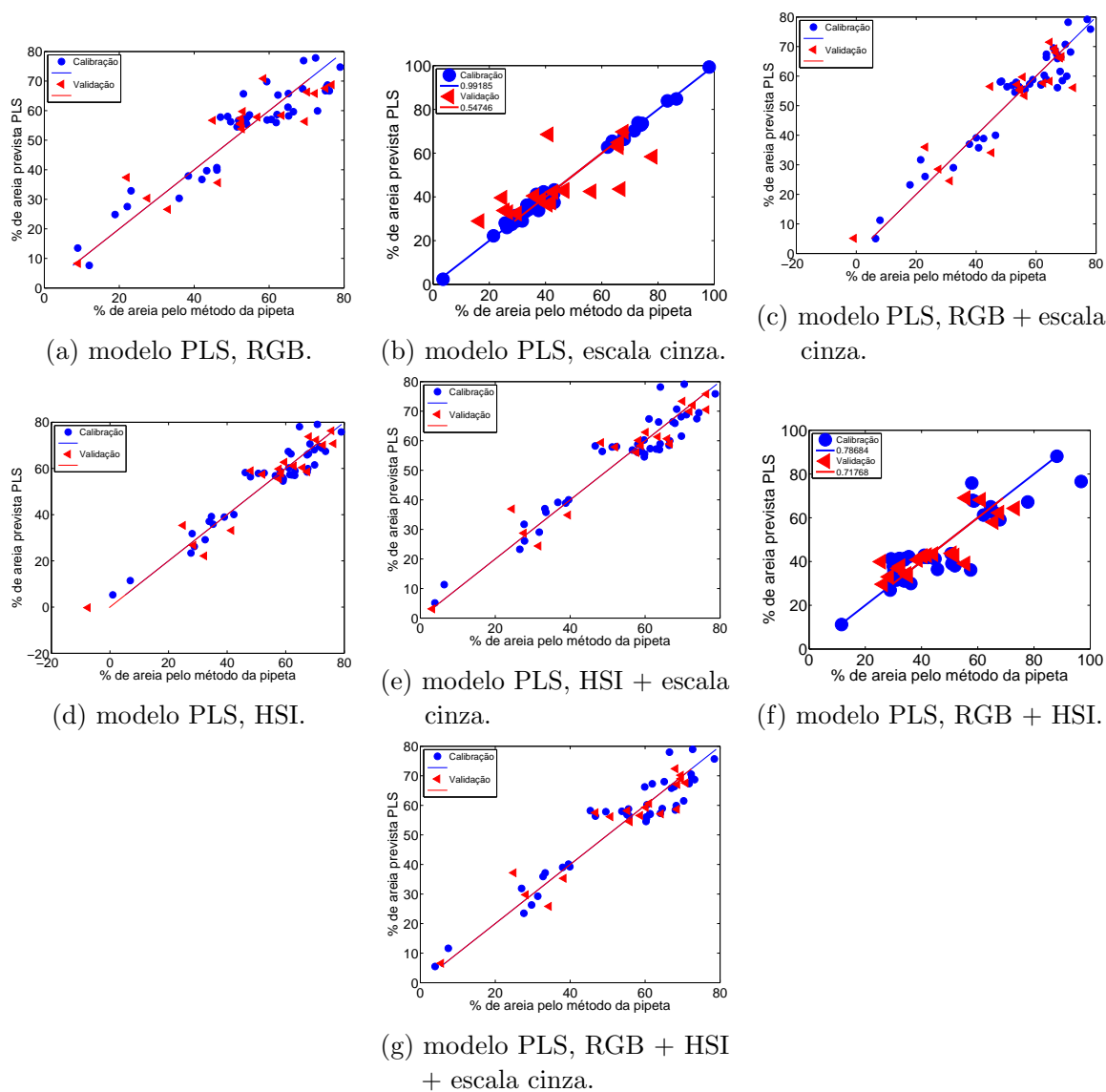


Figura 54 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.



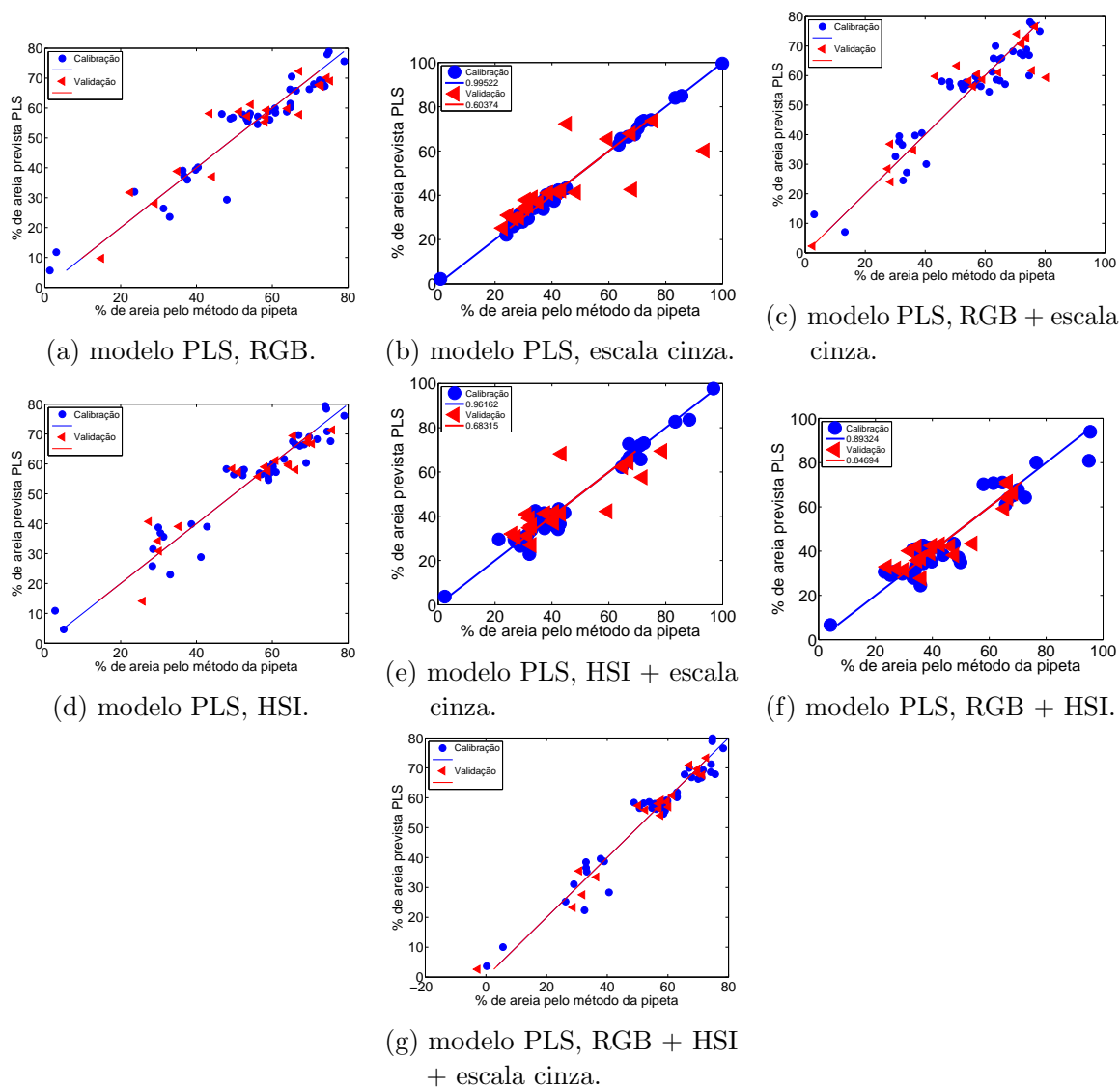


Figura 55 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice G – Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de areia em amostras de solo

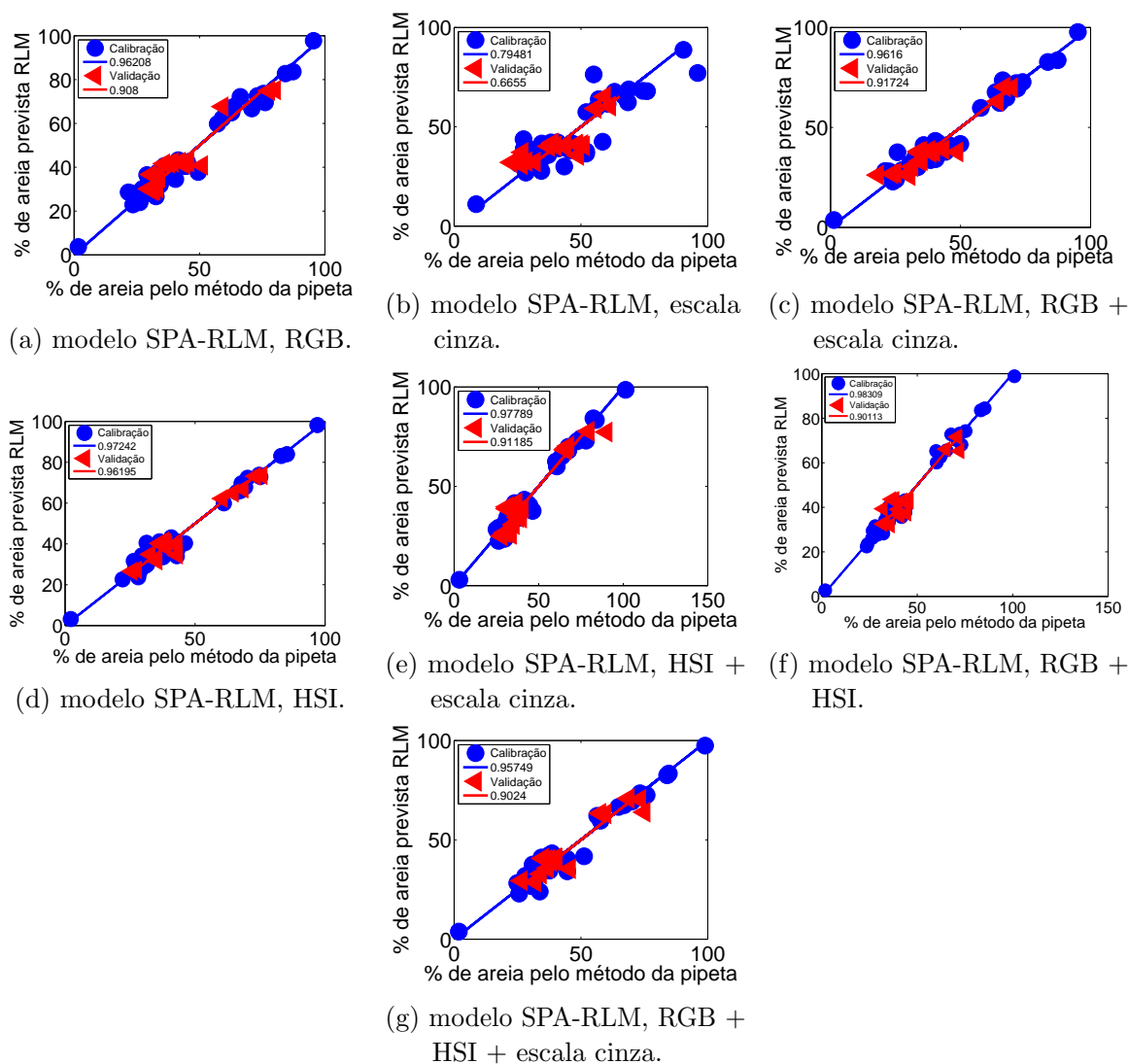


Figura 56 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

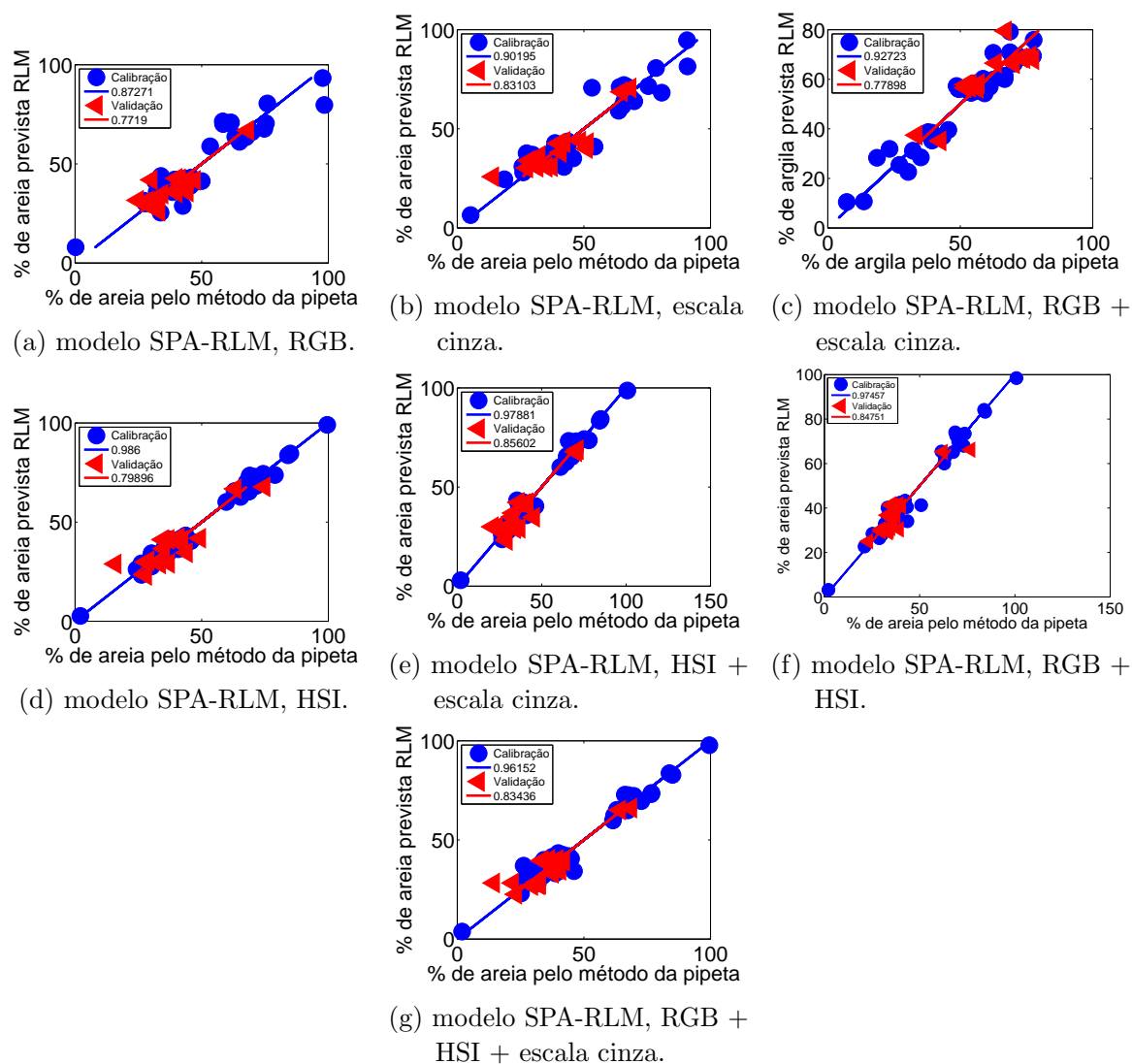


Figura 57 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

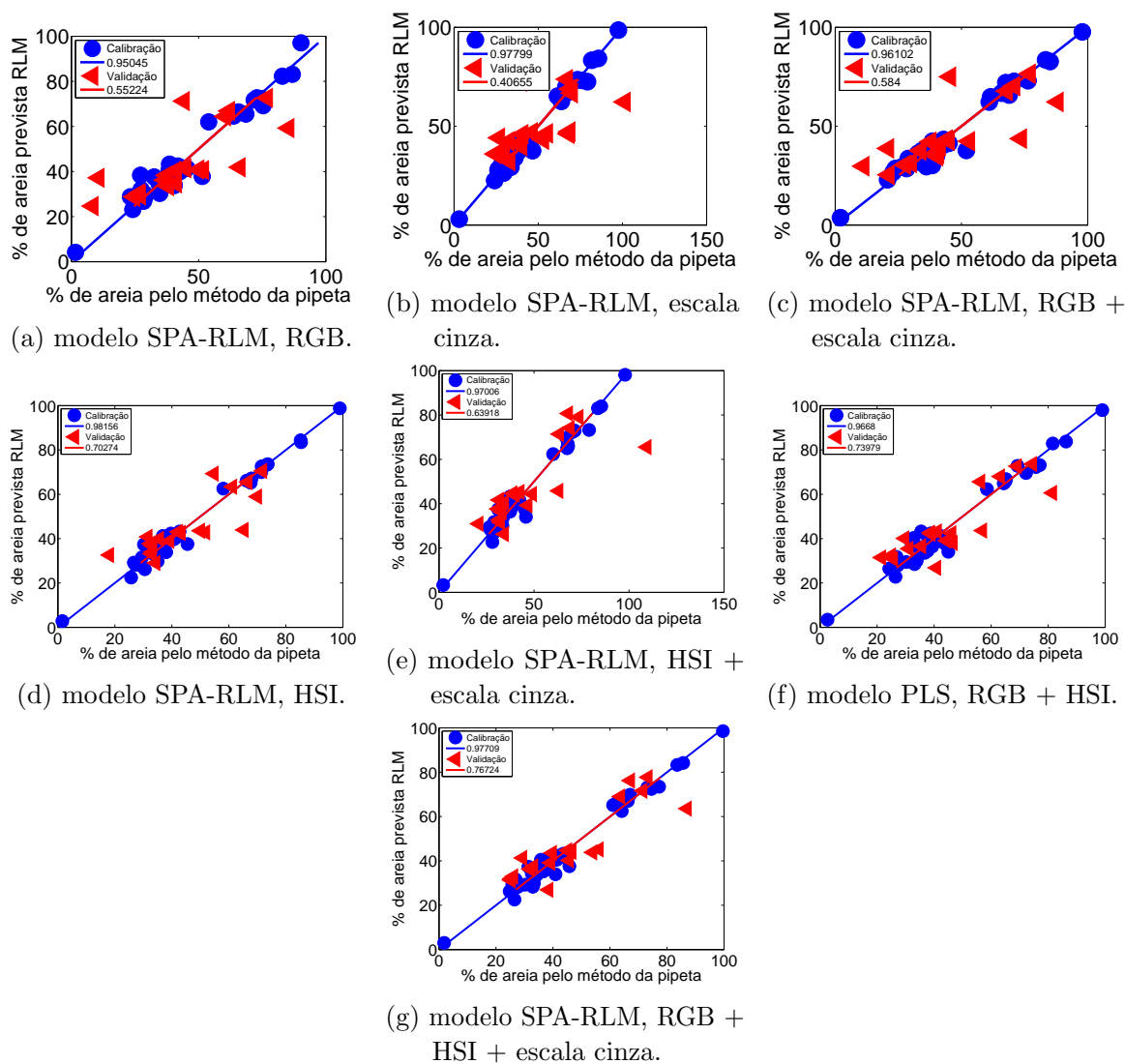


Figura 58 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.

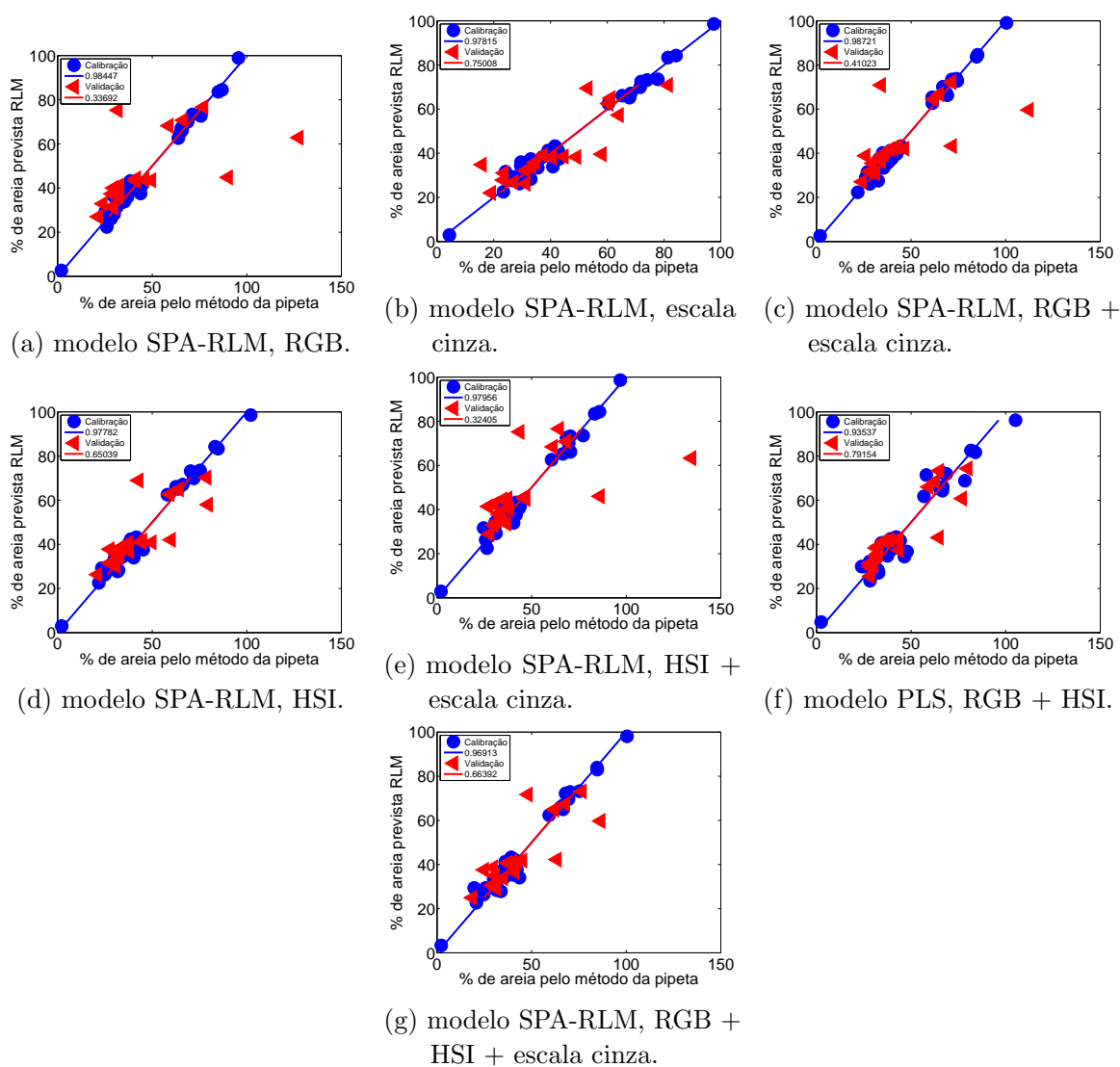


Figura 59 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice H – Relatório de desempenho do algoritmo genético aplicado a modelos PLS referentes as imagens com 6,3x de aumento óptico

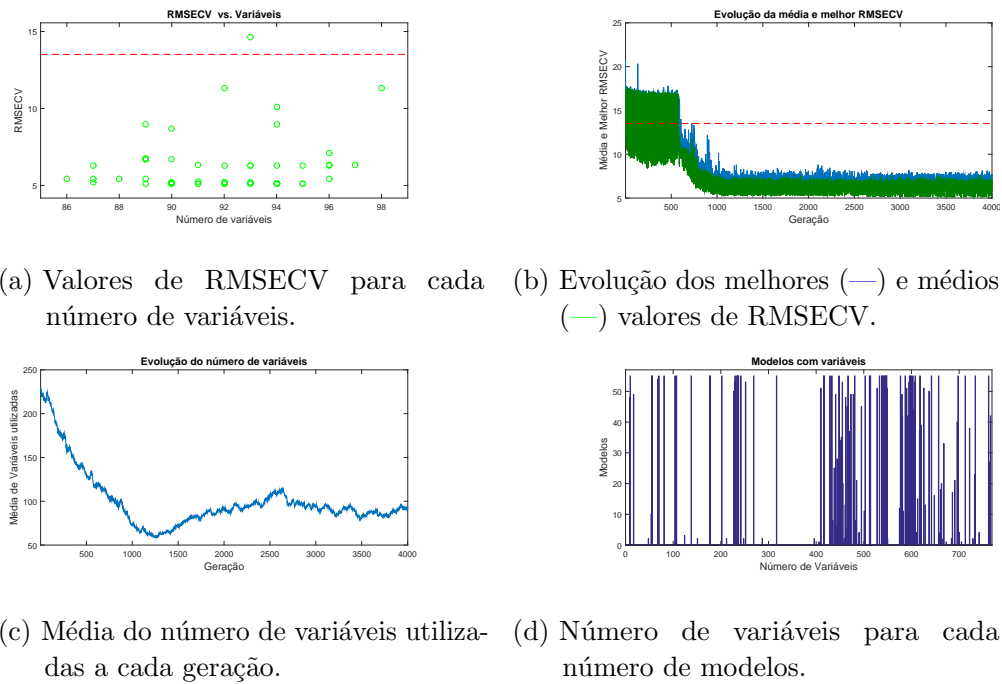


Figura 60 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB.

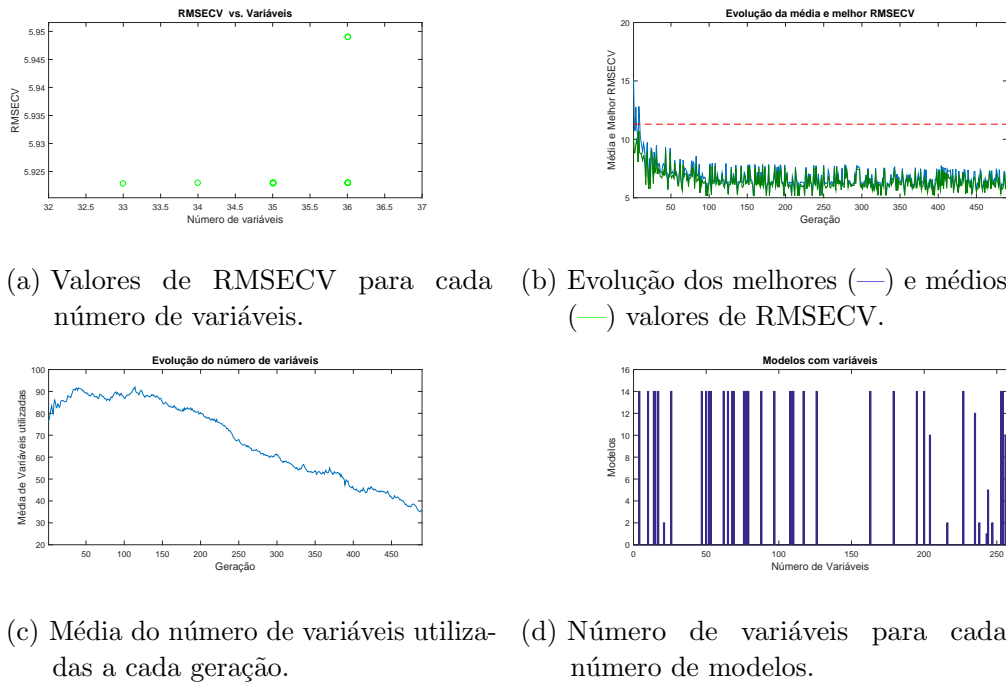


Figura 61 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza.

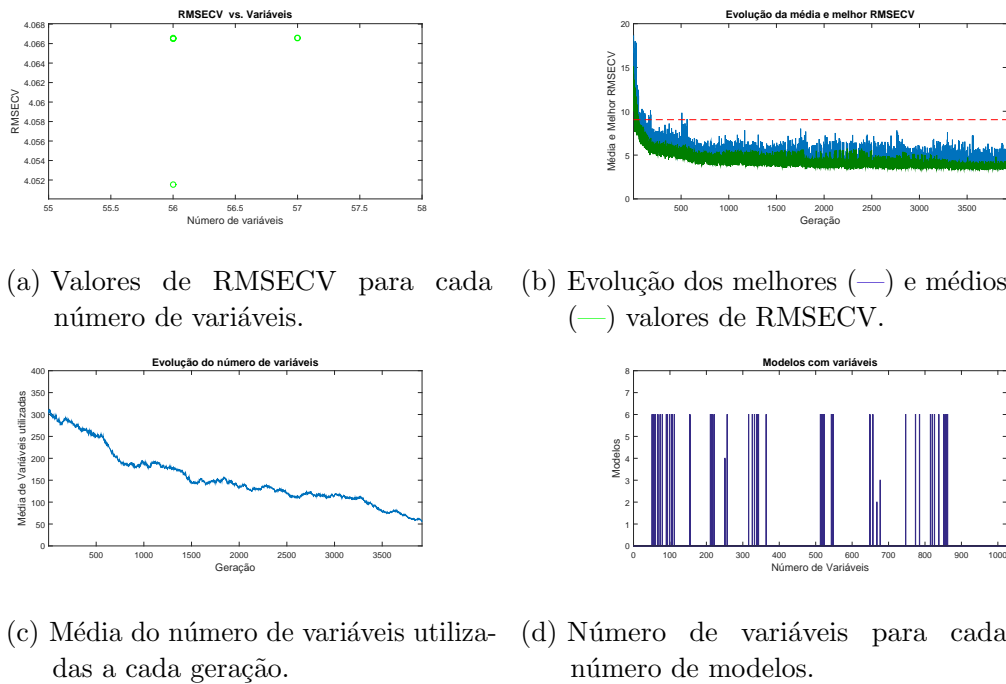


Figura 62 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza.

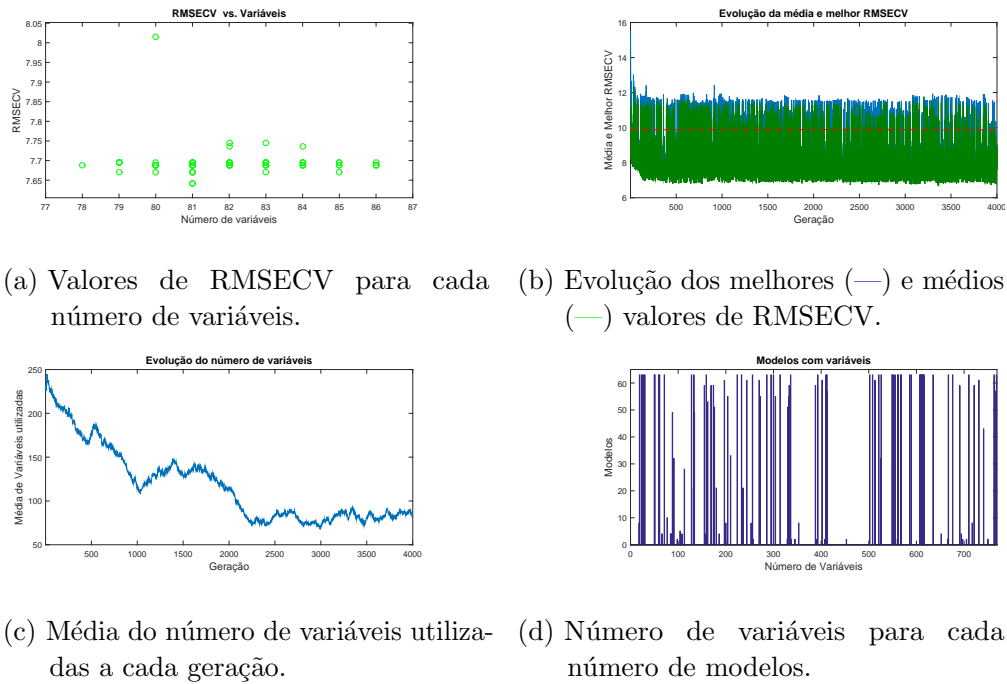


Figura 63 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI.

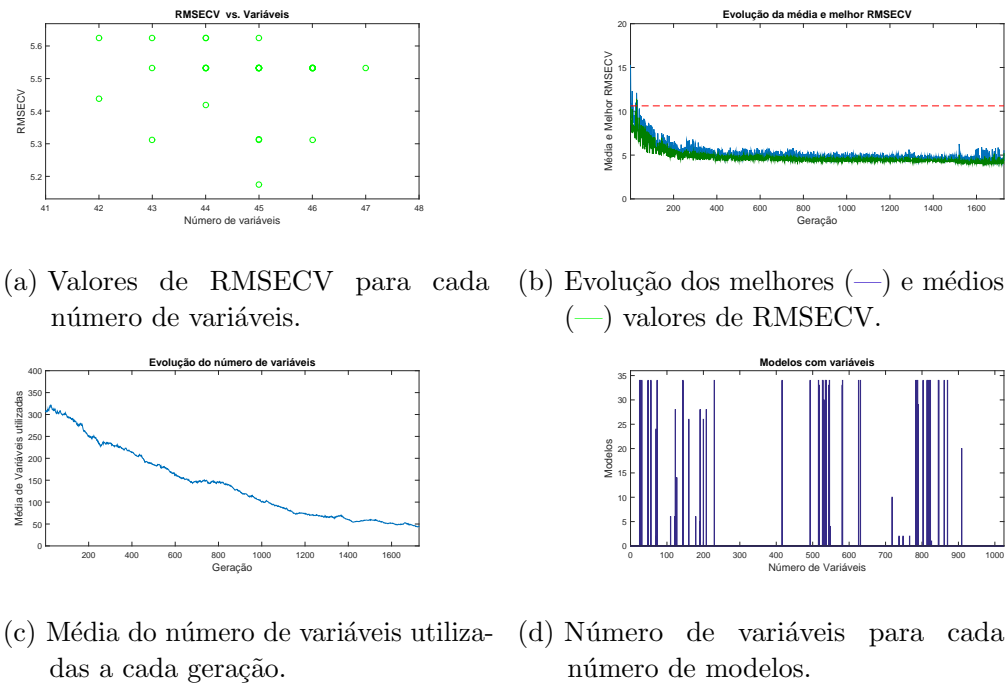
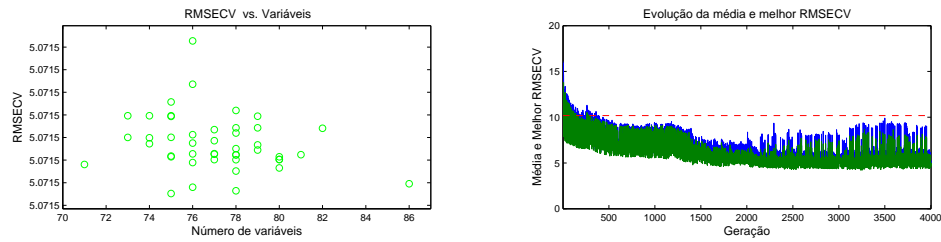
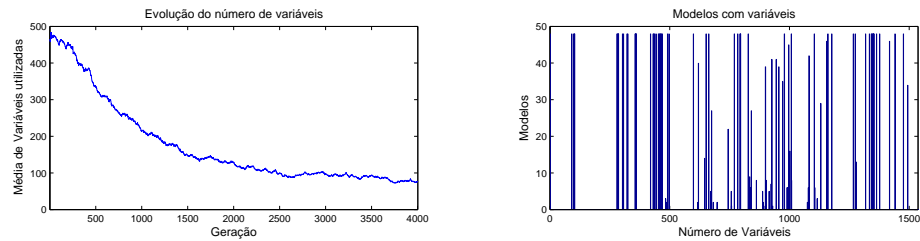


Figura 64 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza.



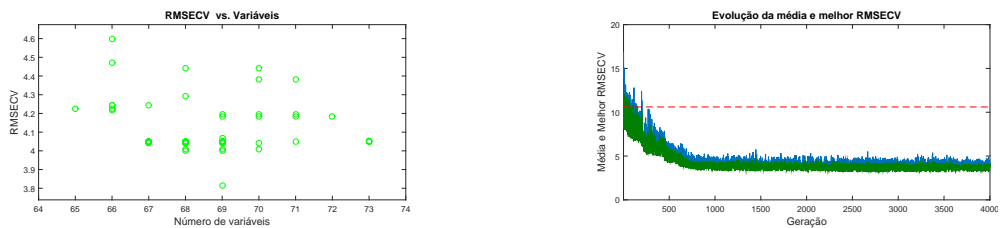


(a) Valores de RMSECV para cada número de variáveis. (b) Evolução dos melhores (—) e médios (—) valores de RMSECV.

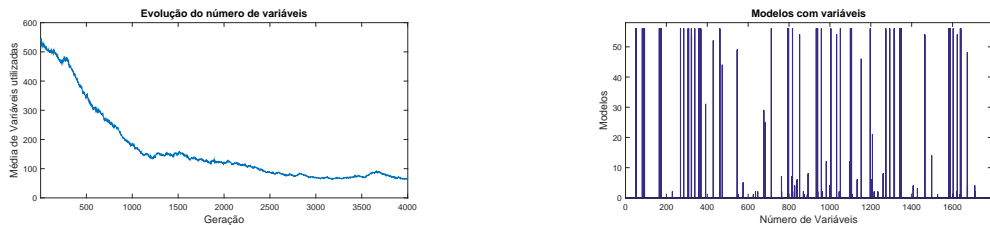


(c) Média do número de variáveis utilizadas a cada geração. (d) Número de variáveis para cada número de modelos.

Figura 65 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI.



(a) Valores de RMSECV para cada número de variáveis. (b) Evolução dos melhores (—) e médios (—) valores de RMSECV.



(c) Média do número de variáveis utilizadas a cada geração. (d) Número de variáveis para cada número de modelos.

Figura 66 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza.

## Apêndice I – Relatório de desempenho do algoritmo genético aplicado a modelos PLS referentes as imagens com 10x de aumento óptico

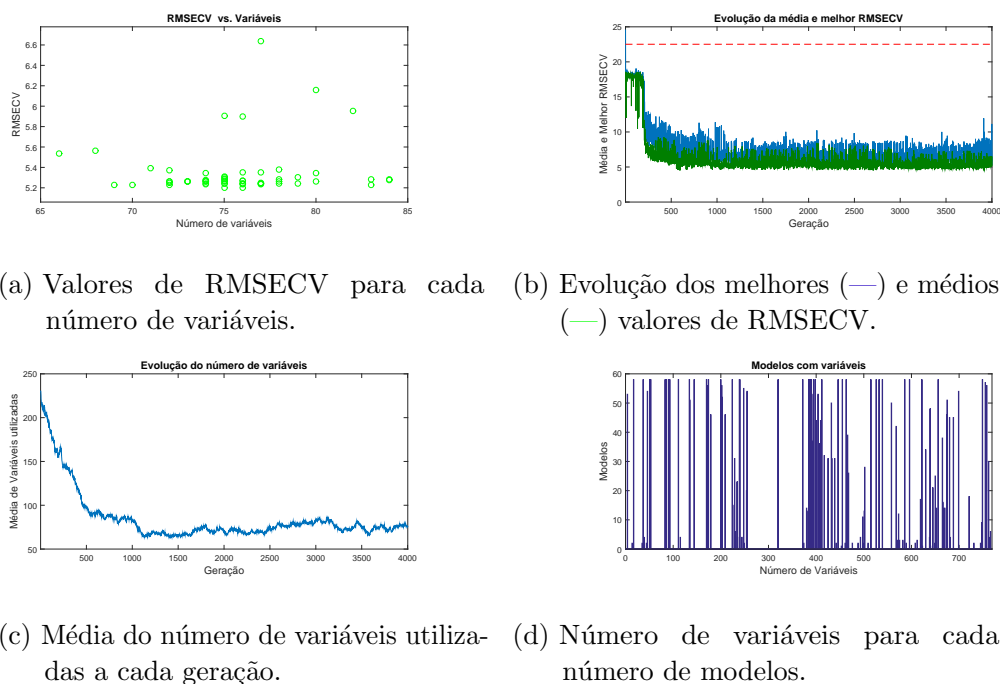


Figura 67 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB.

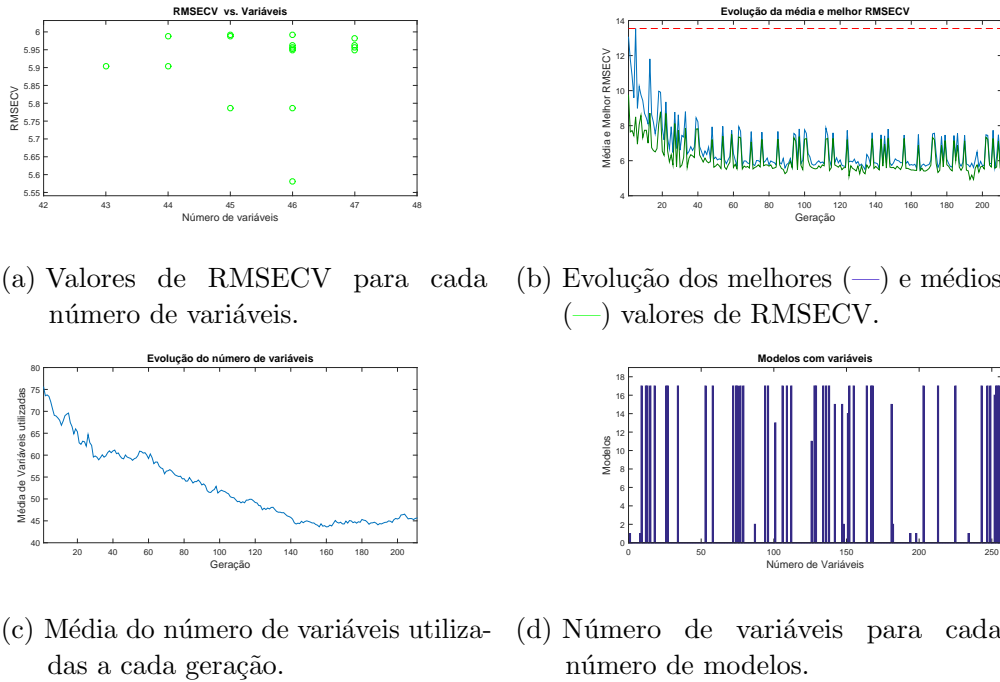


Figura 68 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor referente a escala cinza.

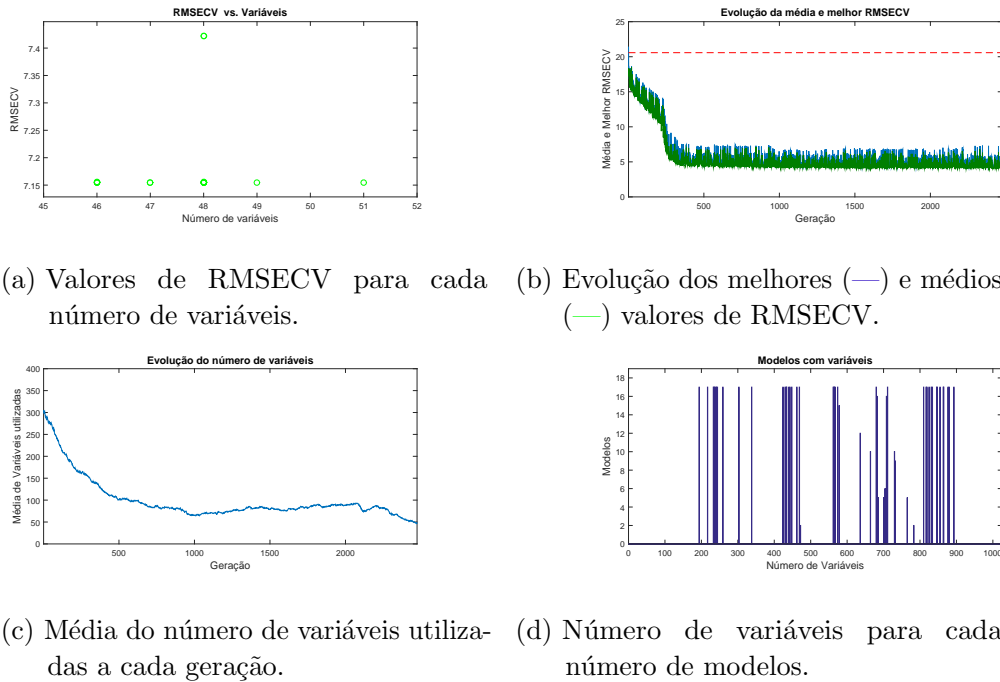


Figura 69 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + escala cinza.

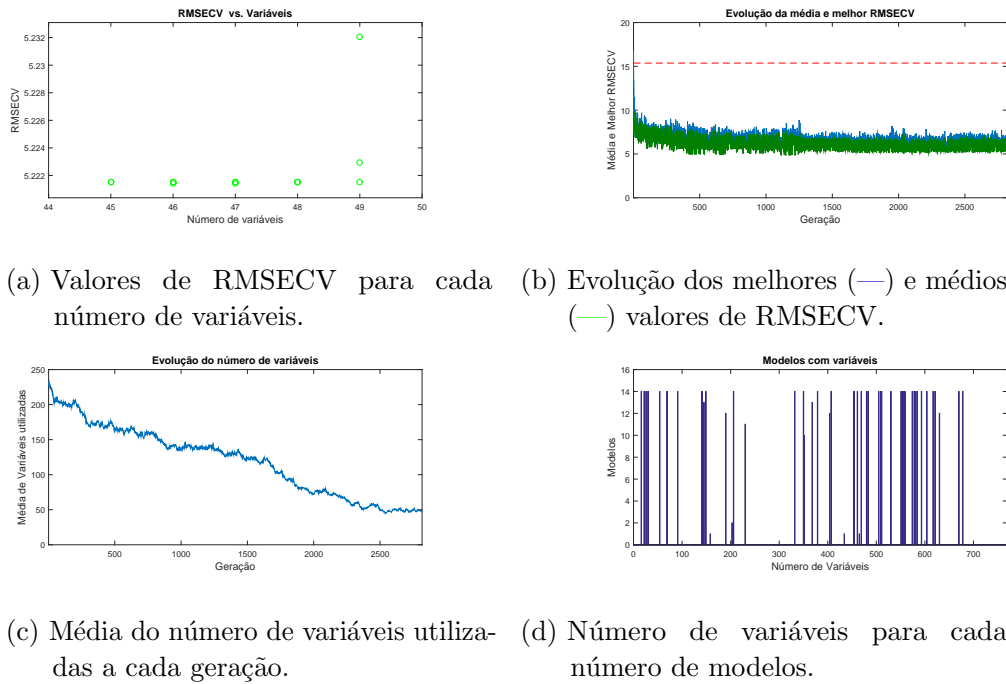


Figura 70 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI.

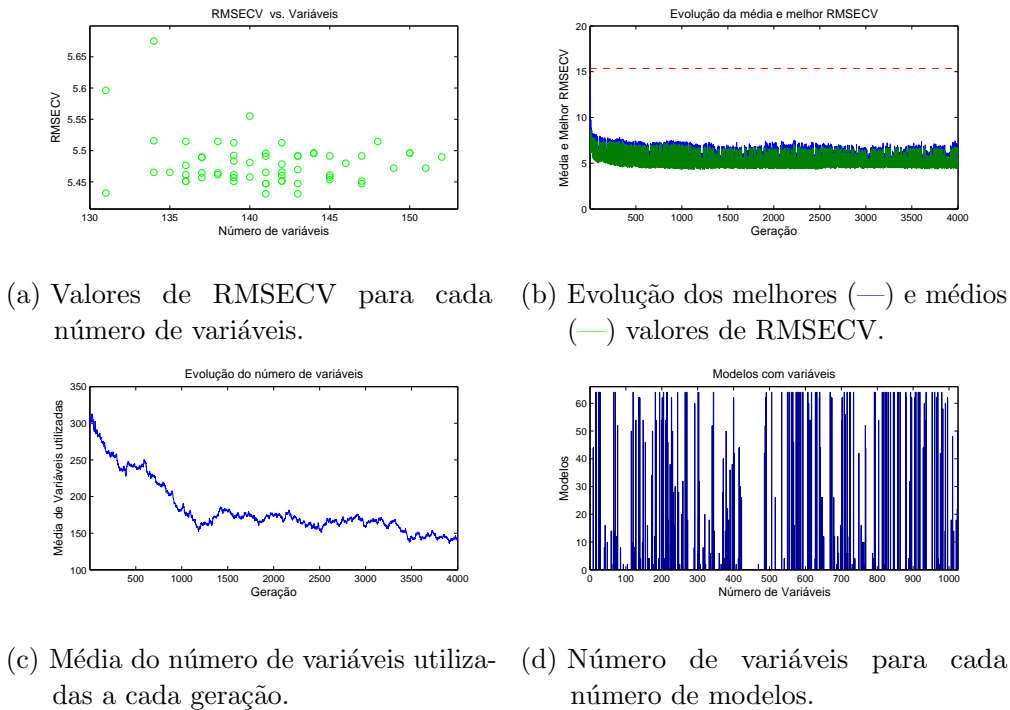


Figura 71 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor HSI + escala cinza.

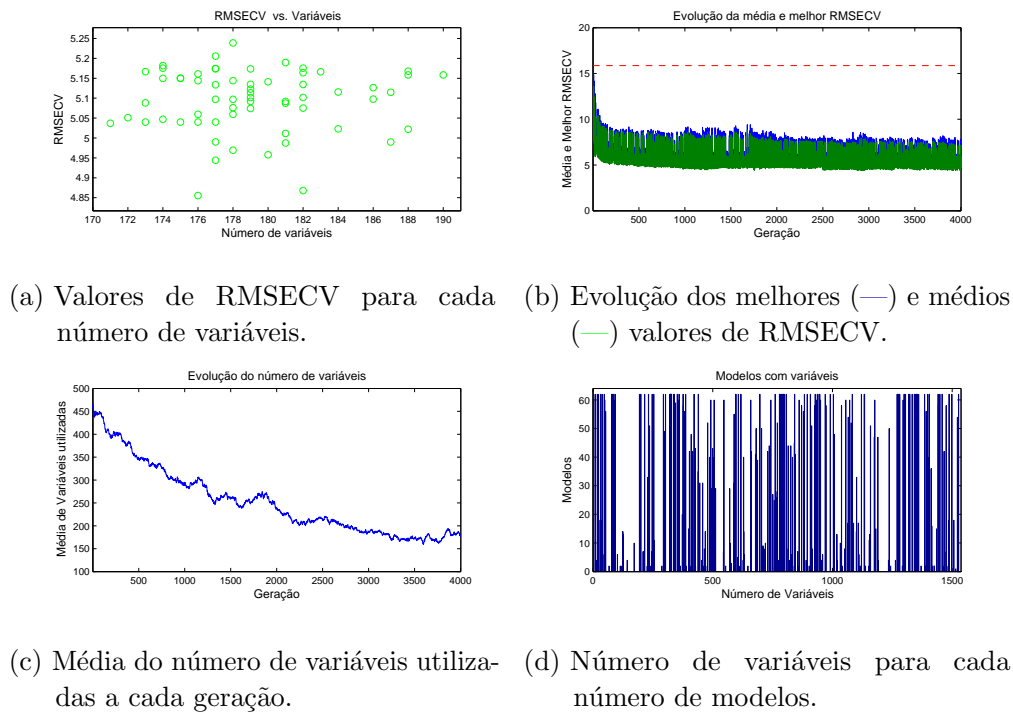


Figura 72 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI.

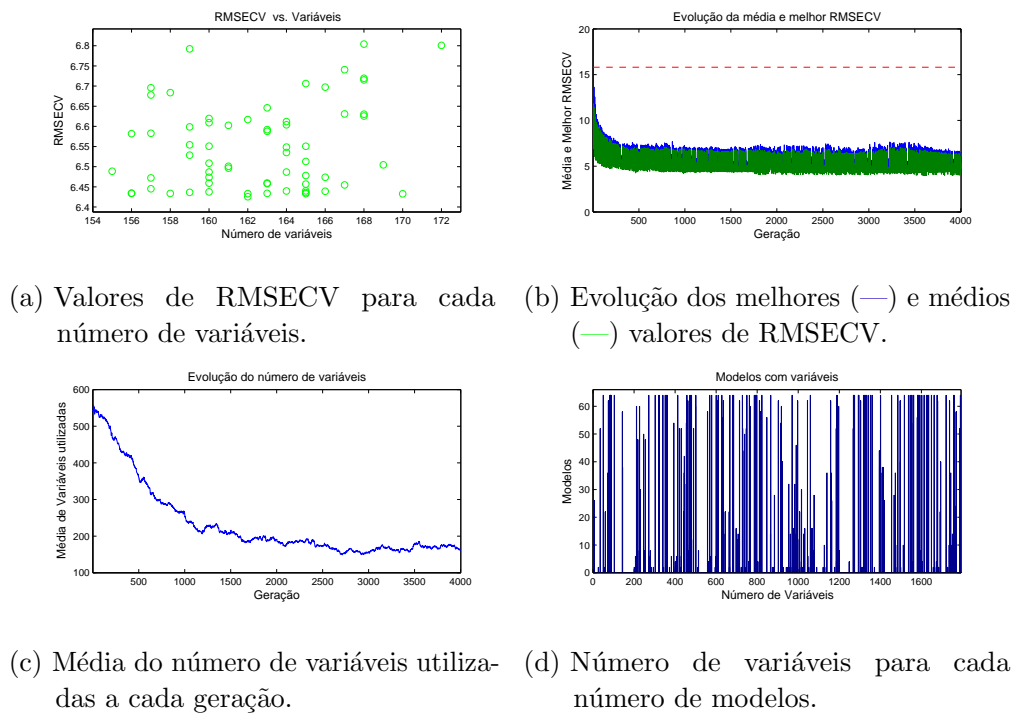


Figura 73 – Relatório de desempenho do algoritmo genético na seleção de variáveis no sistema de cor RGB + HSI + escala cinza.

## Apêndice J – Modelos de regressão GA-PLS para predição dos teores de areia em amostras de solo

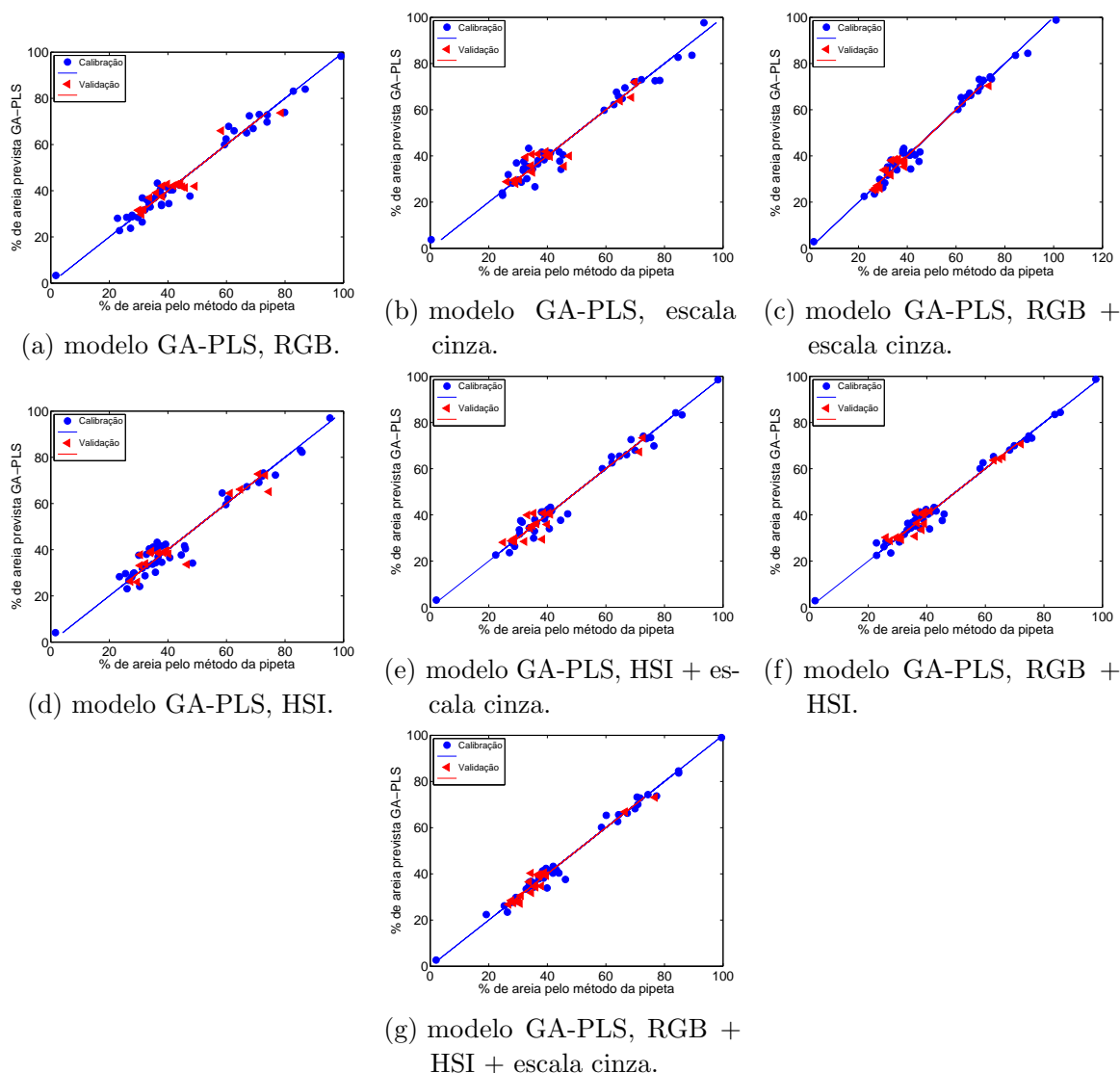


Figura 74 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

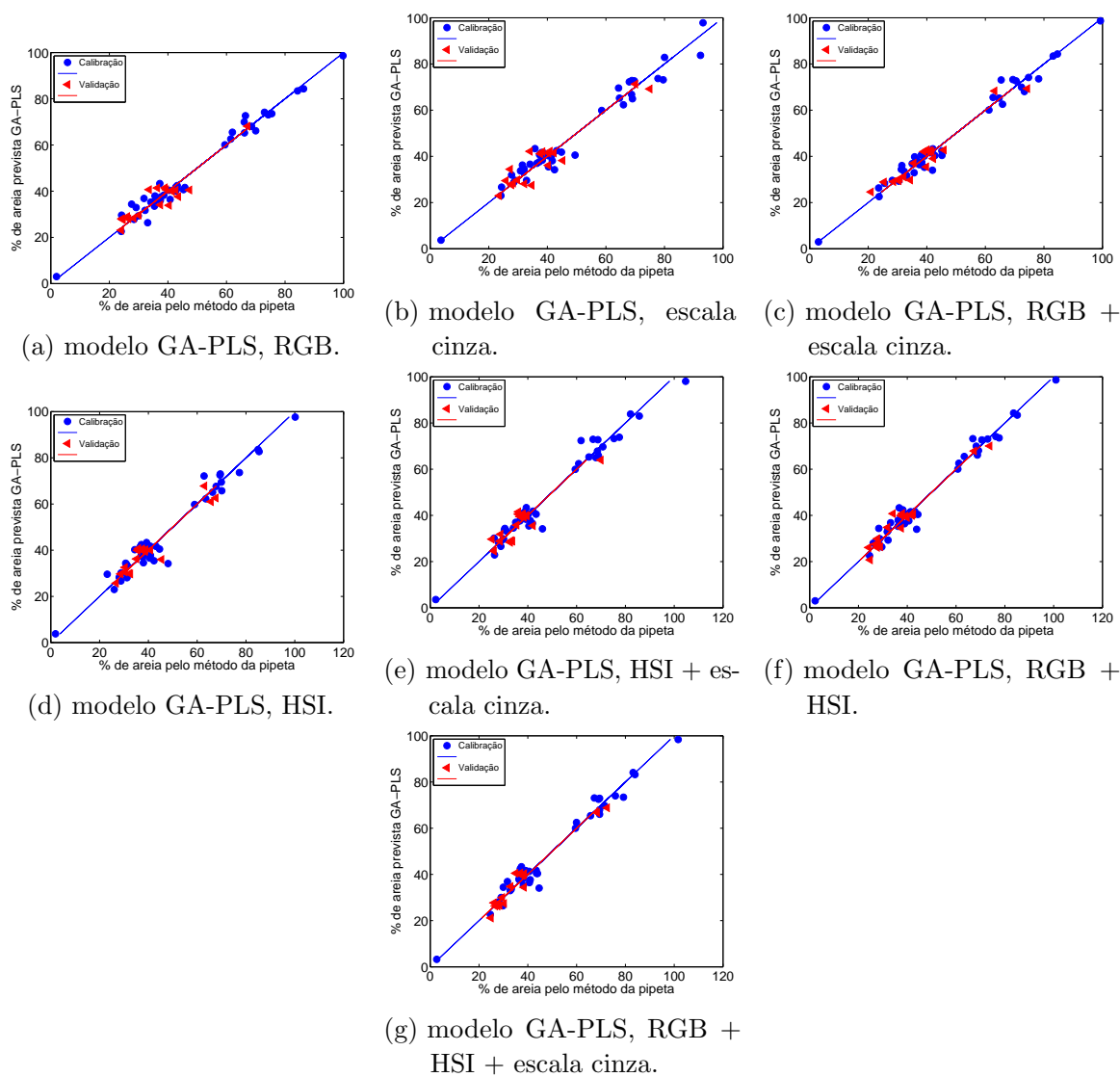


Figura 75 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

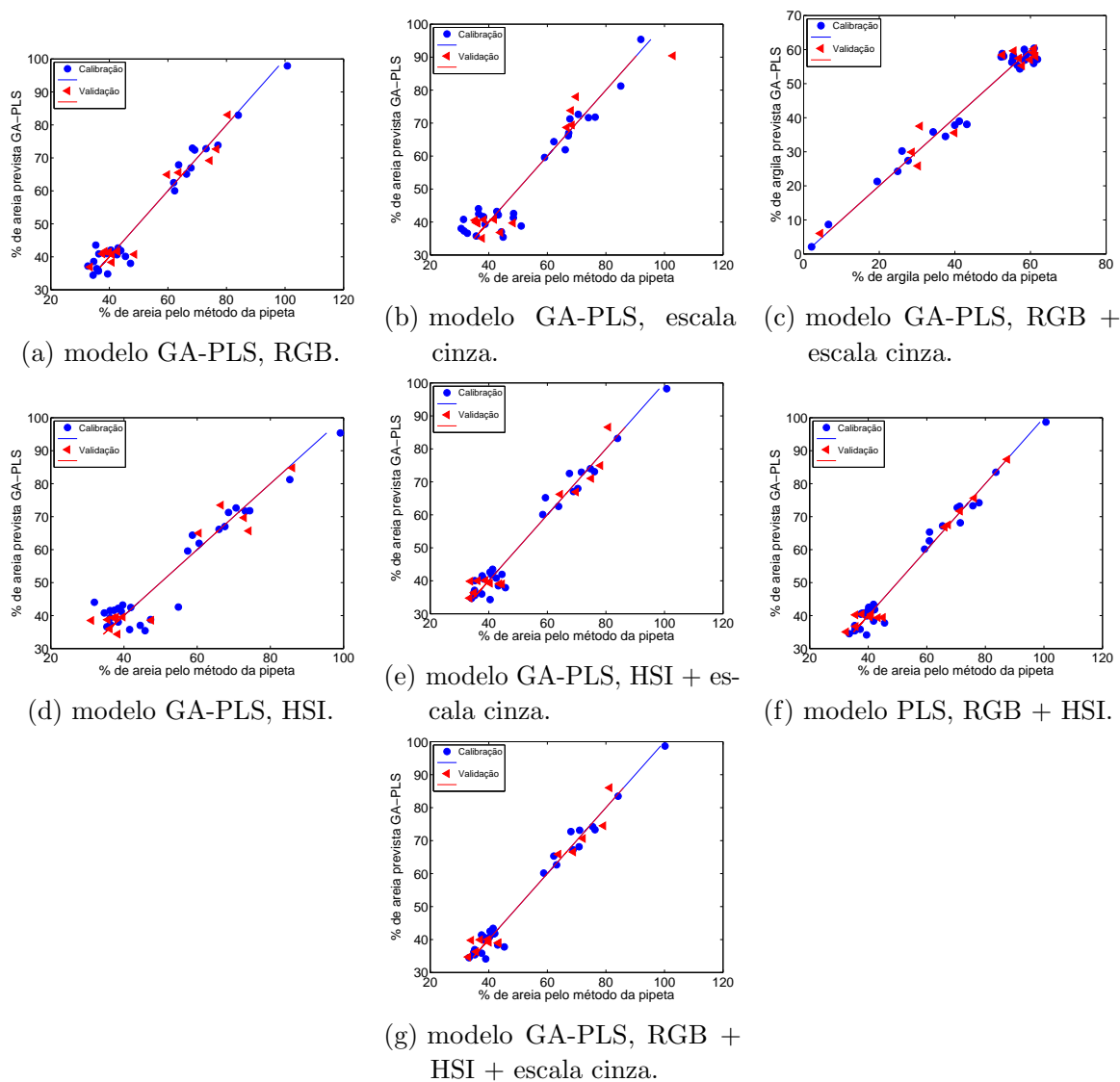


Figura 76 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.



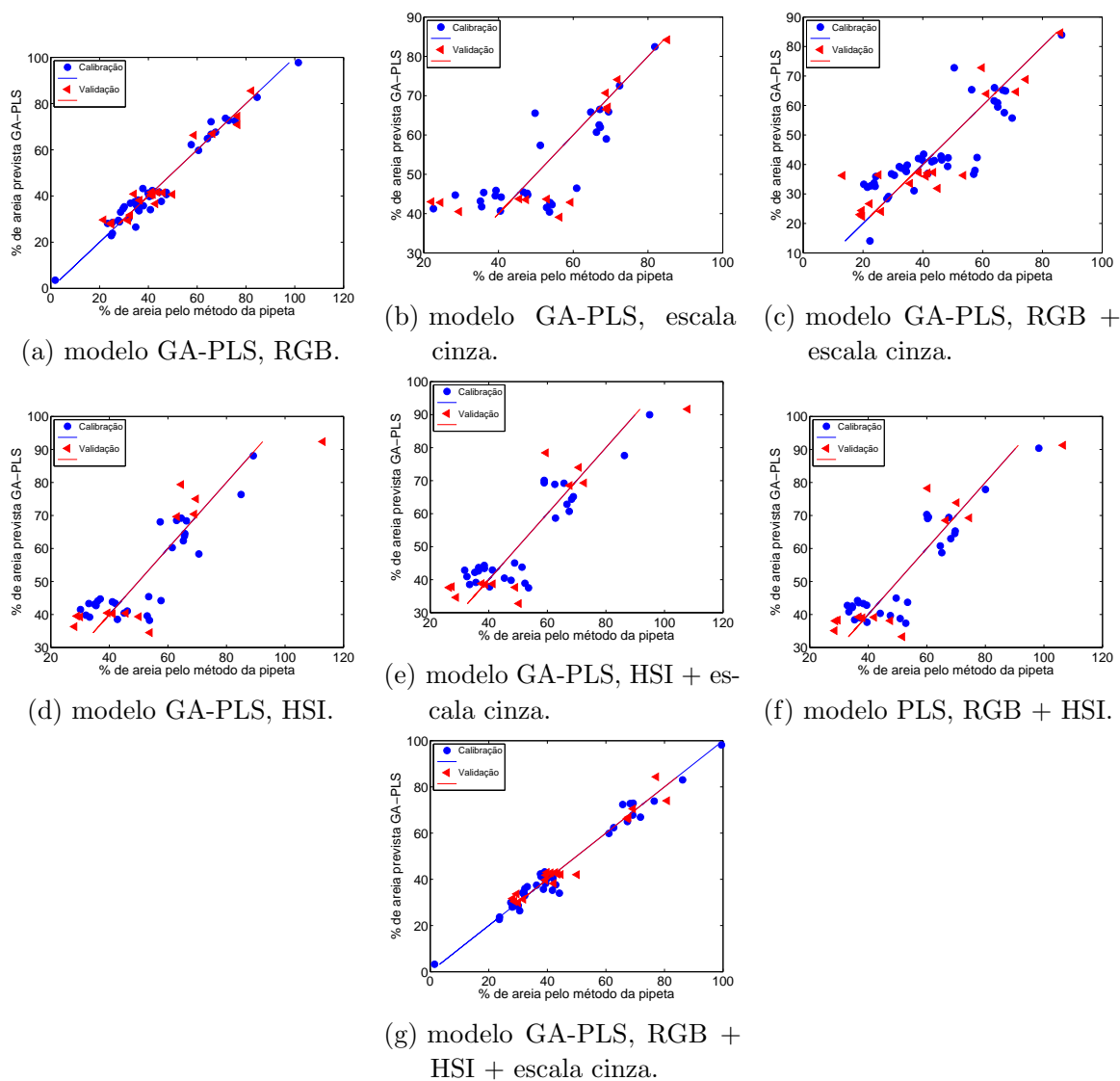


Figura 77 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice K – Modelos de regressão por mínimos quadrados PLS para predição dos teores de argila em amostras argilosas

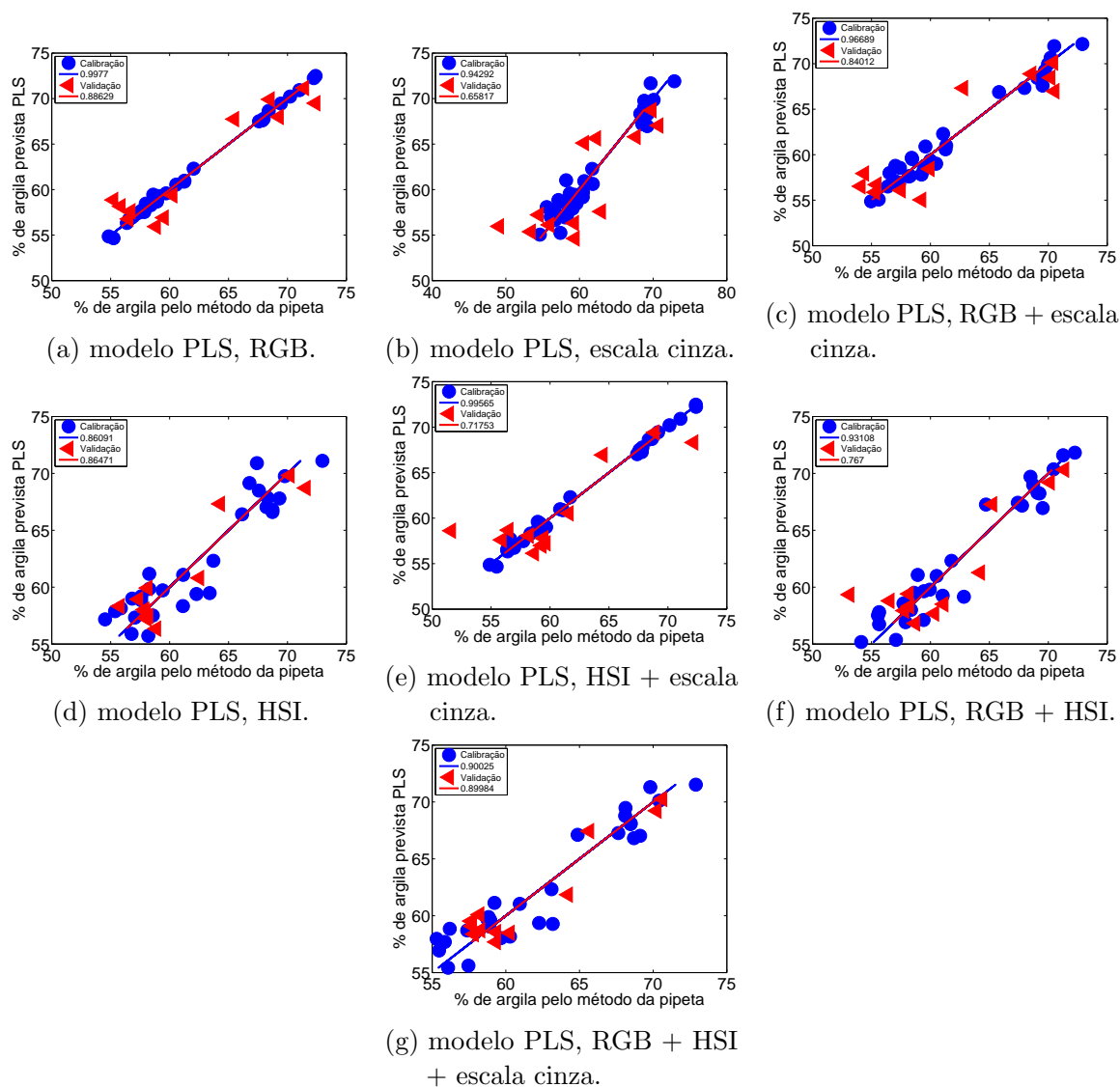


Figura 78 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 6,3x.

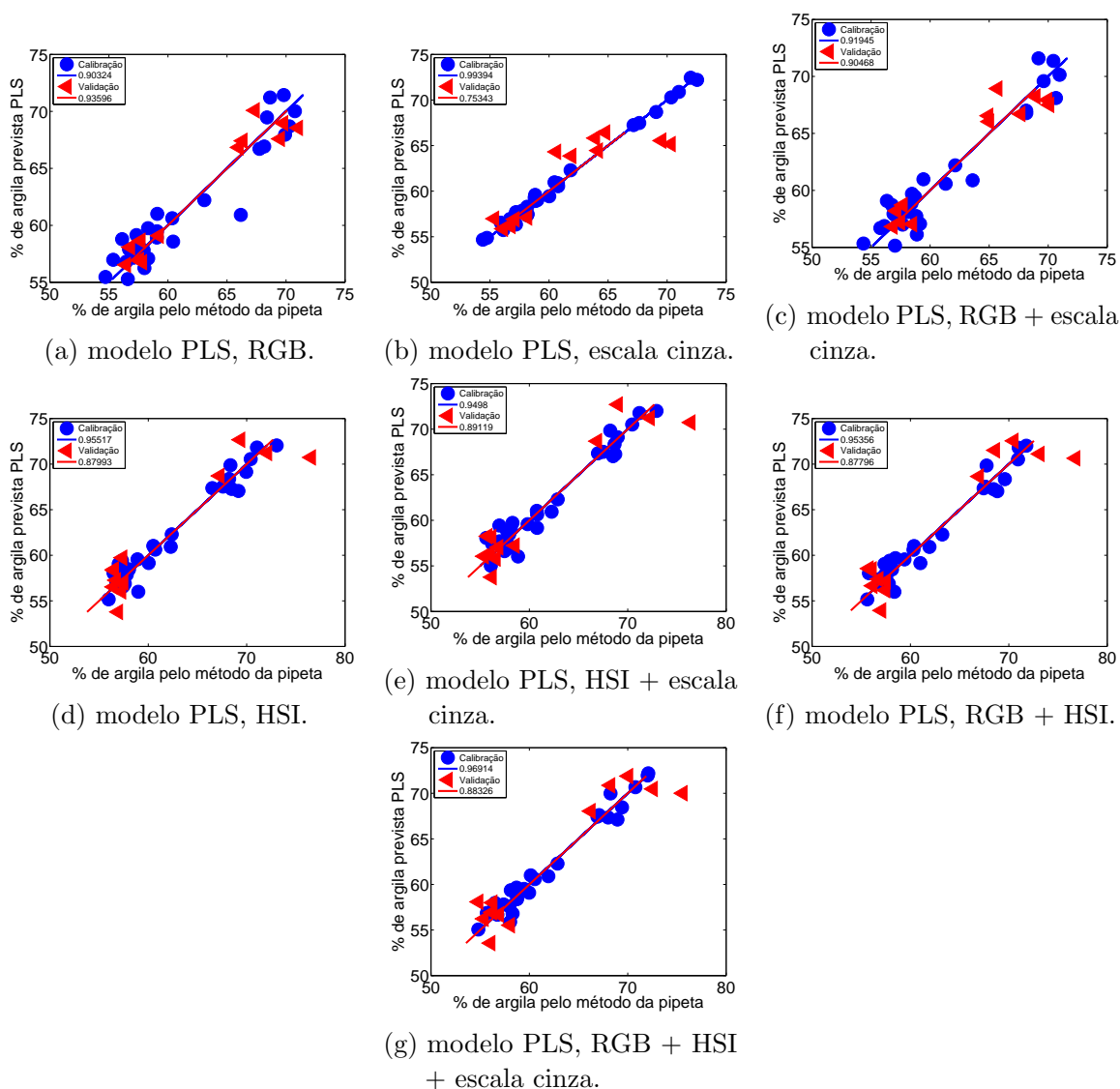


Figura 79 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 10,0x.

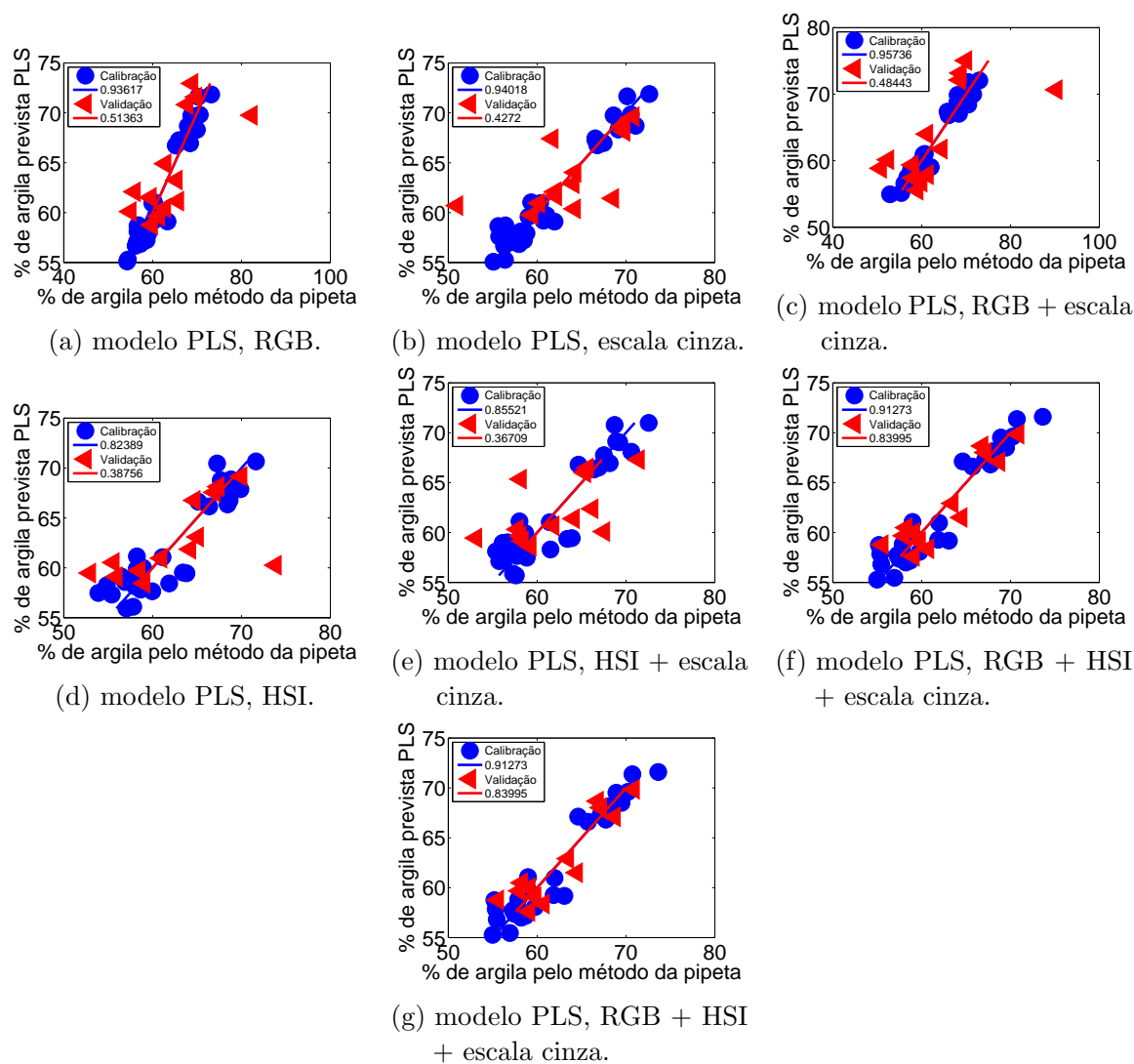


Figura 80 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.

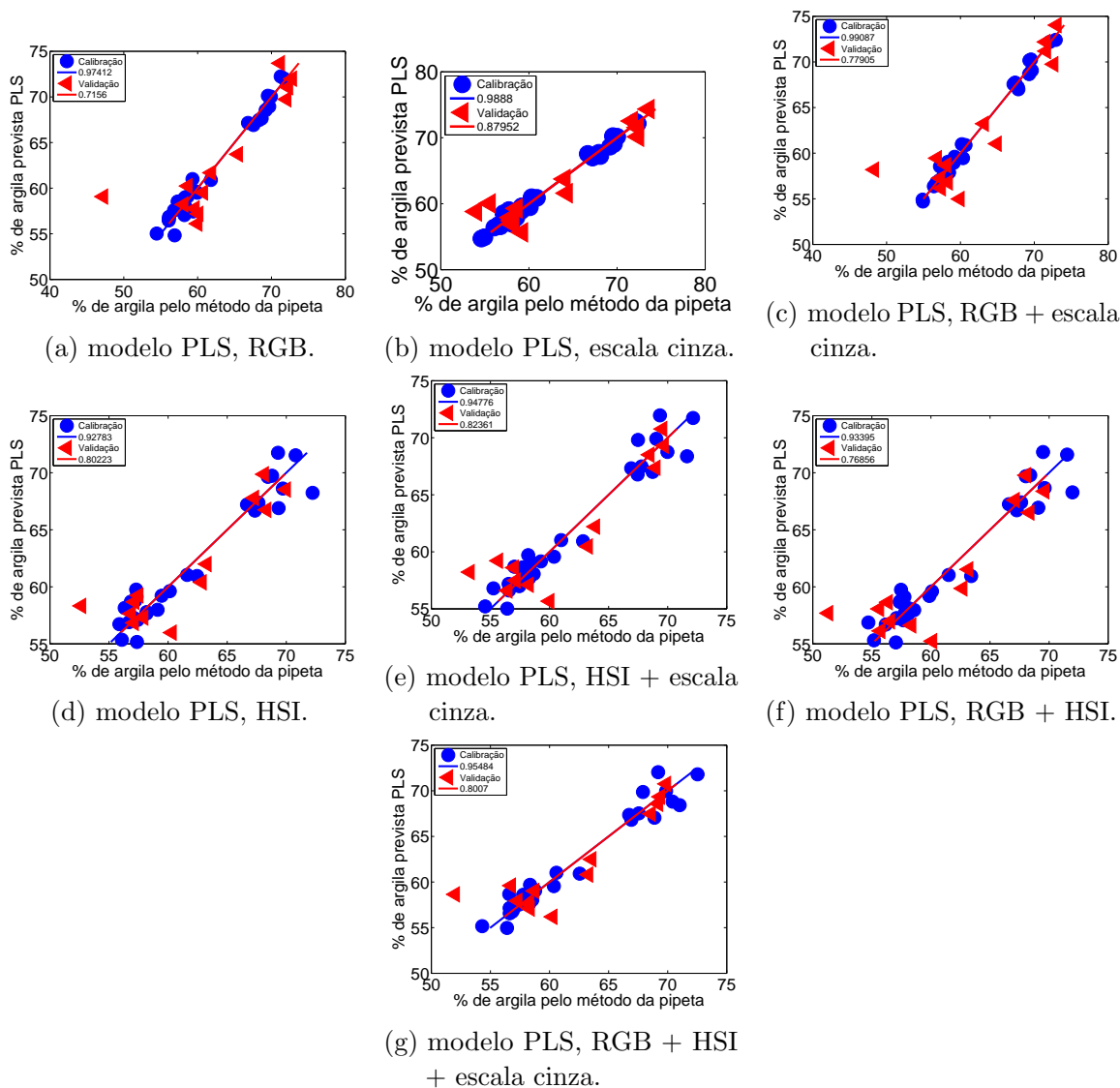


Figura 81 – Porcentagem de argila prevista pelo modelo PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice L – Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de argila em amostras argilosas

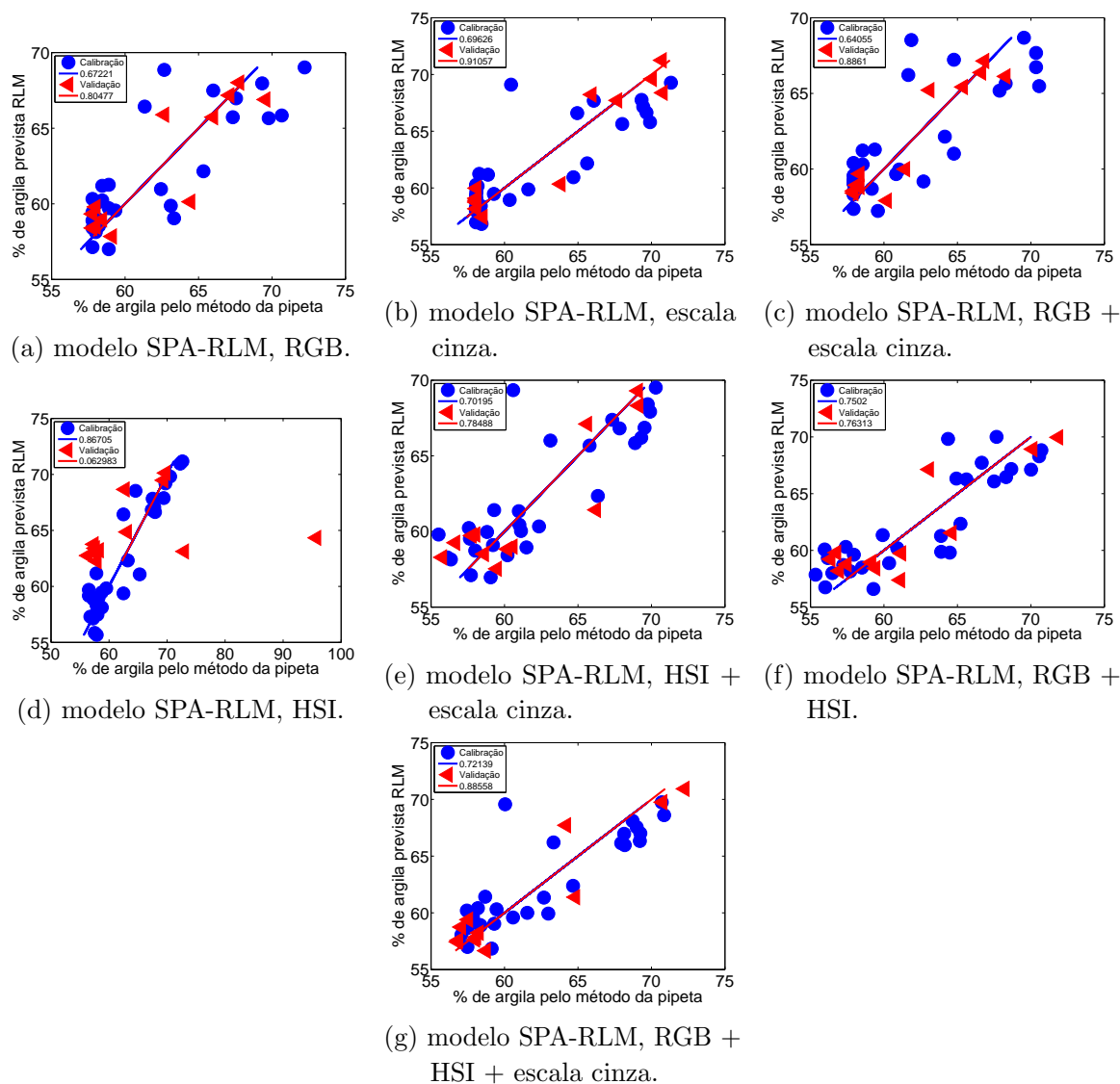


Figura 82 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 6,3x.

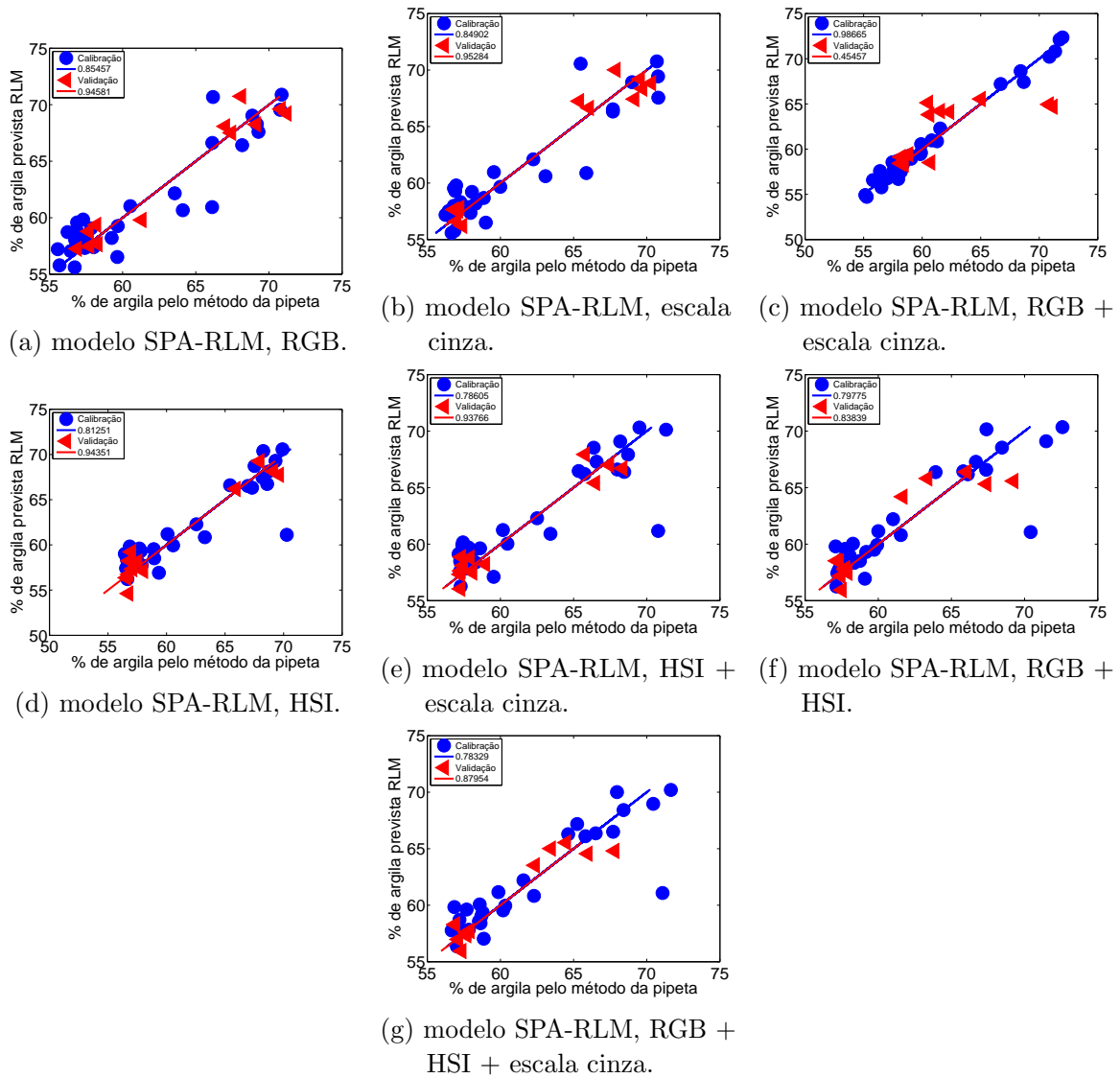


Figura 83 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁) referente ao aumento óptico de 10,0x.

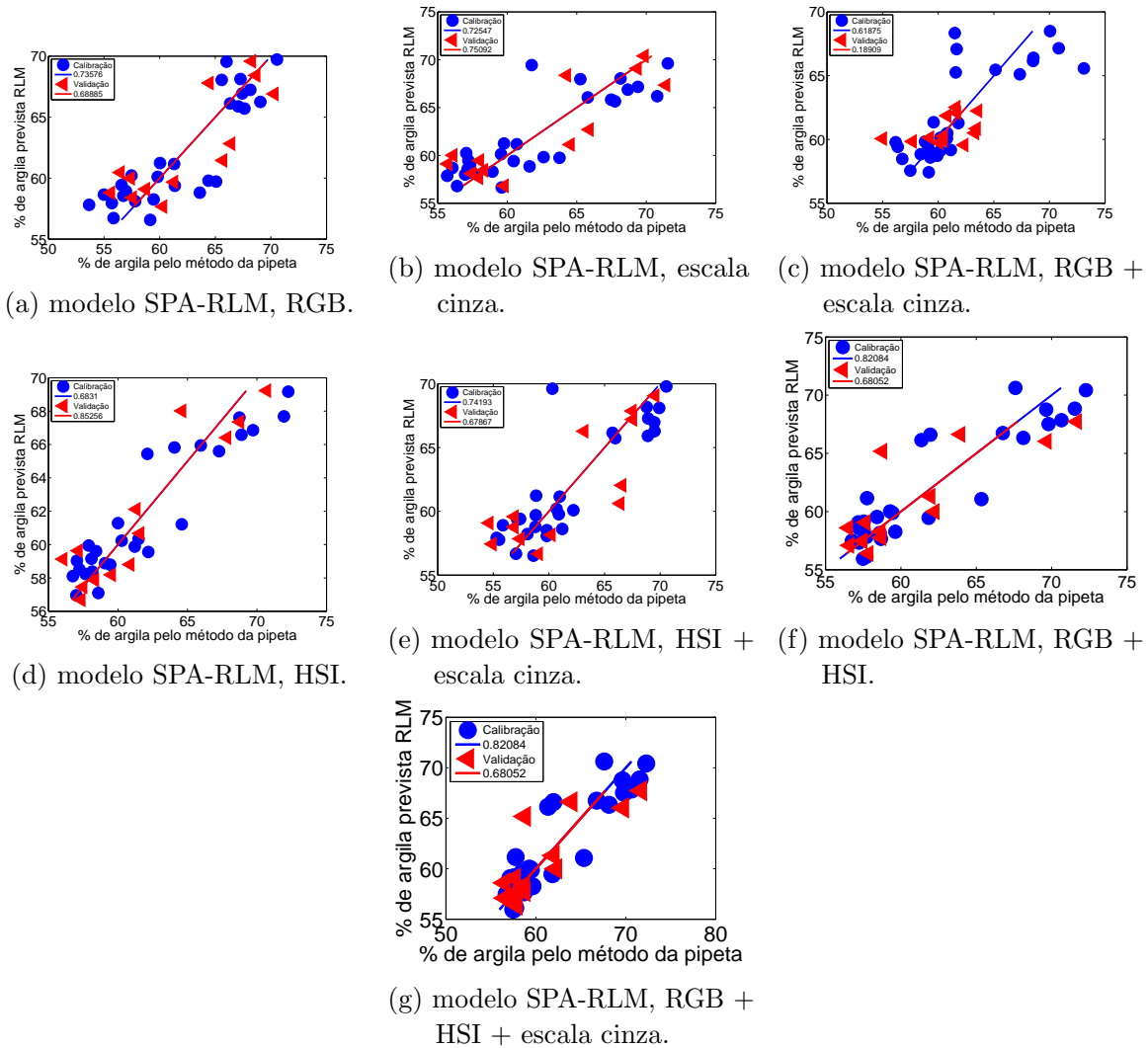


Figura 84 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.



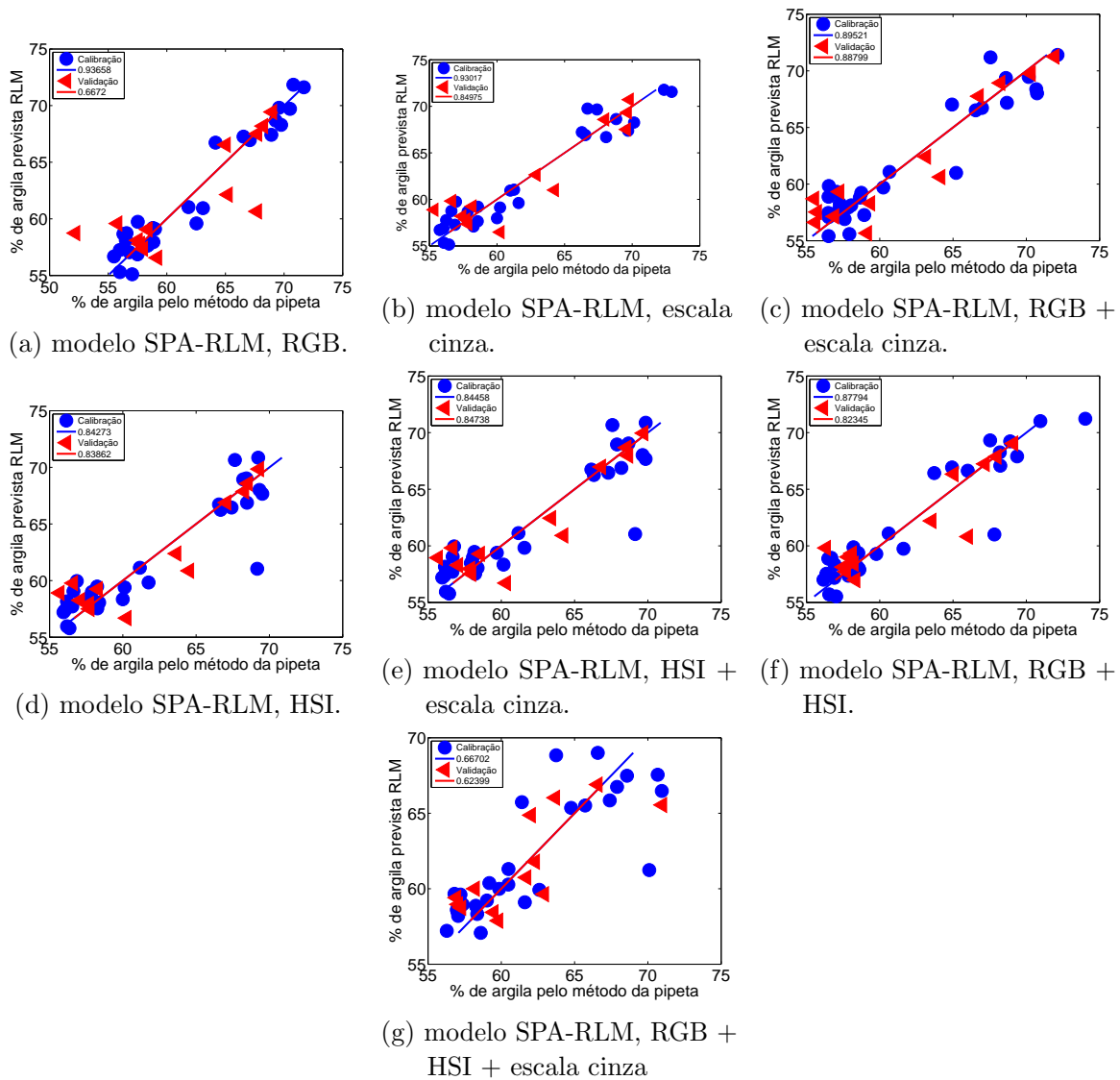


Figura 85 – Porcentagem de argila prevista pelo modelo SPA-RLM versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice M – Modelos de regressão GA-PLS para predição dos teores de argila em amostras argilosas

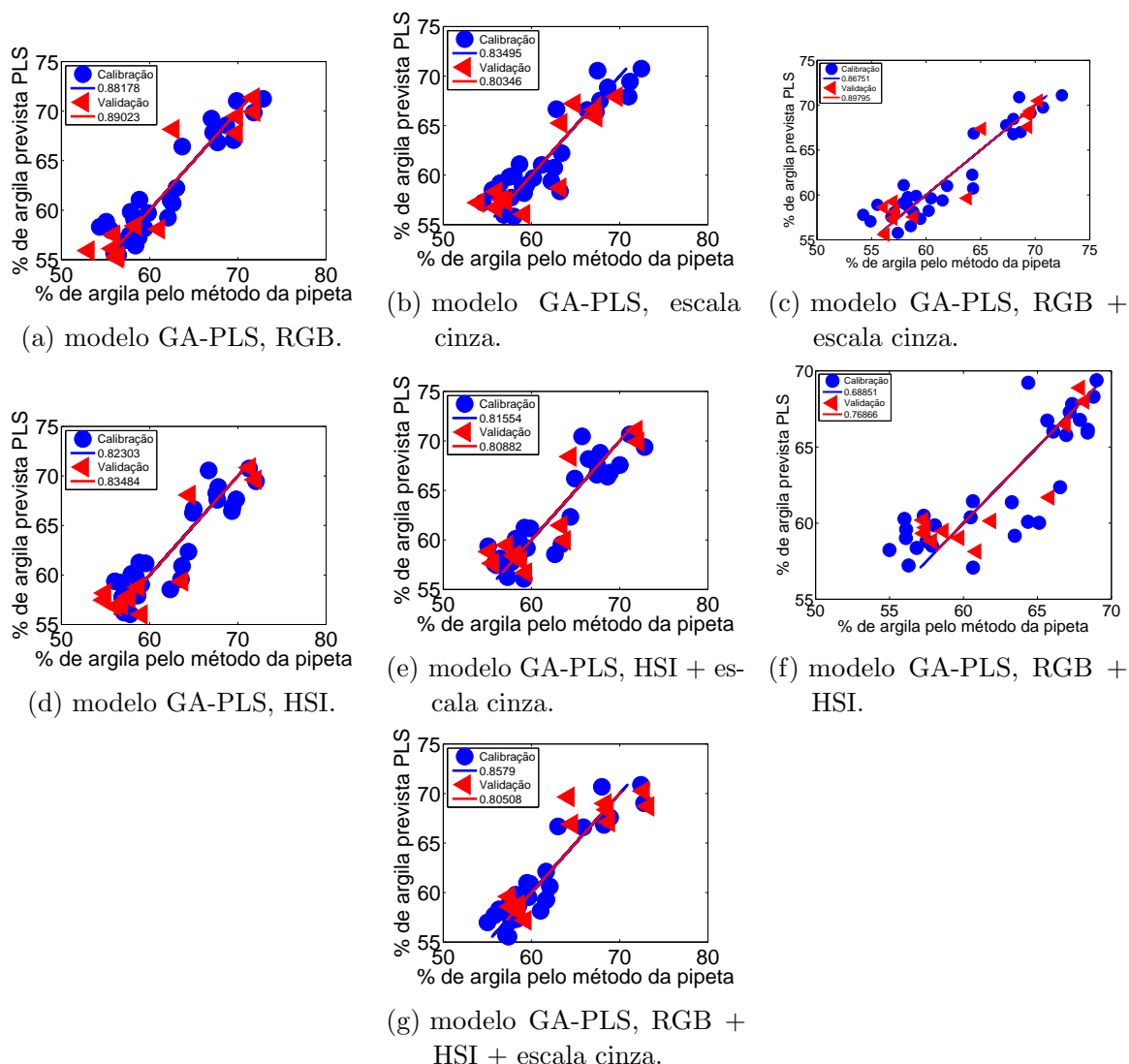


Figura 86 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

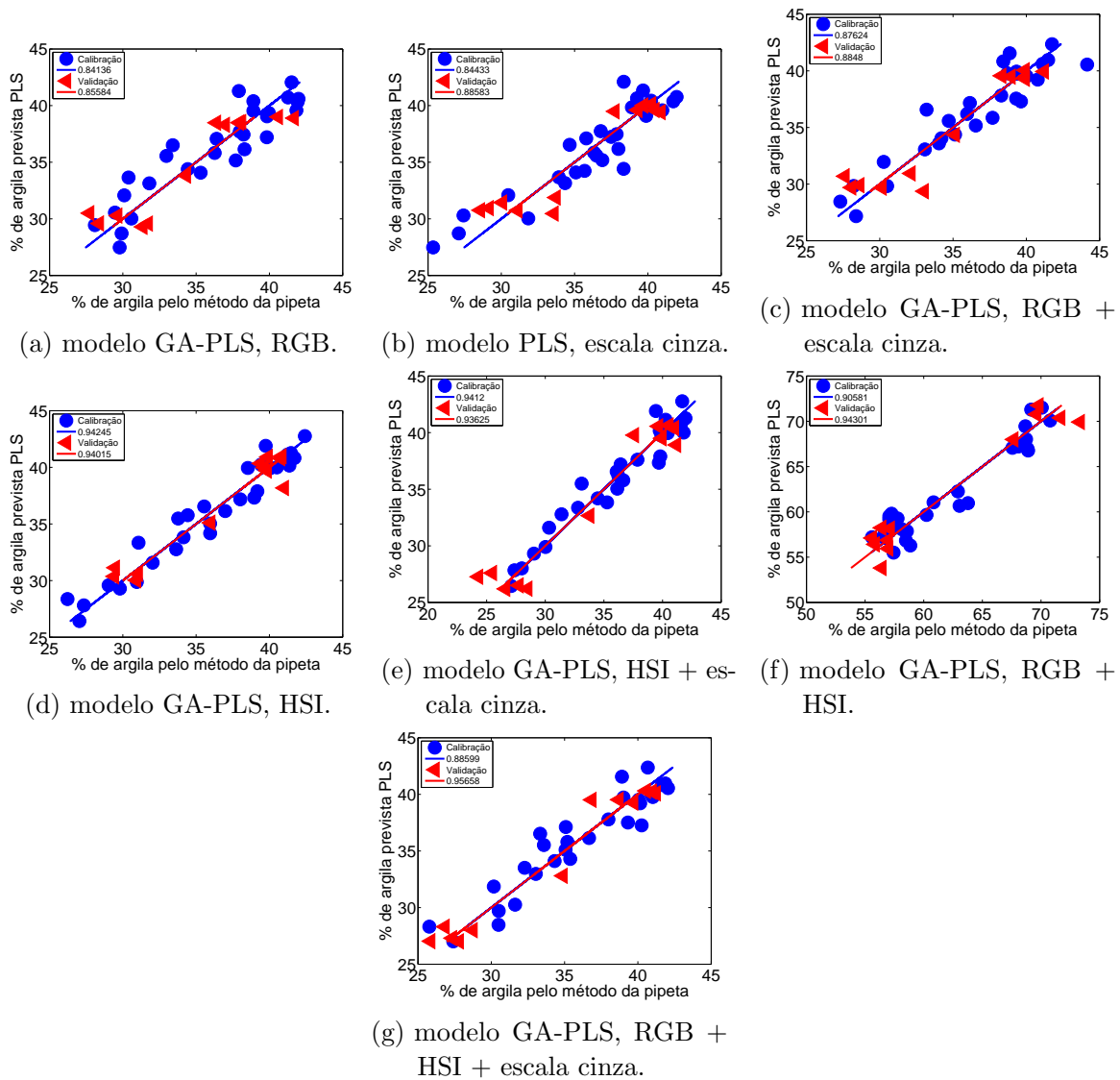


Figura 87 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

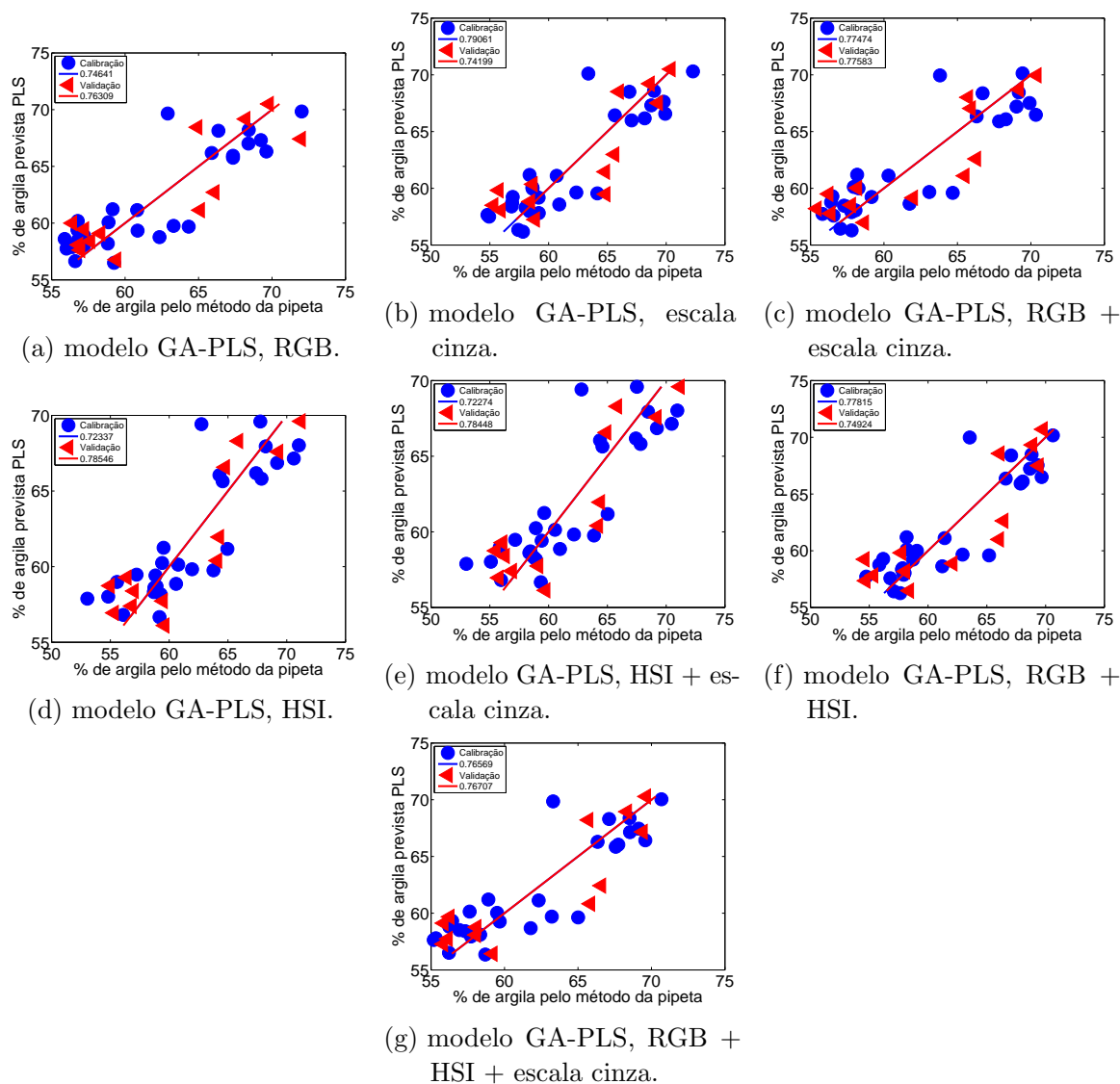


Figura 88 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.

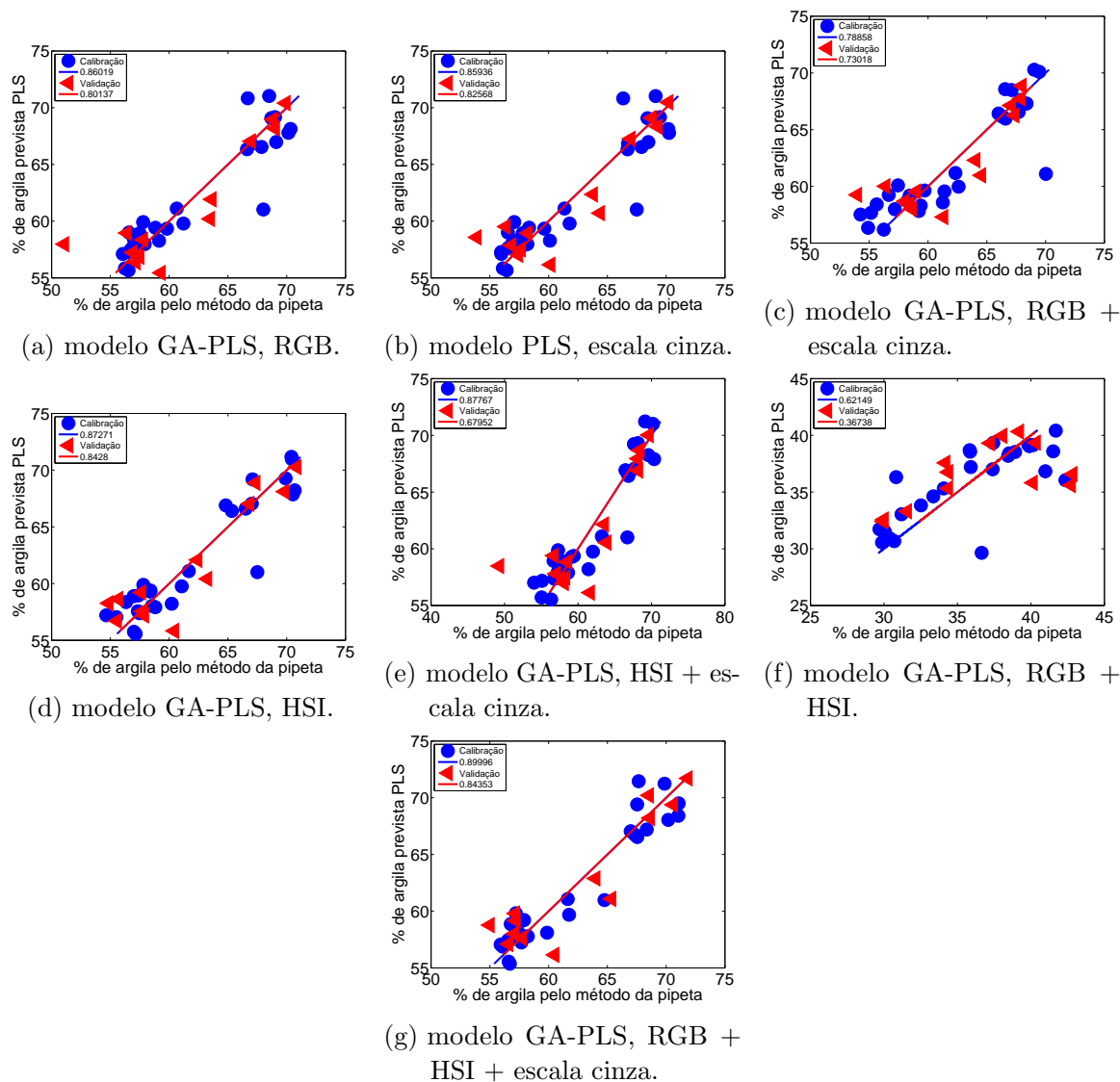


Figura 89 – Porcentagem de argila prevista pelo modelo GA-PLS versus porcentagem de argila determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice N – Modelos de regressão por mínimos quadrados PLS para predição dos teores de areia em amostras argilosas

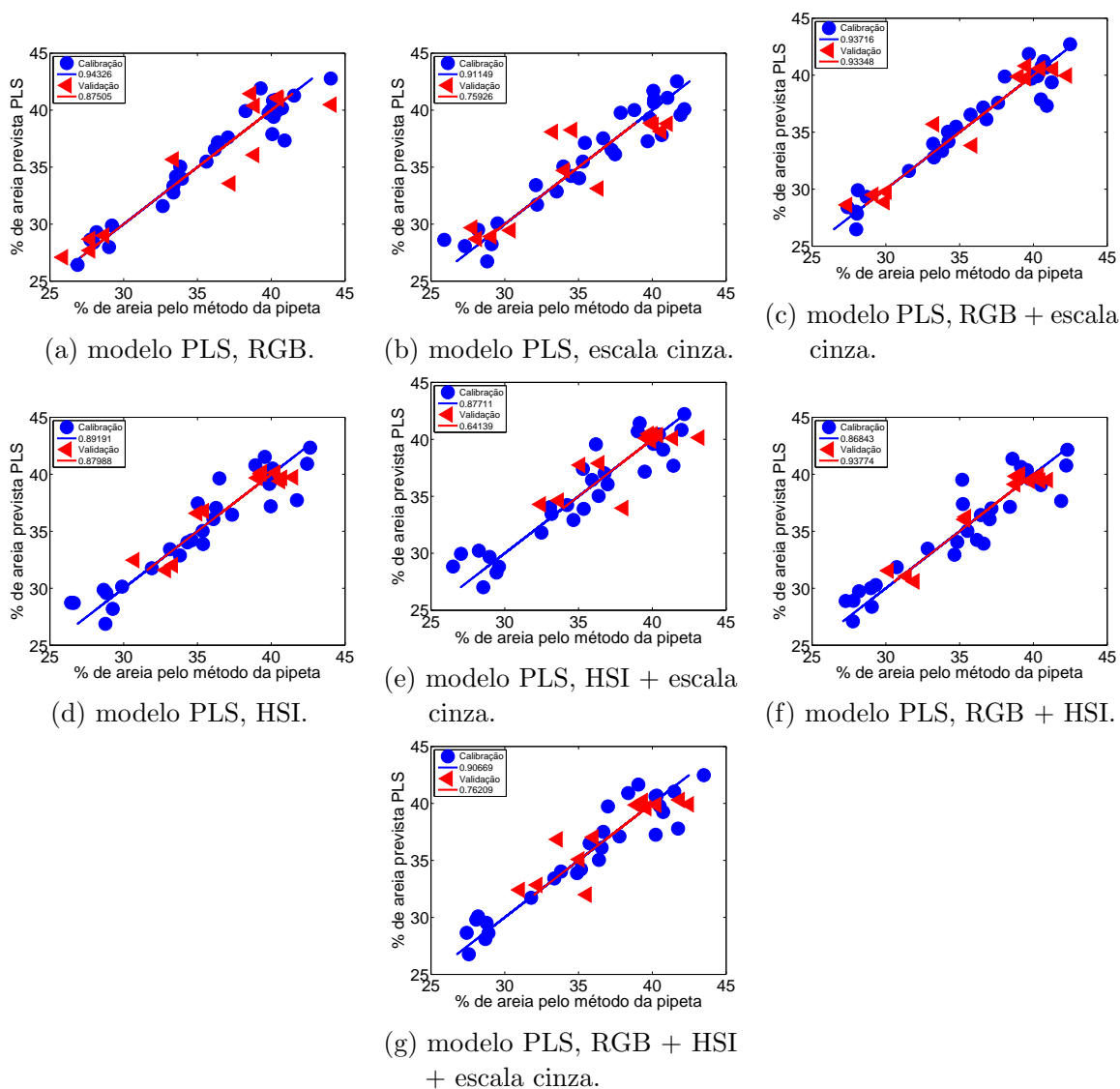


Figura 90 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 6,3x.

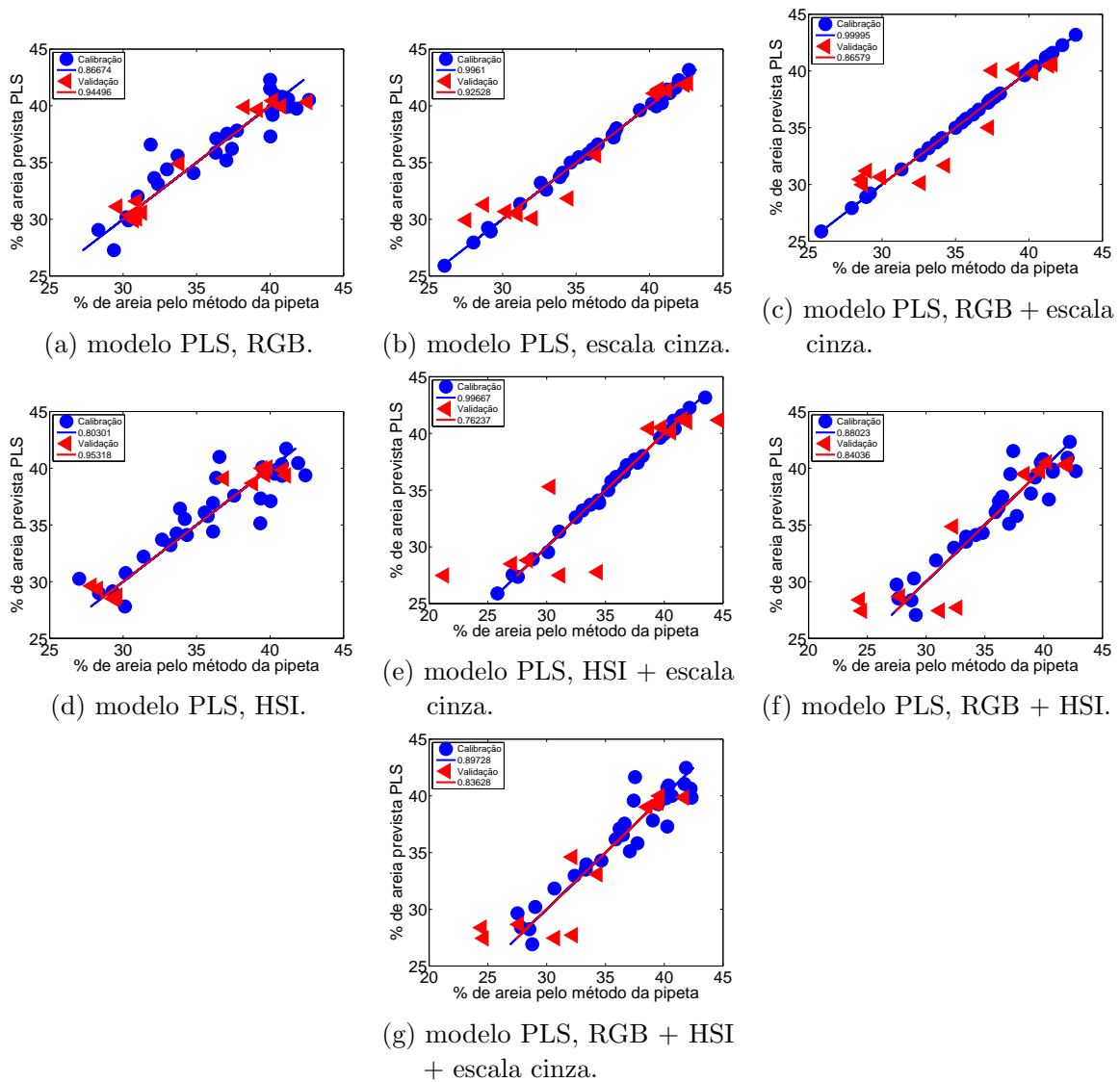


Figura 91 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 10,0x.

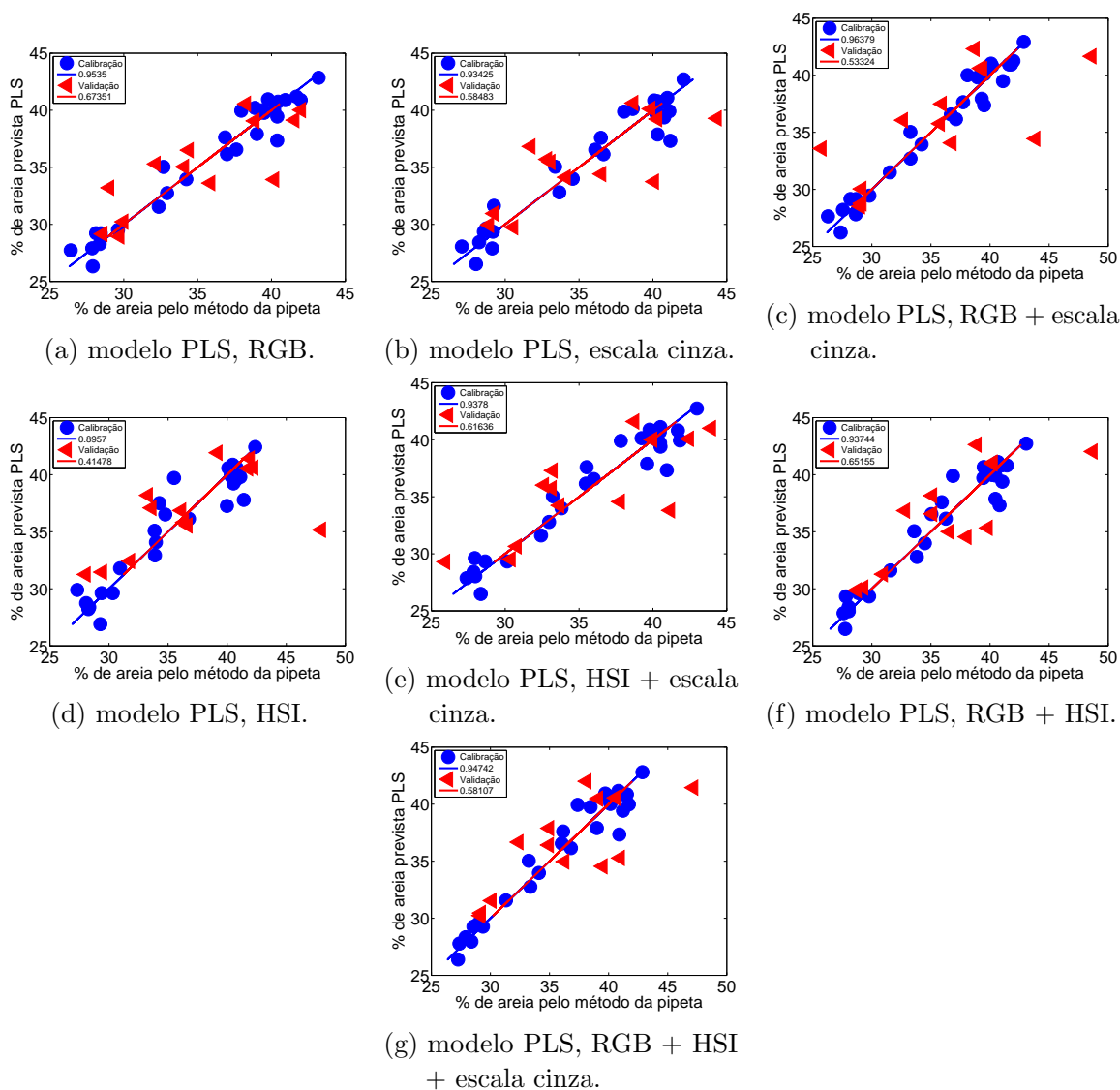


Figura 92 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.



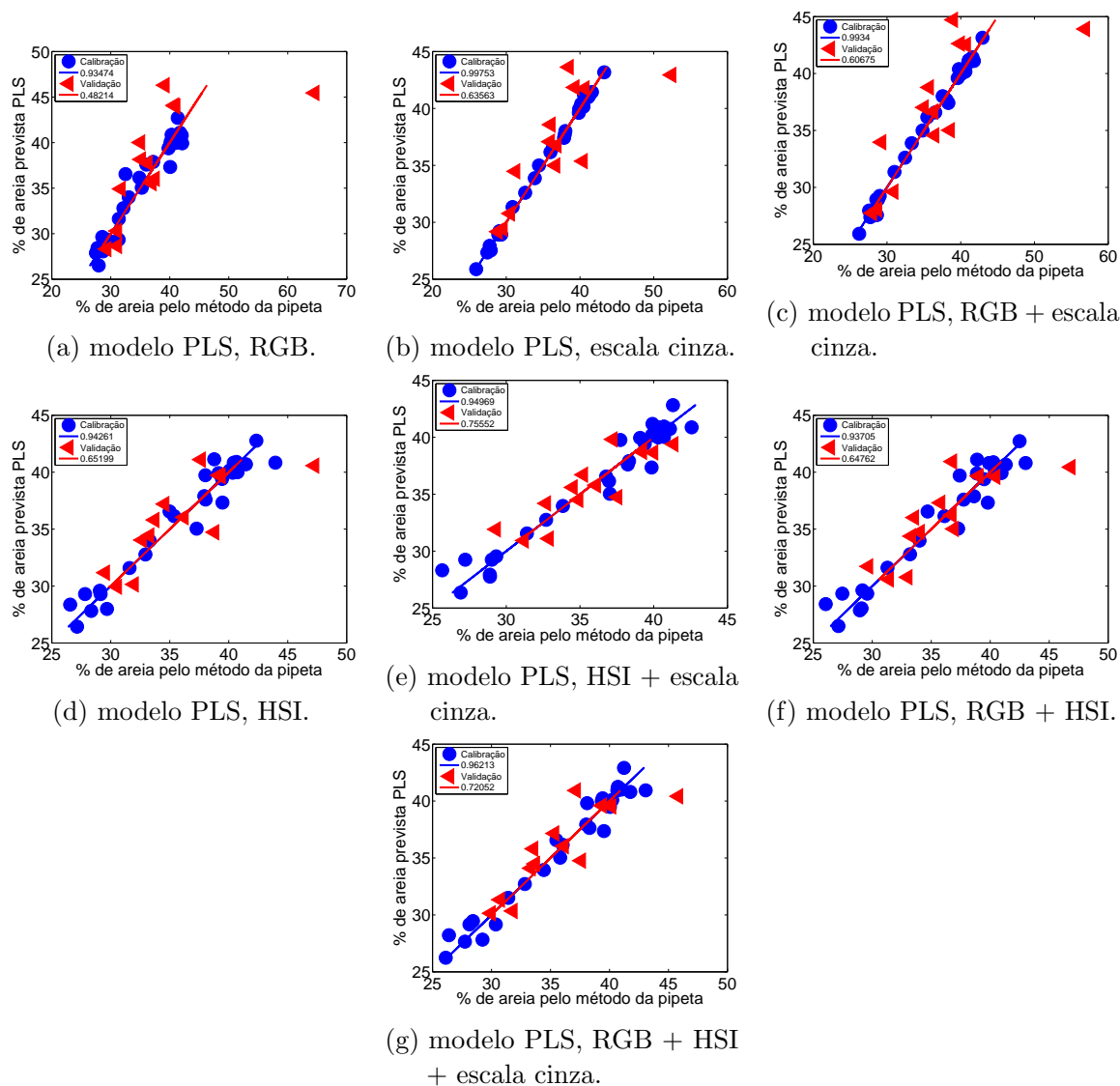


Figura 93 – Porcentagem de areia prevista pelo modelo PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice O – Modelos de regressão linear múltipla associados com o algoritmo das projeções sucessivas (SPA-RLM) para predição dos teores de areia em amostras argilosas

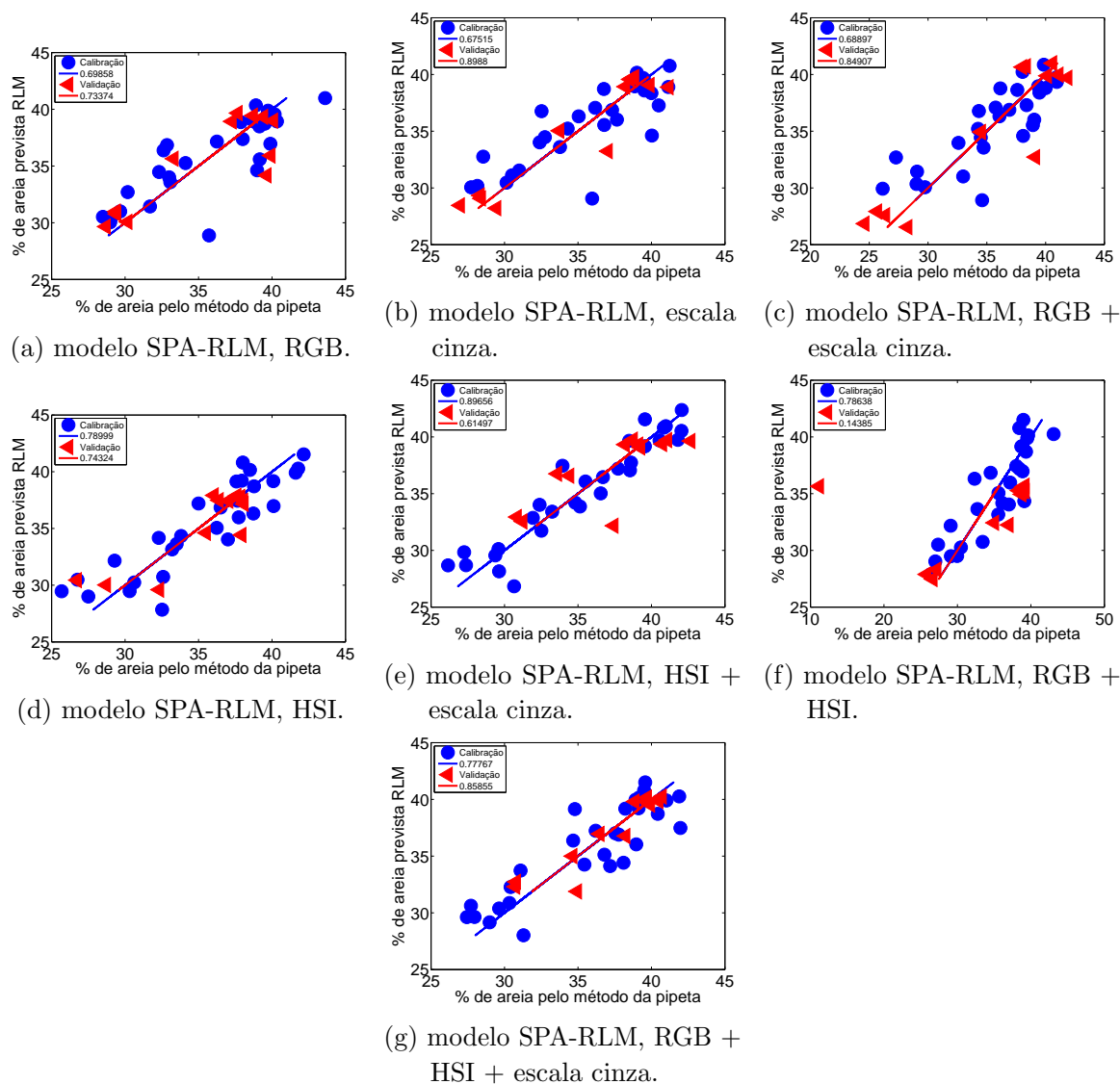


Figura 94 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 6,3x.

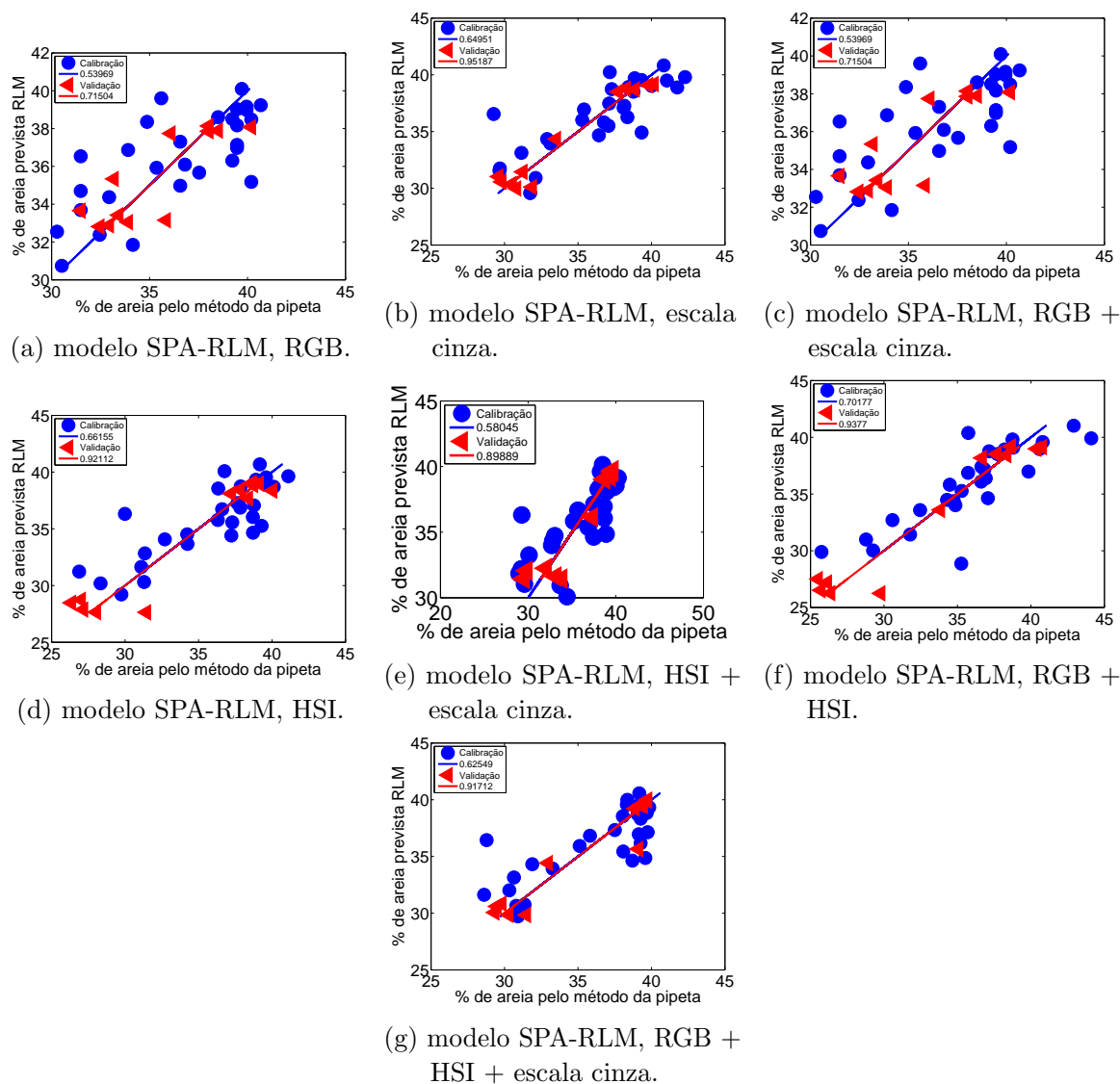


Figura 95 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◄), subdivididas utilizando o algoritmo SPXY e referente ao aumento óptico de 10,0x.

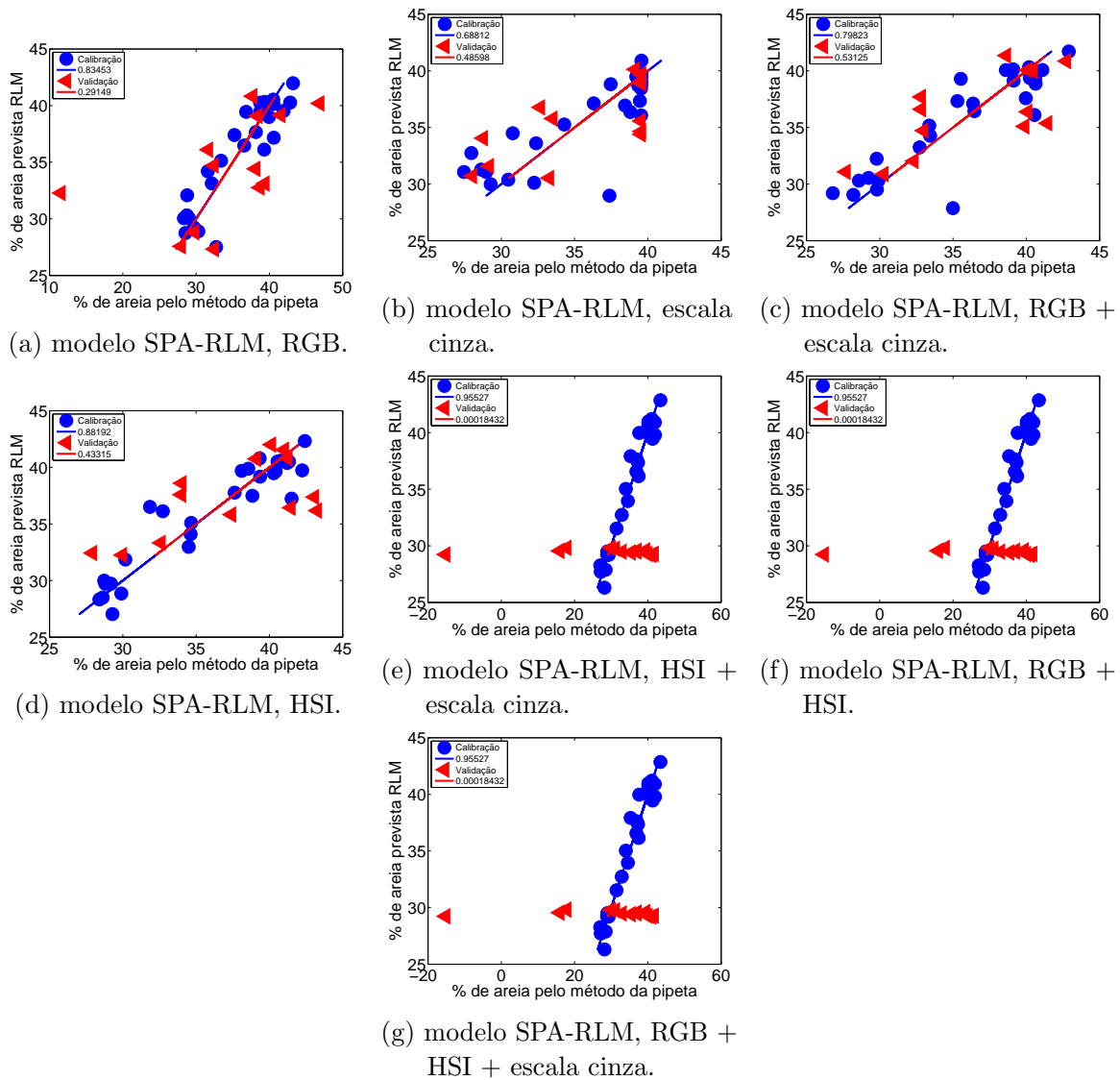


Figura 96 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.

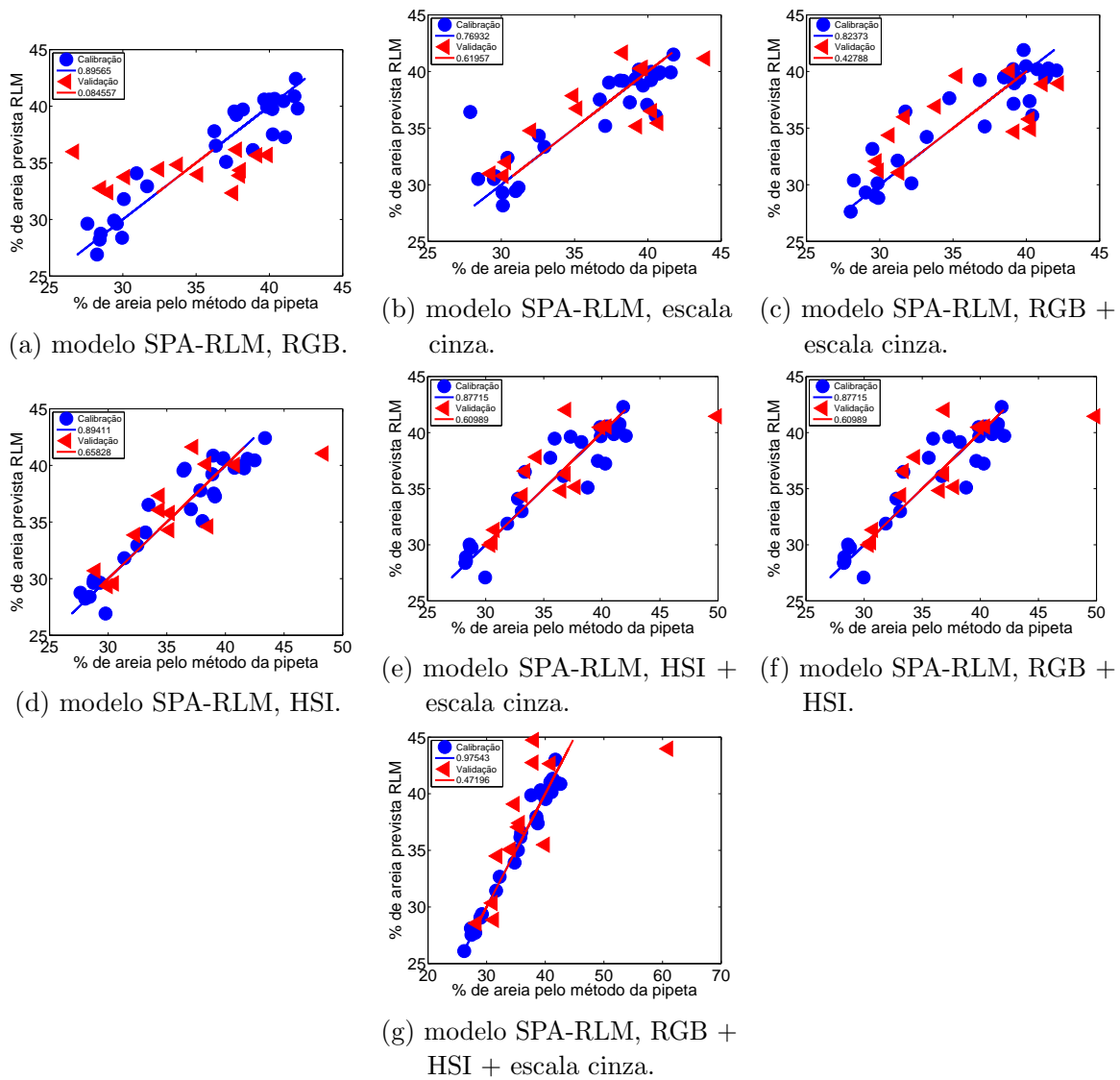


Figura 97 – Porcentagem de areia prevista pelo modelo SPA-RLM versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas utilizando o particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.

## Apêndice P – Modelos de regressão GA-PLS para predição dos teores de areia em amostras argilosas

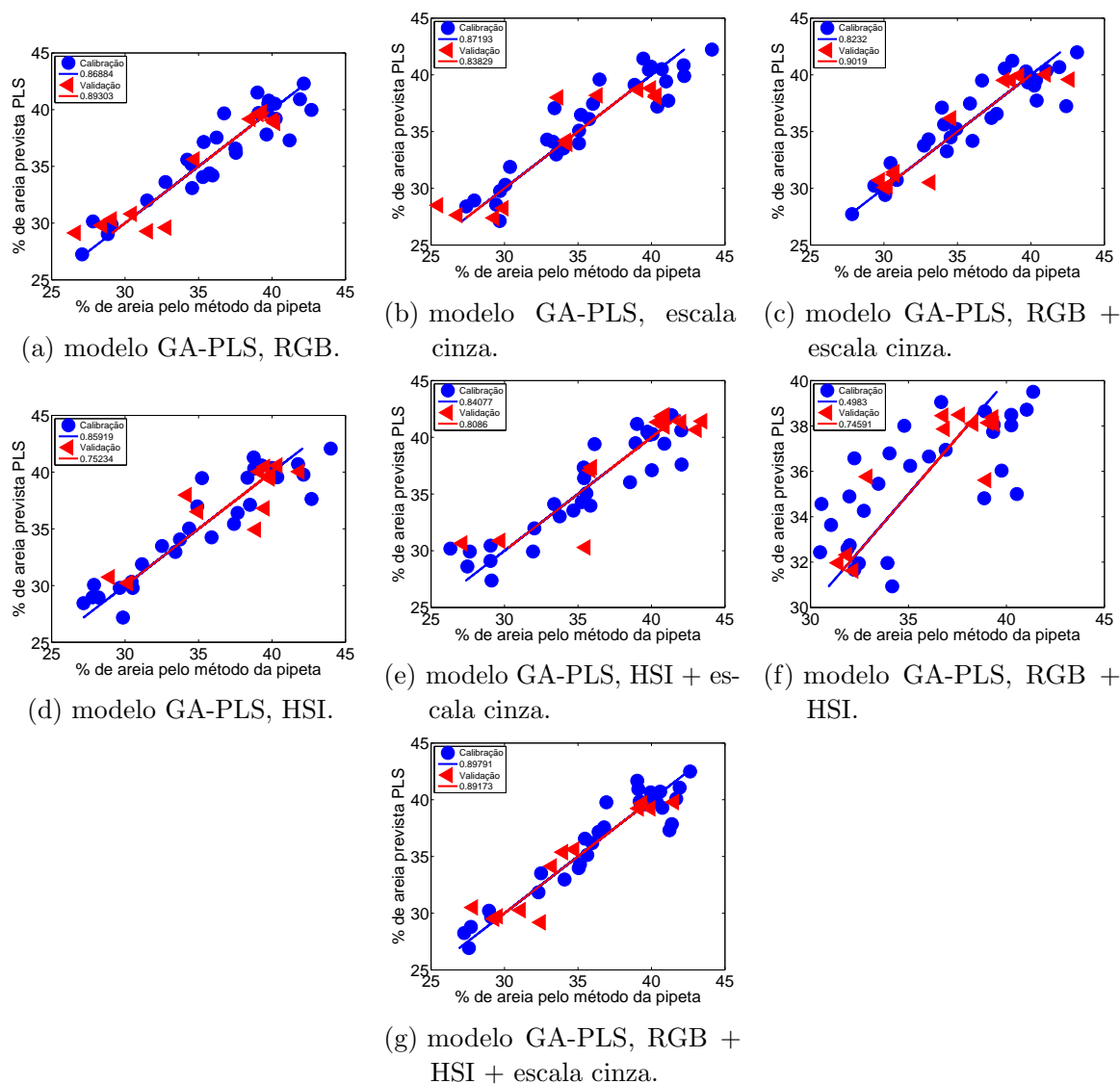


Figura 98 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 6,3x.

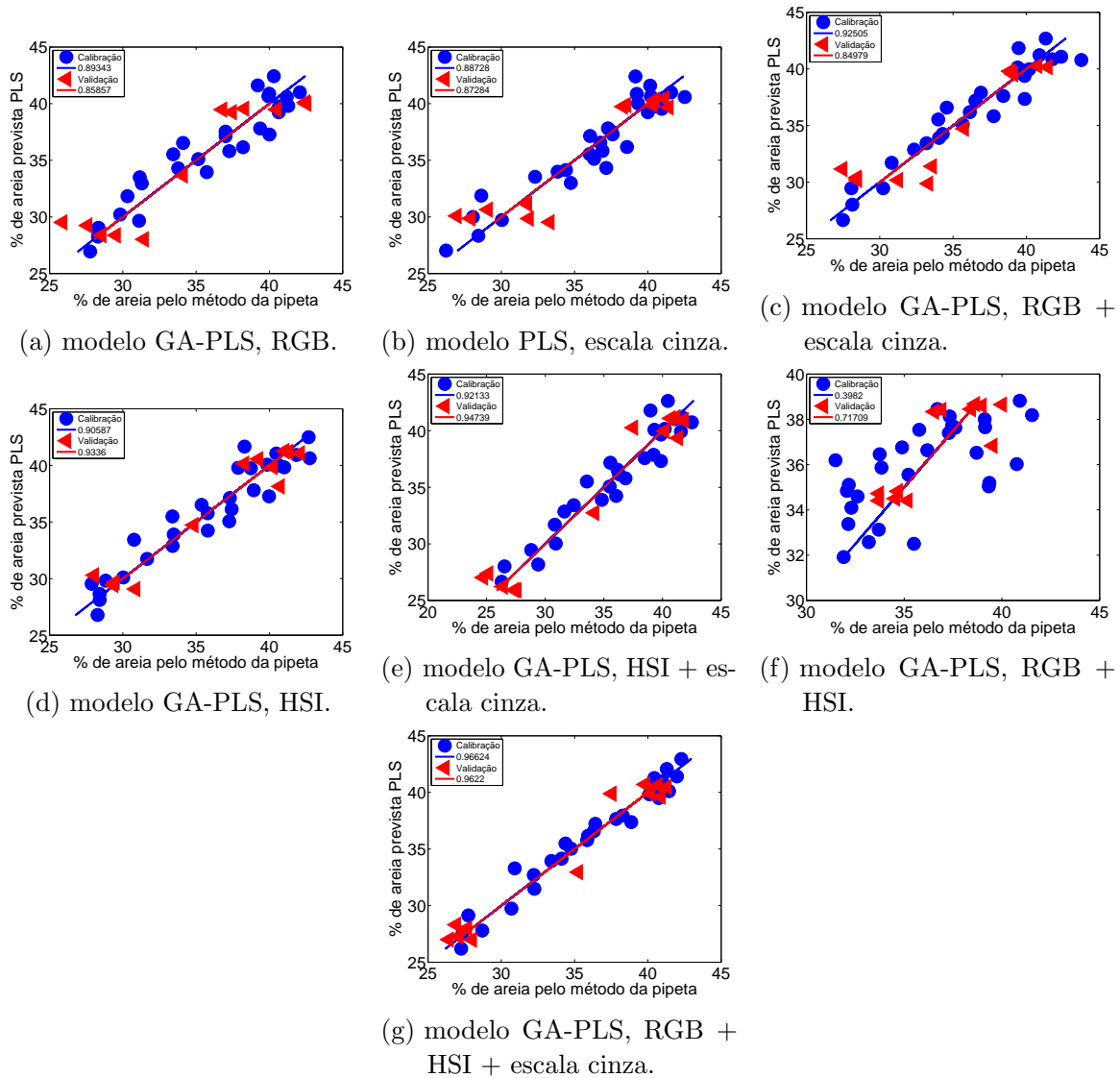


Figura 99 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas pelo algoritmo SPXY e referente ao aumento óptico de 10,0x.

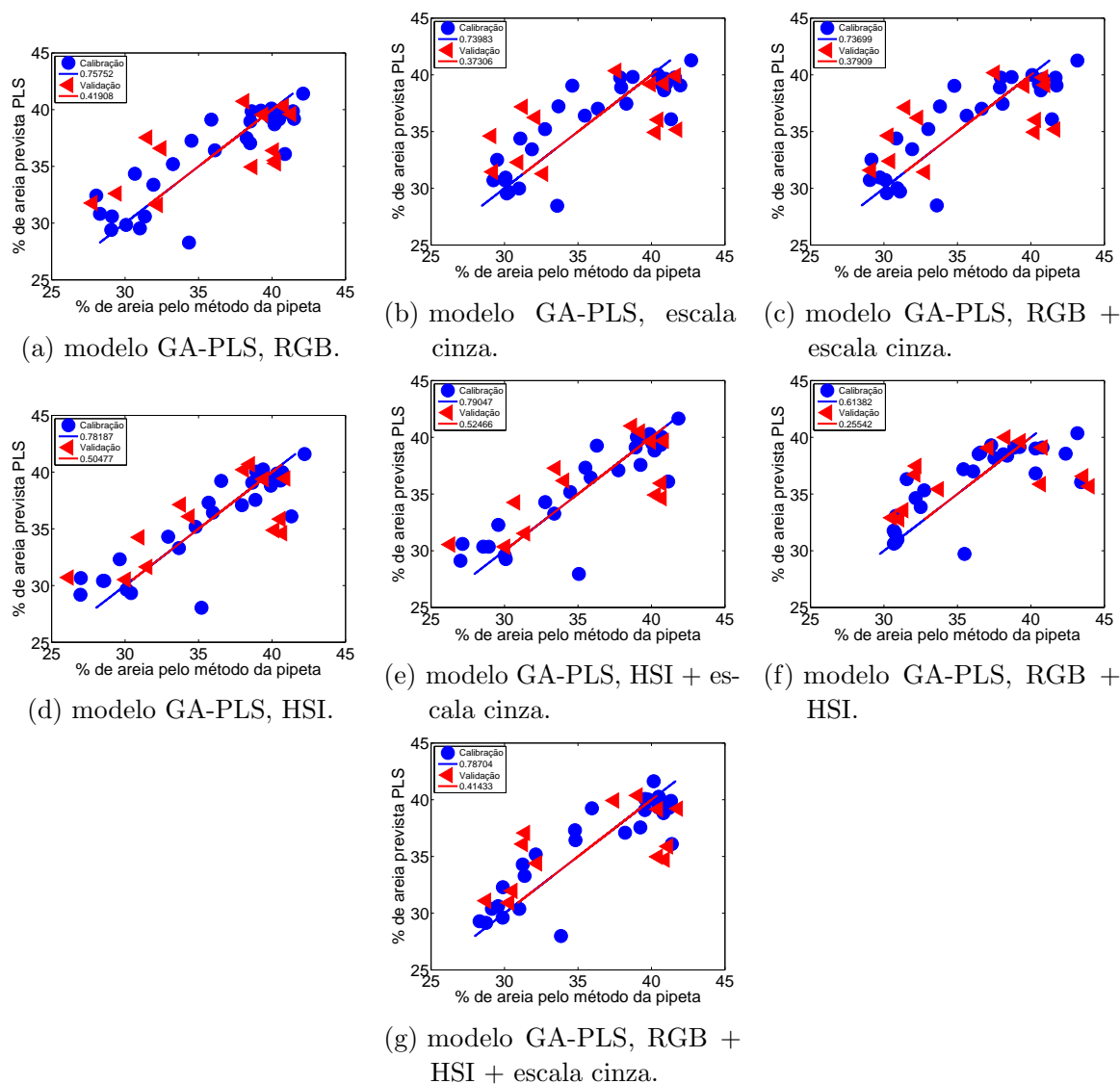


Figura 100 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 6,3x.



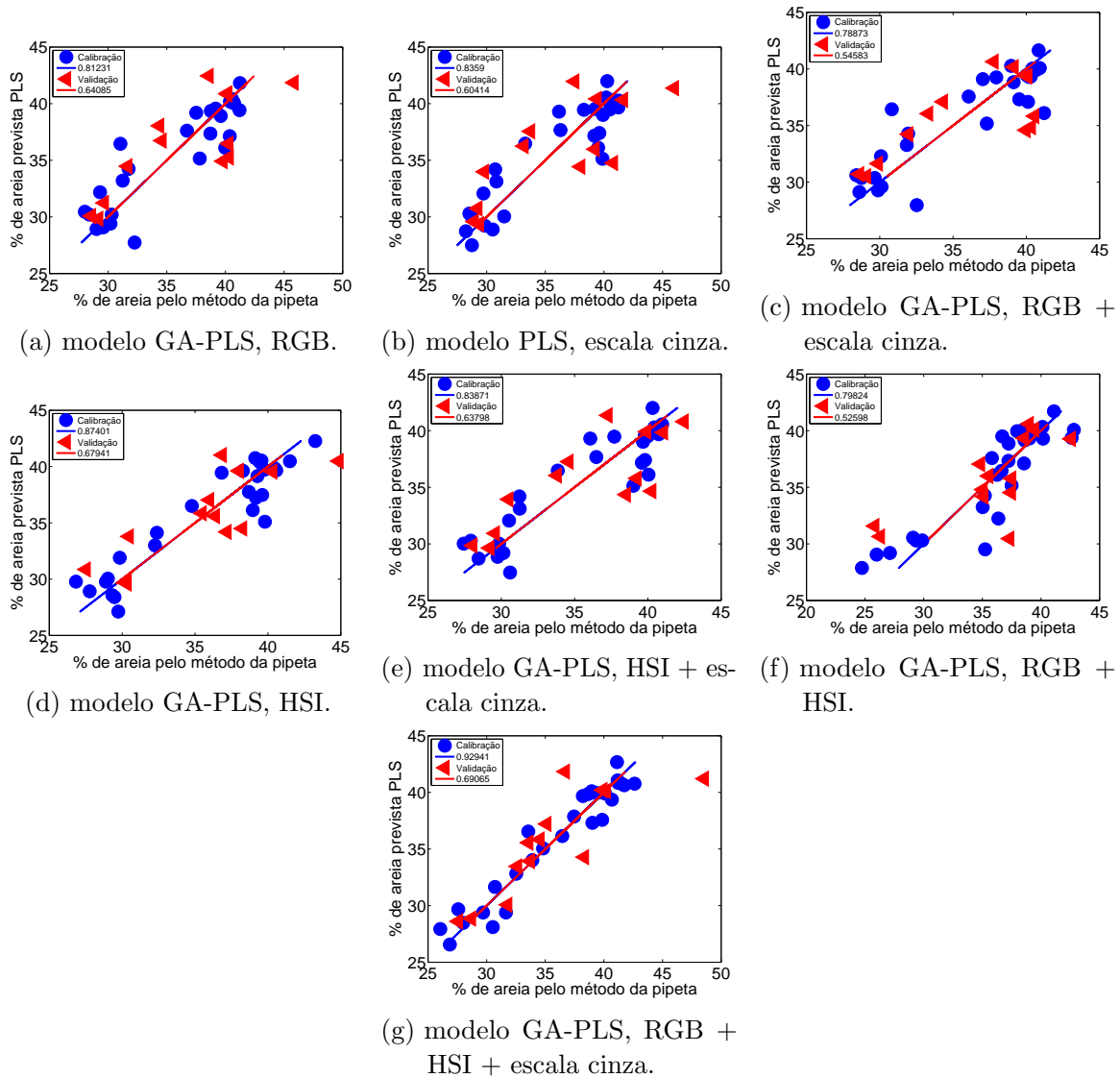


Figura 101 – Porcentagem de areia prevista pelo modelo GA-PLS versus porcentagem de areia determinada pelo método de referência (pipeta) para os conjuntos de calibração (●) e validação (◁), subdivididas por particionamento igualmente espaçado e referente ao aumento óptico de 10,0x.