

ALINE DELFINO GERMANO

**ANÁLISE EXPLORATÓRIA DE IMAGENS ÓPTICAS E DADOS SAR
POLARIMÉTRICOS NO PLANALTO SUL CATARINENSE**

Dissertação apresentada ao curso de Pós-graduação
em Engenharia Florestal do Centro de Ciências
Agroveterinárias da Universidade do Estado de Santa
Catarina como requisito parcial para a obtenção do
título de Mestre em Engenharia Florestal.

Orientador: Dr. Veraldo Liesenberg

Co-orientador: Dr. Marcos Benedito Schimalski

Lages, SC

2018

**Ficha catalográfica elaborada pelo(a) autor(a), com
auxílio do programa de geração automática da
Biblioteca Setorial do CAV/UDESC**

Germano, Aline Delfino
Análise Exploratória de Imagens Óptica e Dados SAR
Polarimétricos no Planalto Sul Catarinense / Aline
Delfino Germano. - Lages , 2018.
135 p.

Orientador: Veraldo Liesenberg
Co-orientador: Marcos Benedito Schimalski
Dissertação (Mestrado) - Universidade do Estado
de Santa Catarina, Centro de Ciências
Agroveterinárias, Programa de Pós-Graduação em
Engenharia Florestal, Lages, 2018.

1. Sensoriamento Remoto. 2. SAR. 3. ALOS/PALSAR.
4. Landsat-5/TM. 5. Classificação Supervisionada .
I. Liesenberg, Veraldo . II. Schimalski , Marcos
Benedito . , .III. Universidade do Estado de Santa
Catarina, Centro de Ciências Agroveterinárias,
Programa de Pós-Graduação em Engenharia Florestal.
IV. Título.

ALINE DELFINO GERMANO

**ANÁLISE EXPLORATÓRIA DE IMAGENS ÓPTICAS E DADOS SAR
POLARIMÉTRICOS NO PLANALTO SUL CATARINENSE**

Dissertação (Mestrado em Engenharia Florestal), Programa de Pós-graduação em Engenharia Florestal, Centro de Ciências Agroveterinárias, Universidade do Estado de Santa Catarina, Mestre em Engenharia Florestal, Área de concentração: Produção Florestal.

Banca examinadora:

Orientador/presidente:

Prof. Dr. Veraldo Liesenberg
(UDESC/Lages-SC)

Membro externo:

Prof. Dr. José Marcato Júnior
(UFMS/Campo Grande- MS)

Membro externo:

Prof. Dr. Roberto Fabris Goerl
(UFSC/Curitibanos- SC)

Membro interno:

Prof. Dr. André Felipe Hess
(UDESC/Lages-SC)

Lages, SC, 16 de fevereiro de 2018

*Aos meus pais, Alexandre e Elisabete, aos meus
irmãos Duh e Der e ao meu afilhado Luã, por todo
apoio e amor recebidos!*

Dedico...

AGRADECIMENTOS

A Deus, por me abençoar e nunca me abandonar.

A presente dissertação de mestrado não poderia chegar a bom porto sem o precioso apoio de várias pessoas.

Em primeiro lugar não poderia deixar de agradecer ao meu orientador, Professor Dr. Veraldo Liesenberg por toda a paciência, empenho e sentido prático com que sempre me orientou. Muito obrigada por me ter corrigido quando necessário.

Aos professores André Hess, Marcos Schimalski, Thiago Stepka e Rosangela Quequetto pelas contribuições inestimáveis e considerações engrandecedoras, tanto em aspectos de existência quanto na vida acadêmica.

À Universidade do Estado de Santa Catarina (UDESC) e ao Programa de Pós-graduação em Engenharia Florestal (PPGEF) pelo amparo durante o período do Mestrado.

À Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC) pelo apoio financeiro por meio de bolsa de estudos (Processo da Chamada Pública FAPESC Nº 05/2015 Recursos Humanos em CTI) e PAP/UDESC (Processo 483/2017).

Gostaria de agradecer aos meus colegas da pós, por todos momentos vividos externamente a universidade, e fazerem da batalha da vida um pouco mais humana.

Aos meus grandes amigos Sandra, João, Mari, Leo, Manu e Felipe que fizeram dos meus dias lageanos, perceber, que mesmo longe de casa, eu tenho uma família nesse cantinho. Agradeço com todo meu coração principalmente a Sandra, João, Mari, Matos, Fer e Nana pela convivência, paciência, amizade, conselhos, risadas e amor compartilhados nesses anos, e que venham muito mais.

Meus irmãos Duh e Der, que mesmo distantes (km) dão conta de alegrar minha humilde vida. Ao meu afilhado Luã que a cada dia me mostra como é bom ser feliz, e que a felicidade é a nossa alma.

E por último não menos importante, e sim, fartamente importante, aos meus pais, Alexandre e Elisabete, por serem inspiração e exemplo para seus filhos e neto, por sempre acreditar, ensinar, batalhar, educar e amar nós com toda as forças. A distância nunca diminuiu este sentimento de gratidão, respeito e amor que sinto por vocês! Amo vocês!

Agradeço a todos que de forma direta ou indireta ajudaram a tornar possível esta conquista.

MUITO OBRIGADA!

RESUMO

GERMANO, Aline Delfino. **Análise Exploratória de Imagens Ópticas e Dados SAR Polarimétricos no Planalto Sul Catarinense.** 2018. 135 p. Dissertação (Mestrado em Engenharia Florestal) – Universidade do Estado de Santa Catarina. Programa de Pós-graduação em Engenharia Florestal, Lages, 2018.

O presente estudo realizou quatro objetivos, sendo o primeiro a realização de uma breve revisão bibliográfica sobre o imageamento de sensores ópticos e de micro-ondas. O segundo, foi avaliar a performance das classificações supervisionadas oriundas dos dados ópticos em diferentes tamanhos amostrais. O terceiro, avaliou a performance de dados óptico e dados SAR, de forma isolada e combinada, para o mapeamento, e o quarto avaliou o tempo computacional de dois classificadores não-paramétricos *Support Vector Machine* (SVM), e o *Random Forest* (RF) para fins de classificação, tanto para diferentes tamanhos de amostras, quanto para diferentes números de variáveis de entrada. Foram consideradas seis classes. Cada classe foi constituída por 1000 pixels. Logo, diferentes tamanhos de amostras de treinamento e validação foram classificados, sendo elas 10% treinamento e 90% validação, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10. A segunda avaliação foi realizada utilizando os dois sensores (ópticos e SAR), e diferentes cenários foram realizados, com os produtos de decomposições (Incoerentes, Coerentes e Texturas) da cena do SAR e com a reflectância de superfície do sensor óptico. Portanto, foram geradas 13 combinações de diferentes variáveis. No geral, foram utilizados métodos de classificação supervisionados tais como o SVM, RF e Maximum Likelihood (ML). A qualidade dos mapas gerados, foram avaliadas usando o índice Kappa. As significâncias dos mapas, foram avaliadas por meio do teste Z. Já para a avaliação do tempo de execução computacional, foram anotados o tempo de término do processamento oferecido pelo aplicativo. Com os resultados obtidos, foi plausível findar que para a avaliação do tamanho de amostra de treinamento e validação, o conjunto de dados 90/10, 60/40 e 40/60 obtiveram os melhores índices Kappa, em ordem para RF, SVM e ML, respectivamente, recomendando a utilização de classificador não-paramétricos em cenas de sensores ópticos, quais são comumente utilizados algoritmos paramétricos. Já para a avaliação da integração dos sensores em diferentes cenários, os resultados mostraram que o índice Kappa máximo encontrado somente para dados oriundos do sensor SAR foi de 86,45%, já para o sensor óptico foi de 94,66% e a integração dos dois sensores atingiu 96,75%, ou seja, evidenciando que a combinação dos dados ópticos com os dados SAR, trazem melhorias para fins de classificação, recomendando a utilização de dados SAR para classificações quando houver dificuldade de se obter imagens de sensores ópticos, livre de nuvens. A avaliação do tempo computacional de ambos classificadores, é legítimo concluir que o RF apresentou menores tempos computacionais e maiores índices Kappa, ao ser comparado com o SVM, recomendando a utilização deste algoritmo de classificação. Portanto, de maneira geral, foi iminente concluir que a utilização do algoritmo RF, nos três capítulos, sobrelevou aos demais. A utilização de imagens oriundas de sensores SAR traz melhorias para classificação de uso e cobertura da terra e deve ser explorada mais intensivamente.

Palavras-chave: Sensoriamento Remoto; SAR; ALOS/PALSAR; Landsat-5; Tamanho de amostra, Tempo computacional, Classificação supervisionada.

ABSTRACT

GERMANO, Aline Delfino. **Exploratory Analysis of optical and data polarimetric SAR image in the Santa Catarina State Southern Plateau.** 2018. 135 p. Dissertation (Master in Forest Engineering – Area: Forest Engineering) – Santa Catarina State University. Forestry Engineering Graduate Program, Lages, 2018.

This research has four main goals, each one is presented as a single chapter, being the first one a brief literature review concerning the imaging process by both optical and microwave sensors. The second one considers the performance of traditional supervised classification methods by changing the size of the data input dataset. The third one evaluated the performance of the supervised classification by using both SAR and optical datasets, alone and combined, for classification purposes. Finally, the fourth one evaluated the computational time of both non-parametric Support Vector Machine (SVM), and Random Forest (RF) classification methods, thus, considering training dataset size and different number of variables as input data. A total of six land use and land cover (LULC) classes were considered. Each LULC class has 1000 pixels. Different sizes of data input datasets, what implies in both training and validation dataset, were considered, such as 10% training and 90% validation, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10. After, 13 different scenarios of data input using both optical (surface reflectance) and SAR data (coherent, incoherent, texture), alone and combined, were evaluated. As supervised classification methods we use SVM, RF and Maximum Likelihood (ML). The quality of the generated maps was evaluated by Kappa index. The significance level by Z-test. For the computational time we consider the processing time of the software. The results show that the best Kappa indices were generated by the training/validation datasets of 90/10, 60/40 and 40/60, in order to RF, SVM e ML, respectively. We do recommend the use of non-parametric classification approaches, even for optical data. The data integration shows that the use of SAR dataset alone achieves a Kappa index of 86.45%, whereas for the optical data alone 94.66%. The combined use of both SAR and optical datasets reached 96.75% what suggest that the use of SAR dataset alone is recommendable when there is an absence of cloud free optical scenes. The computational time analysis reveals that RF presented the lowest time demand and highest Kappa index when compared to SVM. RF also showed the best performance the three first chapters what suggest their use in LULC mapping procedures. The use of SAR data brought us good results for LULC classification and should be explored more intensively.

Keywords: Remote sensing; SAR; ALOS/PALSAR; Landsat-5; Sampling size; Kappa index; computational time; Supervised classification.

LISTA DE FIGURAS

Capítulo 1: FUNDAMENTAÇÃO TEÓRICA DE SENSORIAMENTO REMOTO: DEFINIÇÕES E APLICABILIDADE

Figura 1.1. Divisão do Espectro Eletromagnético (EEM) em diferentes intervalos.	35
Figura 1.2. Influências atmosféricas.....	37
Figura 1.3. Diferentes resoluções espaciais para discriminar áreas.	40
Figura 1.4. Diferenças da resolução radiométrica em uma área urbana.	40
Figura 1.5. Interação das bandas X, C e L em vegetação.....	43
Figura 1.6. Trajetória do campo elétrico referente as polarizações.....	44
Figura 1.7. Tipos de superfície e espalhamento associados: (A) lisa - Especular, (B) rugosa-Difuso, (C) lisa - Reflexão de canto.	45
Figura 1.8. Imagem sob efeito de diferentes tipos de ruídos.....	46
Figura 1.9. Plano Bidimensional de Entropia (H) e Ângulo Alpha (α).	48
Figura 1.10. Esquema de classificação do SVM.	52
Figura 1.11. Esquema estrutural do resultado de uma árvore de decisão.....	53

Capítulo 2: Impacto do Tamanho e Proporção Amostral para Fins de Classificação Supervisionada em Dados Tm/Landsat-5

Figura 2.1. Localização da área de estudo.....	73
Figura 2.2. Assinatura Espectral das classes amostradas para classificação de uso e Cobertura da terra.	76
Figura 2.3. Valores do Coeficiente Kappa e tamanho de amostras de treinamento/Validação.	77
Figura 2.4. Bloxplot- Relação índice Kappa com algoritmos de classificação.	78
Figura 2.5. Cenas Classificadas com diferentes algoritmos de classificação.	79

Capítulo 3: Avaliação da Performance de Dados SAR Polarimétricos (Banda L) e Landsat para Fins de Mapeamento no Planalto Sul Catarinense

Figura 3.1. Localização da área de estudo.....	91
Figura 3.2. Processamento realizado para processamento das cenas dos sensores utilizados..	93
Figura 3.3. Assinatura Espectral das classes amostradas para classificação de uso e Cobertura da terra referente ao sensor TM/Landsat-5.....	101
Figura 3.4. Coeficiente de Retroespalhamento em relação as polarizações	102

Figura 3.5. Plano bidimensional alfa vs. entropia proposto por Cloude e Pottier (1997) para as classes amostrais selecionadas.....	103
Figura 3.6. Decomposições de Cloude e Pottier (1997). (A) Landsat-5; (B) Angulo médio Alpha; (C) Entropia.....	104
Figura 3.7. Índice Kappa com os diferentes cenários de produtos SAR e óptico.....	105
Figura 3.8. Boxplot- Relação índice Kappa com classificadores.	109

Capítulo 4: Análise do Tempo de Execução de Dois Algoritmos Não-Paramétricos para Classificação de Uso e Ocupação da Terra Utilizando os Sensores Landsat-5/TM e ALOS/PALSAR

Figura 4.1. Localização da área de estudo.	120
Figura 4.2. Processamento realizado para processamento das cenas dos sensores utilizados.	121
Figura 4.3.Tamanho de amostra de treinamento e validação em relação ao tempo computacional e índice Kappa.....	127
Figura 4.4.Correlação de Spearman para Níveis de Tratamento	128
Figura 4.5.Cenários em relação ao tempo computacional e índice Kappa.	129
Figura 4.6.Correlação de Spearman em relação aos cenários de combinações.	130

LISTA DE TABELAS

Capítulo 1: Fundamentação Teórica de Sensoriamento Remoto: Definições e Aplicabilidade

Tabela 1.1. Comportamento Espectral da Vegetação em diferentes regiões espetrais.....	42
Tabela 1.2. Descrição de Bandas Utilizadas por Sensores SAR.	44
Tabela 1.3. Mecanismos de Espalhamento das Zonas Bidimensional.	49
Tabela 1.4. Representação da Matriz de Confusão.	54
Tabela 1.5. Categorização de uma classificação a partir do Índice Kappa.	56

Capítulo 2: Avaliação de algoritmos paramétricos e não-paramétricos para classificação de uso e cobertura da terra no Planalto Sul Catarinense

Tabela 2.1. Características espetrais e espacial do satélite Landsat 5.....	74
Tabela 2.2. Descrição das Classes de Uso e Cobertura da terra.	75
Tabela 2.3. Valores de Índice Kappa obtidos para os diferentes algoritmos de classificação. 78	
Tabela 2.4. Representatividade das diferentes classes de uso e cobertura da terra, para a área de estudo, considerando três distintos classificadores e conjunto amostral.	79
Tabela 2.5.Valores do Teste Z para diferentes conjuntos amostrais avaliados para os classificadores (a) Random Forest, (b) Support Vector Machine e (c) Maximum Likelihood.80	
Tabela 2.6 Valores do Teste Z comparando os classificadores com os respectivos tamanhos de amostras.	81
Tabela 2.7. Valores do Teste Z comparando os classificadores com os respectivos tamanhos de amostras de treinamento.	81
Tabela 2.8. Valores do Teste Z comparando os classificadores com os respectivos tamanhos de amostras de treinamento.	82

Capítulo 3: Avaliação da performance de dados SAR polarimétricos e ópticos para fins de mapeamento no planalto sul catarinense

Tabela 3.1. Resoluções espetrais e espacial do Satélite Landsat -5.....	92
Tabela 3.2. Atributos Incoerentes.....	95
Tabela 3.3. Atributos Coerentes	97

Tabela 3.4. Atributos Texturais calculados a partir dos coeficientes de retroespalhamento dos dados SAR.....	98
Tabela 3.5. Descrição das seis classes de uso e cobertura da terra.	99
Tabela 3.6. Combinação dos cenários para classificação de uso e cobertura da terra.	100
Tabela 3.7. Valores do Teste Z comparando os classificadores com seus respectivos cenários.	107
Tabela 3.8. Valores do Teste Z, comparando as combinações para o classificador SVM.	107
Tabela 3.9. Valores do Teste Z, comparando as combinações para o classificador RF.	108

Capítulo 4: Análise do Tempo de Execução de Dois Algoritimos Não-Paramétricos para Classificação de Uso e Cobertura da terra no Planalto Sul Catarinense

Tabela 4.1. Resoluções espectrais e espacial do Satélite Landsat -5.	119
Tabela 4.2. Atributos Incoerentes	122
Tabela 4.3. Atributos Coerentes.....	123
Tabela 4.4. Atributos Texturas.....	124
Tabela 4.5. Descrição das Classes de Uso e cobertura da terra.	124
Tabela 4.6. Combinação dos cenários para classificação de uso e ocupação do solo.	125

LISTA DE ABREVIATURAS E SIGLAS

ANOVA	Analise de variânciа
DOS	Dark object subtraction
EEM	Espectro eletromagnético
e.g.	Exempli gratia (por exemplo)
eg	Exatidão geral
eu	Exatidão do usuário
ep	Exatidão do produtor
FLAASH	<i>Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes</i>
GPS	Global Positioning System (Sistema de posicionamento global)
i.e.	id est (isto é)
IV	Índice de Veregação
K	Índice Kappa
L	Radiância
ML	Maximum likelihood
NASA	Agência Espacial Americana
ND	Niveis digitais
MODIS	<i>MOderate Resolution Imaging Spectroradiometer</i>
OOB	out-of-bag
RADAR	Detecção e telemetria por Rádio (Radio Detection and Ranging)
REM	Radiação Eletromagnética
RF	Random Forest
S	Cenários
SAR	Synthetic Aperture Radar
SC	Santa Catarina
SERE	Sensoriamento Remoto
SIGSC	Sistema de informação geográfica do Estado de Santa Catarina
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
TOA	Refectância do topo da atmosfera
UDESC	Universidade do Estado de Santa Catarina
USGS	<i>U.S. Geological Survey</i>

SUMÁRIO

INTRODUÇÃO GERAL	23
HIPÓTESE	25
OBJETIVOS	26
ESTRUTURA DO DOCUMENTO	27
REFERÊNCIAS BIBLIOGRÁFICAS	28

CAPÍTULO 1

Fundamentação Teórica de Sensoriamento Remoto: Definições e Aplicabilidade	31
Resumo	31
Abstract.....	32
1.1. INTRODUÇÃO.....	33
1.2. DESENVOLVIMENTO.....	34
1.2.1.Sensoriamento Remoto	34
1.2.2.Imageamento Óptico.....	35
1.2.3.Imageamento Micro-ondas.....	43
1.2.4.Métodos De Classificações para Geração de Mapas Temáticos	50
1.2.5.Aplicação do Sensoriamento Remoto na Ciência Florestal.....	56
1.3. CONCLUSÕES	58
REFERÊNCIAS BIBLIOGRÁFICAS	59

CAPÍTULO 2

Avaliação de algoritmos paramétricos e não-paramétricos para classificação de uso e cobertura da terra no Planalto Sul Catarinense.....	69
Resumo	69
Abstract.....	70
2.1. INTRODUÇÃO.....	71
2.2. MATERIAL E MÉTODOS.....	72
2.2.1.Descrição da área de estudo	72
2.2.2. Pré-processamento de imagens.....	74
2.2.3. Classificações dos dados remotamente situados	74

2.3. RESULTADOS E DISCUSSÃO.....	76
2.4. CONCLUSÕES	82
REFERÊNCIAS BIBLIOGRÁFICAS.....	84

CAPÍTULO 3

Avaliação da performance de dados SAR polarimétricos e ópticos para fins de mapeamento no planalto sul catarinense	87
Resumo.....	88
Abstract	88
3.1. INTRODUÇÃO	99
3.2. MATERIAL E MÉTODOS	90
3.2.1. Descrição da área de estudo	90
3.2.2. Sensores Orbitais.....	92
3.2.3. Pré-processamento das cenas.....	93
3.2.4. Descrição das classes de uso e cobertura da terra.....	98
3.3. RESULTADOS E DISCUSSÃO	100
3.4. CONCLUSÕES	110
REFERÊNCIAS BIBLIOGRÁFICAS.....	111

CAPÍTULO 4

Análise do Tempo de Execução de Dois Algoritmos Não-Paramétricos para Classificação de Uso e Cobertura da terra no Planalto Sul Catarinense.....	115
Resumo.....	115
Abstract	116
4.1. INTRODUÇÃO	117
4.2. MATERIAL E MÉTODOS	118
4.2.1. Caracterização da área de estudo	118
4.2.2. Sensores Orbitais.....	118
4.2.3. Pré-processamento e Processamento das cenas.....	119
4.2.4. Descrição das classes de uso e cobertura da terra.....	124
4.2.5. Classificação dos dados remotamente situados	125

4.3. RESULTADOS E DISCUSSÃO.....	127
4.3.1. Tempo computacional para diferentes tamanhos de amostra de treinamento e validação	127
4.3.2. Tempo Computacional para diferentes cenários	128
4.4 CONCLUSÕES	130
REFERÊNCIAS BIBLIOGRÁFICAS	131
CONSIDERAÇÕES FINAIS	135

INTRODUÇÃO GERAL

O uso crescente de sensores remotamente situados, na forma de imagens, tem sido aplicado para fins de mapeamento de uso e cobertura da terra, e apresentam uma alternativa menos custosa que os trabalhos tradicionais a campo (JANOTH et al., 2007; BARBIER et al. 2012). De tal modo, os mapas tornam-se ferramentas importantes para monitorar, entender e prever os resultados da interação natureza-homem, sendo ela em escala local, regional e ou global (CLARK et al., 2010). Desta forma, dados oriundos do sensoriamento remoto envolvem diversas aplicações em diferentes áreas de conhecimento, e têm se destacado na tomada de decisões para gestão de recursos naturais (FLORENZANO, 2002).

A utilização de técnicas de mapeamentos aplicando sensores orbitais ópticos, como o da série LANDSAT, consiste em uma alternativa viável e vêm-se utilizando em larga escala para monitoramentos florestais (ROSOT, 2001). No entanto, os sensores ópticos nem sempre apresentam imagens em boa qualidade, devido à presença de nuvens e aerossóis, o que pode deixar áreas da superfície terrestre sem informações (SOARES et al., 1997). Prontamente, os sensores micro-ondas orbitais vêm preencher este espaço, já que o processo de aquisição de cenas independe das condições climáticas (PALME, 1997).

Sensores micro-ondas, por serem ativos, apresentam vantagem de operar em todas as condições de iluminação (PONZONI, SHIMABUKURO e KUPLICH, 2012). Suas ondas eletromagnéticas estão na faixa espectral das micro-ondas, se tornando uma ferramenta promissora para o recobrimento de regiões sujeitas à constante presença de nuvens, brumas e chuvas (SILVA, 2010). Kasischke, Melack e Dobson (1997) afirmaram que por falta de experiência de interpretação dos dados de micro-ondas, estes não foram amplamente estudados como os dados de imagens ópticas. Após duas décadas, muitos estudos estão sendo realizados, o que tem apontado um crescimento acelerado no desenvolvimento de produtos Synthetic Aperture Radar (SAR) (PONZONI, SHIMABUKURO e KUPLICH, 2012).

A demanda por produtos dos sensores SAR teve um aumento no setor de pesquisa, entretanto, ainda são relativamente baixos ao ser comparados com os sensores ópticos. Um exemplo a ser verificado é no Simpósio Brasileiro de Sensoriamento Remoto (SBSR) realizado no ano de 2018, onde, ao se pesquisar sobre mapeamento de uso e cobertura da terra, acha-se 64 publicações, entretanto, apenas 4 dessas foram realizadas com sensores micro-ondas. A geociência no geral é a quinta área mais pesquisada na região sul do Brasil, sendo que no estado de Santa Catarina é a quarta (PLATAFORMA LATTES, 2018).

Estudos de classificação de uso e cobertura da terra estão sendo aplicados com a utilização de sensores SAR, como a pesquisa de Muthukumarasamy, Shanmugam e Kolanuvada (2017), que utilizaram o sensor na banda L para classificações, ou como Mishra et al. (2017) que empregaram sensor na banda C. Além desta utilização, os sensores SAR oferecem parâmetros estrutural dos objetos da superfície terrestre. Assim, Viet Nguyen et al. (2016b) utilizaram a banda L para caracterizar as estruturas de florestas tropicais. Martins et al. (2016) pesquisaram a sensibilidade do sensor ALOS/PALSAR na banda L para detecção de áreas degradadas com fogo no norte da Amazônia.

É notório que diversos estudos estão sendo realizados a fim de identificação, classificação do uso e cobertura da terra, tanto com sensores SAR e Ópticos; entretanto, ainda é vago estudo com aplicação da combinação desses sensores, tanto para classificações quanto para extração de propriedade biofísicas.

HIPÓTESE

Estudos como de Muthukumarasamy, Shanmugam e Kolanuvada (2017) que avaliaram decomposições de alvos de sensores micro-ondas separadamente para classificações de uso e cobertura da terra, obtiveram resultados significativos. Mishra et al. (2017) investigaram a utilização de dados micro-ondas na banda C, empregando a textura para verificação de melhorias nas classificações. Assim, estudos que visam a melhoria de classificações de uso e cobertura da terra, utilizando sensores micro-ondas e seus produtos, são de suma importância. Este argumento é endossado para regiões aonde a cobertura de nuvens é frequente, especialmente na região que compreende o planalto sul catarinense, objeto deste estudo.

A hipótese que norteia esta dissertação de mestrado é a de que o uso de dados SAR polarimétricos, na banda L, integrado com dados de um sensor óptico pode melhorar as classificações de mapeamentos de uso e cobertura da terra. Para testar a validade da hipótese, serão analisados métodos desenvolvidos especificamente para dados SAR polarimétricos (decomposição dos alvos) com a integração dos dados ópticos, seguindo abordagens que buscarão alcançar os objetivos listados a seguir.

OBJETIVOS

Fundamentado no exposto, o objetivo geral desta dissertação de mestrado foi avaliar o uso isolado e combinado de dados de diferentes sensores imageadores (óptico e de micro-ondas) com a finalidade mapear o uso e a cobertura da terra de ambientes rurais no planalto sul catarinense. Como objetivos específicos podem-se listar:

- Avaliação de diferentes classificadores *Random Forest* (RF), *Support Vector Machine* (SVM) e *Maximum Likelihood* (ML)) para diferentes tamanhos de amostras de treinamento e validação (i.e., 10% treinamento e 90% validação, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10) utilizando imagens do sensor TM/Landsat-5.
- Aferir a performance dos dados ópticos (Landsat-5/TM), bem como de dados SAR polarimétricos (ALOS/PALSAR), de forma individual e em diferentes combinações de dados para dois classificadores não-paramétricos.
- Investigar o tempo computacional de dois classificadores não-paramétricos para fins de classificação de uso e cobertura da terra para diferentes números de variáveis de entrada e combinações, utilizando dados dos sensores Landsat-5/TM e ALOS/PALSAR.

ESTRUTURA DO DOCUMENTO

Este documento está organizado em quatro capítulos, com os seguintes conteúdos:

Capítulo 1 – apresentação da fundamentação teórica acerca das técnicas empregadas nesta pesquisa. Parte-se inicialmente de uma apresentação simplificada do funcionamento de sensores ópticos e micro-ondas. Posteriormente, incluindo informações sobre técnicas de classificação de uso e ocupação da terra e ferramentas de avaliação de mapas temáticos. Finalmente, com explanação sobre a utilização do sensoriamento remoto aplicado a setores florestais;

Capítulo 2 – apresentação de estudo sobre o impacto do tamanho e proporção de amostras de treinamento e validação para fins de classificação do uso e cobertura da terra aplicado a dados do sensor TM/LANDSAT-5;

Capítulo 3- explanação da performance da integração de dados oriundo de sensores operando na região de micro-ondas e óptico, para fins de mapeamento de ambientes rurais no planalto sul catarinense, aplicando dois algoritmos não-paramétricos; e

Capítulo 4 – esclarecimento sobre a análise do tempo de execução de dois algoritmos não-paramétricos para fins de classificação de uso e ocupação da terra utilizando os sensores Landsat-5/TM e ALOS/PALSAR.

REFERÊNCIAS BIBLIOGRÁFICAS

BARBIER, N.; COUTERON, P.; GASTELLY-ETCHEGORRY, J. P.; PROISY, C. Linking canopy images to forest structural parameters: Potential of a modeling framework. **Annals of Forest Science**, v. 69, n. 2, p. 305–311, 2012.

CLARK, M. L.; AIDE, T. M.; GRAU, H. R.; RINER, G. A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America. **Remote Sensing of Environment**, v. 114, n. 11, p. 2816–2832, 2010.

DUTRA, L. V.; MURA, J. C.; FREITAS, C. D. C.; SANTOS, J. R. DOS; ELMIRO, M. T. **Processamento de Imagens de Radar de Abertura Sintética – Princípios e Aplicações**. Anais do IV Workshop em Tratamento de Imagens. Minas Gerais. p.4–13, 2003.

FLORENZANO, T. G. **Imagens de Satélite para Estudos Ambientais**. Oficina de Textos, p. 1–97, 2002.

JANOTH, J. et al. **Procedimentos Baseados em Segmentação para a Análise de Mudanças e Classificação Florestais com Dados de Satélite de Alta Resolução**. cap. 9. In: BLASCHKE, T.; KUX, H. Sensoriamento Remoto e SIG Avançados. 2. ed.- São Paulo: Oficina de textos. 2007.

KASISCHKE, E. S.; MELACK, J. M.; DOBSON, M. C. The Use of Imaging Radars for Ecological Applications A Review. **Remote Sensing of Environment**, v. 4257, n. 96, p. 141–156, ,1997.

MISHRA, V. N.; PRASAD, R.; KUMAR, P.; GUPTA, D. K.; SRIVASTAVA, P. K. Dual-polarimetric C-band SAR data for land use/land cover classification by incorporating textural information. **Environmental Earth Sciences**, v. 76, n. 1, p. 1–16, 2017.

MUTHUKUMARASAMY, I.; SHANMUGAM, R. S.; KOLANUVADA, S. R. SAR polarimetric decomposition with ALOS PALSAR-1 for agricultural land and other land use/cover classification: case study in Rajasthan, India. **Environmental Earth Sciences**, v. 76, n. 13, p. 1–13, 2017.

PALME, U. **Radares imageadores em sensoriamento remoto**. Anuário Fator GIS 1997: o guia de referência do geoprocessamento, Curitiba, v.1, p.c28-c34. 1997

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento Remoto da Vegetação**. 2º ed. São Paulo: Oficina De Texto, 2012.

ROSOT, N. C. . **Integração de Imagens de Sensores de Microondas e Ópticos para Fins de Mapeamento e Classificação de Reflorestamentos no Sul do Brasil**, 2001 p.207. Programa de Pós- Graduação em Engenharia de Produção (Universidade Federal de Santa Catarina), 2001.

SILVA, A. DE Q. DA. **Avaliação de Imagens SAR no Mapeamento de Crostas Lateríticas do Depósito N1, Província Mineral de Carajás, Pará**. (Tese de Doutorado do Curso de Pós-Graduação em Sensoriamento Remoto), nstituto Nacional de Pesquisas

Espaciais. São José dos Campos, p.233, 2010.

SOARES, S. M.; SIQUEIRA SANT'ANNA, S. J.; KUNTSCHIK, G. Integração de dados multiespectrais landsat-5 TM com imagens SAR JERS-1, para fins cartográficos. **European Space Agency, (Special Publication)**, n. 407, p. 115–123, 1997.

CAPÍTULO 1

Fundamentação Teórica de Sensoriamento Remoto: Definições e Aplicabilidade

RESUMO

A utilização de imagens coletadas por sensores remotos ópticos é bastante difundida no meio acadêmico e no mercado de trabalho, todavia, a utilização de dados coletados por sensores operando na região micro-ondas, também conhecidos como RADAR, ainda é incipiente. Na ciência florestal, trabalhos empregando dados de micro-ondas têm apresentado resultados bastante promissores, os quais normalmente, ajudam na redução de mão de obra e consequentemente nos custos de projetos. A integração de sensores ópticos e de micro-ondas vêm sendo aplicados para melhorar a acurácia de classificação bem como na melhoria da recuperação de parâmetros biofísicos. O objetivo geral deste capítulo foi realizar uma breve revisão bibliográfica sobre o imageamento de sensores ópticos e de micro-ondas para fins de classificação de uso e cobertura da terra conduzidos em ambientes florestais complexos.

Palavras-Chave: Sensoriamento Remoto, Imageamento Óptico, Imageamento por Micro-ondas.

CHAPTER 1

Theoretical Rationale of Remote Sensing: Definitions and Applicability

ABSTRACT

The use of optical images is well known in the academy and market. However, the use of microwave data, i.e. Synthetic Aperture Radar (SAR) is still incipient. In Forestry Science, several studies have been applying microwave data showing very promising perspectives that may result in reducting time consuming and operational cost. The combined use of both optical and SAR sensors has been employed to enhance classification accuracy as well as in improving the retrieval of biophysical properties. The main goal of this chapter is to present a brief literature review concerning the combined use of both optical and SAR systems for classification purposes conducted in complex forest environments worldwide.

Keywords: Remote Sensing; Optical Image Imaging; Microwave Imaging.

1.1. INTRODUÇÃO

A utilização do sensoriamento remoto para aplicações em ambientes florestais tem se desenvolvido nos últimos anos. Inicialmente, foi aplicado para mapeamento do uso e cobertura da terra de extensas regiões e em locais de difícil acesso (JANOTH et al., 2007). Posteriormente, com o refinamento da resolução espacial, os dados começaram a ser usados para delineamento de talhões florestais (ARGÜELLO et al., 2010), identificação de falhas nas linhas de plantio (ALVES et al., 2015; RAMESH et al., 2016), estimativas de áreas agrícolas plantadas (LUIZ e EPIPHANIO, 2001; ALVARES et al., 2015) entre outras aplicações. No entanto, existem diferentes sensores, os quais apresentam distintas características, qualificando ou quantificando os mesmos quanto a sua aplicação (JENSEN, 2009).

A principal aplicação até então, de sensores remotamente situados, está na confecção de mapas temáticos, onde determinadas classes de uso da terra são rotuladas em áreas homogêneas (FIGUEIREDO, 2005). Para a confecção de mapas temáticos, é utilizado, em geral, imagens advindas de sensores remotos ópticos, justamente por serem difundidos e por apresentarem um maior número de sensores em órbita, e, por consequência, fornecerem imagens, o qual o usuário consegue visualizar de forma análoga ao sistema visual humano, os objetos na superfície terrestre. Todavia, imagens coletadas por sensores remotos ópticos apresentam como principal limitação, a obtenção de imagens, livres de nuvens e aerossóis. A utilização de sensores operando na região do micro-ondas surge como uma alternativa na geração de informações, onde os sensores remotos ópticos não conseguem obter informações dos alvos na superfície terrestre.

Sensores operando na região de micro-ondas foram desenvolvidos inicialmente para uso militar e posteriormente “olvidados” para aplicações usuais (MENESES e ALMEIDA, 2012). Atualmente, estes dados podem servir de informações complementares, seja em aplicações envolvendo classificações (CHEN et al., 2016, SANLI et al., 2016; GOU, LI e YANG, 2016), seja na recuperação de parâmetros biofísicos (GAO et al., 2016; VIET NGUYEN et al., 2016b). Desta forma, os sensores micro-ondas vem sendo utilizado como alternativas que visam melhorar o uso do sensoriamento remoto para diversas aplicações.

O objetivo geral deste capítulo foi realizar uma breve revisão bibliográfica sobre o imageamento de sensores ópticos e de micro-ondas, de forma isolada e combinada, fins de classificação e recuperação de parâmetros biofísicos conduzidos em ambientes florestais complexos.

1.2. DESENVOLVIMENTO

1.2.1. Sensoriamento Remoto

Dados e imagens de dados remotamente situados têm sido utilizados em diferentes áreas do conhecimento. Estas informações, em geral, na forma de mapas, têm se destacado na tomada de decisões nas ciências agrárias e na gestão de recursos naturais (FLORENZANO, 2002). De acordo com Jensen (2009), a utilização do sensoriamento remoto, na forma de imagens, permitiu que diferentes variáveis ambientais pudessem ser investigadas (e.g., solo, água, vegetação, etc.).

Diversos estudos e aplicações nas ciências agrárias vêm sendo adotadas por meio do uso do sensoriamento remoto. Como exemplo, pode ser citado os sensores ópticos, que fornecem dados em diferentes intervalos espectrais na forma de imagens que permitem o cômputo de índices de vegetação. Estes índices de vegetação podem ser usados para fornecer informações a respeito do estágio de desenvolvimento da cultura do feijão (EPIPHANIO et al., 1996), na estimativa de parâmetros dendrométricos para reflorestamentos (OLIVEIRA, 2012), monitoramento da produção de cana-de-açúcar (PICOLI, 2012), mapeamento de biomassa florestal (VIET NGUYEN et al., 2016b), classificação de cobertura da terra com sensores ativos (MISHRA et al., 2017), entre outros.

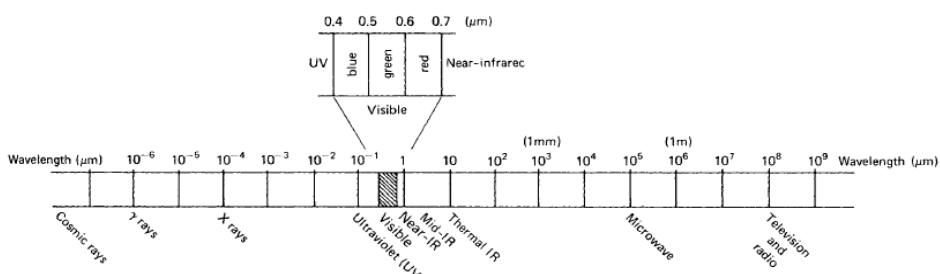
Segundo Novo (2010), o sensoriamento remoto pode ser definido como sendo a utilização de aeronaves, espaçonaves e modernos sensores, com o objetivo de estudar o ambiente terrestre através das interações entre a Radiação Eletromagnética (REM) e os objetos existentes na superfície terrestre. O emprego de sensores a bordo de satélites (Sensores Remotos Orbitais) possibilitou importantes avanços no mapeamento e monitoramento dos ecossistemas terrestres. A obtenção de dados, em diferentes faixas do espectro eletromagnético, e a interação dos dados na biosfera abriram uma nova fase de conhecimento (ABER e MELLILO, 2001).

Dependendo do estudo a ser realizado, deve-se escolher o (s) sensor (es) remoto (s) que apresenta (m) características desejáveis e viáveis para obtenção da (s) variável (is) desejada (s). Assim, os sensores podem ser classificados de acordo com a região do Espectro Eletromagnético (EEM) em que operam (e.g., Sensores ópticos e micro-ondas) (FREITAS et al., 2007). No entanto, é importante especificar que os sensores ópticos são os que registram a REM entre o intervalo do espectro eletromagnético que compreende de 400 nm a 24000 nm. Por outro lado, os sensores operando na região dos micro-ondas atuam no intervalo de 1 cm a 1m (JENSEN, 2009).

1.2.2. Imageamento Óptico

Os sensores ópticos podem oferecer informações importantes dos alvos, principalmente, sobre as suas características biofísicas, e sobretudo para estudos sobre detecção de mudanças (JENSEN, 2009). A Figura 1.1 mostra o espectro eletromagnético. Pode ser observado, que os sensores ópticos estão limitados pela região espectral óptica, ou seja, do visível ao infravermelho. As divisões mostradas na Figura 1.1 são as propostas por Lillesand e Kiefer (2008).

Figura 1.1. Divisão do Espectro Eletromagnético (EEM) em diferentes intervalos espetrais.



Fonte: Lillesand e Kiefer (2008).

Para qualquer estudo envolvendo dados ou imagens de sensoriamento remoto, conhecer a energia radiante (a radiometria) dos objetos da superfície terrestre é de suma importância. De acordo com Figueiredo (2005), a radiação eletromagnética (REM) conduz informações intrínsecas dos objetos. Logo, as principais medidas radiométricas referem-se a irradiância, radiância, absorância, reflectância e transmitância (NOVO, 2010).

As imagens coletadas por sensores remotamente situados são armazenadas em formato digital (níveis digitais – ND). Estas podem apresentar deteriorações radiométricas devido aos diferentes erros que podem ocorrer na calibração dos sensores (i.e., erros na transmissão de dados, influências atmosféricas e distorções geométricas) bem como na depreciação do aparelho eletrônico distorcendo a relação sinal/ruído. Toutin (2004) afirmou que as fontes de distorção podem ser agrupadas em duas grandes categorias: os observados (atmosfera e Terra) ou referente aos sistemas de aquisição (plataforma, sensor e outros instrumento de medição).

Portanto, técnicas de pré-processamento, seja de dados quanto de imagens, têm a finalidade de minimizar estes erros, afim de não comprometer o resultado final do processamento. Figueiredo (2005) resumiu em três principais tipos de pré-processamento que são utilizados (i.e., Geométrica, Radiométrica e Atmosférica). Desta forma, afim de poder realizar o comportamento espectral dos alvos analisados em imagens digitais, devem ser realizadas diferentes etapas de pré-processamento para a obtenção da variável reflectância de

superfície. Os procedimentos rotineiros para a conversão de dados brutos, ou seja, em ND para reflectância de superfície serão descritos no próximo seguimento.

1.2.2.1. Pré-Processamento de Imagens

A radiância (L) é a principal medida radiométrica coletada por um sensor remotamente situado. Para a sua obtenção, faz-se necessário uma boa calibração radiométrica dos detectores pré- e pós-lançamento (PONZONI et al., 2015). De tal modo, a determinação da radiância é fundamental para a obtenção da reflectância de superfície. Para se obter valores de reflectância de superfície, faz-se necessário a conversão de ND para L e posteriormente para reflectância de superfície. De acordo com Ponzoni, Shimabukuro e Kuplich (2012), a conversão tem como objetivo, permitir usufruir das características espectrais de objetos, bem como a elaboração de operações aritméticas com diferentes bandas ou sensores. A conversão de ND para radiância pode ser realizada por meio de uma interpolação linear como demonstrado na Equação 1 (Eq.1).

$$L\lambda = \left(\frac{L_{max\lambda} - L_{min\lambda}}{ND_{max} - ND_{min}} \right) * (ND - ND_{min}) + L_{min\lambda} \quad (\text{Eq.1})$$

Onde: $L\lambda$: Radiância espectral aparente ($\text{W}/(\text{m}^2 \cdot \text{sr} \cdot \mu\text{m})$); $L_{max\lambda}$: a radiância espectral máxima que o sensor é capaz de registrar ($\text{W}/(\text{m}^2 \cdot \text{sr} \cdot \mu\text{m})$); $L_{min\lambda}$: é a radiância espectral mínima que o sensor é capaz de registrar ($\text{W}/(\text{m}^2 \cdot \text{sr} \cdot \mu\text{m})$); ND_{max} : é o número digital correspondente a $L_{max\lambda}$; e ND_{min} : é o número digital correspondente a $L_{min\lambda}$ do faixa de radiação espectral pós-calibração.

A partir da radiância, pode-se realizar a conversão para reflectância aparente e/ou reflectância de superfície. Este procedimento facilita a comparação entre diferentes bandas e cenas. Entretanto, o primeiro, ou seja, a reflectância aparente, ainda possui contribuição da atmosfera. A conversão de L para reflectância aparente busca remover as diferenças causadas pela geometria solar e porções de energia solar. Por meio da Equação 2 (Eq.2), descrita por Chander, Markham e Helder (2009) obtém-se a Reflectância Aparente (i.e., trata-se da reflectância calculada a partir do Topo da Atmosfera). Os mesmos autores reportaram tabelas que fornecem informações necessárias para a conversão radiométrica de vários sensores.

$$\rho\lambda = \frac{\pi \cdot L\lambda \cdot d^2}{E_{sol\lambda} \cdot \text{sen}(\theta)} \quad (\text{Eq.2})$$

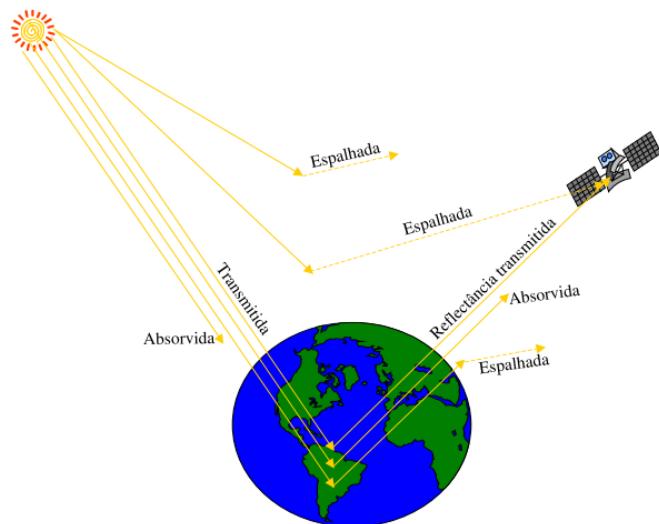
$\rho\lambda$: reflectância aparente ou topo da atmosfera (adimensional); $L\lambda$: radiância bidirecional aparente ($\text{W}/(\text{m}^2 \cdot \text{sr} \cdot \mu\text{m})$); d : distância sol- terra (em unidades astronômicas); $E_{sol\lambda}$: radiância solar emitida pelo sol no topo da atmosfera ($\text{W}/(\text{m}^2 \cdot \mu\text{m})$); θ : ângulo de elevação solar (graus decimais).

A reflectância é denominada pela razão da quantidade de energia radiante que deixa o objeto terrestre (radiância) por a quantidade de energia incidente da mesma área (irradiância) (MENESES e ALMEIDA, 2012). Conforme descrito por Ponzoni e Shimabukuro (2007), existem duas geometrias envolvidas na aquisição de dados, sendo uma delas a geometria de iluminação (formada pelos ângulos zenital e azimuthal da fonte de iluminação) e a geometria de visada (formada pelos ângulos zenital e azimuthal do sensor).

De acordo com Siqueira (2015), há três vantagens em usar a reflectância do topo da atmosfera (TOA): primeiro, este procedimento minimiza o efeito do ângulo zenital; a reflectância de TOA compensa diferentes valores da irradiância solar resultante de diferenças de intervalos espectrais da banda espectral; e a reflectância TOA corrige a variação da distância Terra-Sol entre diferentes datas de aquisição de dados. Song et al., (2001) afirmaram que a correção atmosférica pode ou não influenciar nos resultados de classificação, entretanto, a mesma é necessária quando se realiza uma comparação entre sensores ou entre datas de aquisições diferentes.

A interferência atmosférica é um dos principais fatores de deterioração nas cenas, sendo que a magnitude deste efeito está sujeita ao comprimento de onda considerado. Portanto, ele afeta de maneira distinta a cada uma das bandas de um sensor (FIGUEIREDO, 2005). De acordo com Vermote et al. (1997) antes da obtenção dos registros de sensores remotamente situados, ocorrem influências atmosféricas no trajeto Sol-Superfície e Superfície-Sensor, como apresentado na Figura 1.2.

Figura 1.2. Influências atmosféricas causadas na trajetória Sol – Atmosfera – Alvo – Atmosfera - Sensor.



Fonte: Figueiredo (2005).

Como supracitado, a reflectância é a razão entre a radiação refletida pela radiação incidente de um objeto na superfície terrestre. Esta é calculada inicialmente a partir do topo da atmosfera, desconsiderando efeitos provocados no trajeto do sensor até a superfície terrestre e vice-versa. Desta forma, os valores da Equação 2, representam os valores de reflectância dos objetos na superfície terrestre, acrescidos ainda da contribuição aditiva e negativa da atmosférica.

Diferentes aplicativos e técnicas foram gerados para corrigir os efeitos atmosféricos sobre cenas de sensores remotamente situados. Chandrasekhar (1960) apud Vermote et al., (1997) explicaram modelos baseados na teoria da transferência radiativa 6S (*Second Simulation of the Satellite Signal in the Solar Spectrum*). Este modelo requer parâmetros da atmosfera local no horário da aquisição da cena.

Chavez (1988 e 1989) propôs a correção atmosférica sobre os valores de reflectância com o método denominado *Dark Object Subtraction* (DOS). Esta metodologia assume que em qualquer imagem, existem pixels que deveriam admitir o valor “0” em alvos escuros (i.e., sombra de encostas, e corpos d’água). Valores acima de “0” nestas áreas seriam resultantes da contribuição aditiva da atmosfera. Portanto, admitindo que estas áreas deveriam apresentar valores “0”, este método usa a diferença encontrada para subtrair toda a cena o valor excedente. Assim, o método considera que a interferência atmosférica é uniforme em toda a extensão da imagem.

Lavreau (1991) desenvolveu um método de correção atmosférica que consiste na minimização do efeito de aerossóis como, fumaça, bruma, nuvens, entre outros. Este método utiliza as informações da própria imagem, baseando-se na faixa espectral do infravermelho, o qual são diferentes dos efeitos atmosféricos na faixa do visível.

De acordo com Felde et al. (2003) muitas cenas exigem a eliminação dos efeitos atmosféricos causado pela dispersão molecular e particulada. Partindo desta hipótese, criaram o *Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes* (FLAASH) que é derivado de um algoritmo baseado na física do código de transferência radiativa. Portanto, esta ferramenta de correção atmosférica corrige os comprimentos de onda nas regiões visíveis através do infravermelho próximo e do infravermelho de ondas curtas.

Além das correções radiométricas e atmosféricas, eventualmente faz-se necessário a correção de erros geométricos que, de acordo com Toutin (2004) desde o lançamento do primeiro satélite em 1970, o necessidades e requisitos para geração de imagens para uso do sensoriamento remoto processadas geometricamente mudou drasticamente.

Figueiredo (2005) afirma que os erros geométricos podem ser causados pelas oscilações do satélite em torno de eixos definidos por um sistema cartesianos posicionado no mesmo, (eixos x, y, z), variação da altitude do satélite (i.e., afeta a escala da imagem), variação da velocidade do satélite (i.e., provoca uma superposição ou afastamento de varreduras consecutivas) e movimento de rotação da Terra (i.e., provoca deslocamentos laterais gradual das linhas ao longo da imagem). Estas distorções podem ser corrigidas, ou seja, minimizadas, por meio de modelos matemáticos que descrevem as distorções existentes. As distorções geométricas, diminuem a precisão espacial das informações captadas pelos sensores remotamente situados.

Determinados termos e conceitos fundamentais do sensoriamento remoto, necessitam ser definidos para o futuro processamento, interpretação e bom emprego dos dados. Conforme Jensen (2009) o principal conceito fundamental faz referência à resolução de um sensor. Essa resolução trata sobre a “capacidade do sistema sensor em discriminar informações de detalhe”. Desta forma, a qualidade e a natureza da resolução é afetada em quatro dimensões: espacial, espectral, radiométrica e temporal (CAMPBELL, 1996). Estas resoluções serão explicitadas a seguir.

1.2.2.2. Resoluções dos sensores ópticos

As imagens ópticas adquiridas pelos sensores, possuem quatro resoluções indispensáveis para sua caracterização (i.e., resolução espectral, espacial, radiométrica e temporal).

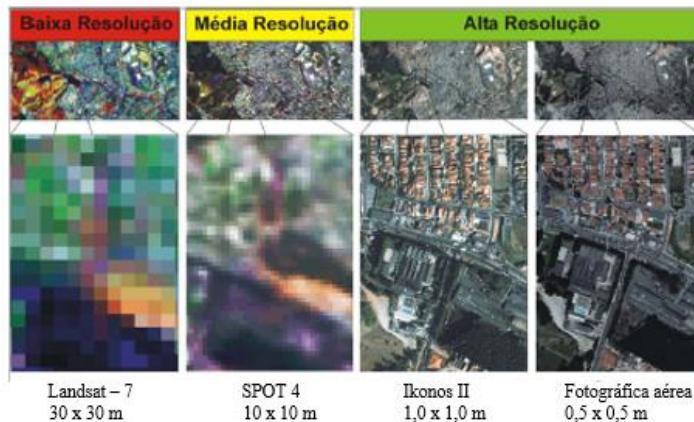
- Resolução Espectral

A resolução espectral é uma medida da largura das faixas espetrais e da sensibilidade do sensor em distinguir os níveis de intensidade do retorno do sinal (NOVO, 2010). Portanto, de acordo com Melo (2002) quanto maior o número de bandas e menor a largura do intervalo, maior será a discriminação dos objetos da superfície terrestre. Em outras palavras, melhor a resolução espectral.

- Resolução Espacial

Melo (2002) descreve a resolução espacial do sensor como à habilidade em distinguir e medir os objetos na superfície terrestre. Esta se baseia na projeção geométrica e pela área do campo de visada em uma certa altitude e em um determinado momento, ou seja, o tamanho do pixel. A Figura 1.3 ilustra as diferenças entre imagens de alta, média e baixa resolução espacial.

Figura 1.3. Diferentes resoluções espaciais para discriminar áreas.

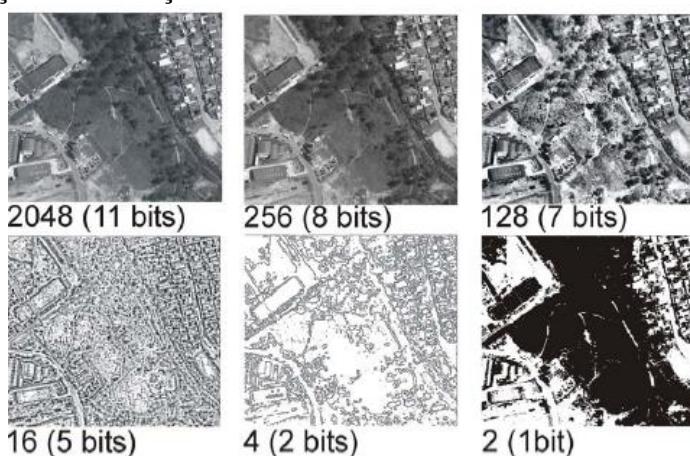


Fonte: Melo (2002).

- Resolução Radiométrica

Segundo Steffen (2017), a resolução radiométrica é caracterizada pelo número de bits, ou seja, níveis de cinza de uma dada imagem. Meneses e Almeida (2012) definiram a resolução radiométrica como sendo a intensidade de radiância detectada pelos sensores. A Figura 1.4 detalha visualmente a composição dos níveis de cinza. Além disso, a radiância de cada pixel passa por codificações digitais obtendo-se um valor numérico, ou seja, Número Digital (ND) (MELO, 2002).

Figura 1.4. Diferenças da resolução radiométrica em uma área urbana.



Fonte: Melo (2002).

- Resolução Temporal

A resolução temporal é definida pela frequência de passagem de um determinado sensor pelo mesmo local. Isso se deve pelas características deste, como a altura da plataforma, campo de visada instantânea, e sua capacidade de apontamento lateral. Esta resolução é importante para estudos que relacionam as mudanças na superfície da Terra e seu monitoramento com o

passar do tempo (BRYS, 2008). Estas resoluções, podem ser analisadas em conjunto ou individualmente, definindo qual o melhor sensor para ser usado em um determinado estudo.

1.2.2.3. Comportamento Espectral

O comportamento espectral dos alvos pode ser descrito pelo estudo da reflectância de superfície de cada material (i.e., solo, água, vegetação e etc.), sendo que cada material reflete de forma diferenciada em diferentes comprimentos de ondas. Do mesmo modo, com base no perfil espectral de um pixel, é possível inferir quais objetos compõem uma determinada cena. A REM ao incidir na superfície terrestre, permite três interações fundamentais, sendo elas, absorção, reflexão e transmissão (MENESES, MADEIRA NETTO, 2001).

Lillesand e Kiefer (2008) afirmaram que as diferenças das interações permitem a distinção das muitas feições presentes em uma imagem. Portanto, o conhecimento do comportamento espectral é muito importante para a escolha da região do espectro sobre a qual pretende-se trabalhar. O conhecimento dos espectros de reflectância é fundamental na definição de novos sensores (NOVO, 2010).

1.2.2.3.1. Comportamento Espectral da Vegetação

Estudos para extração de informações sobre ambientes utilizando o sensoriamento remoto, com vegetação têm apresentado cada vez mais frequentes (BLASCHKE e KUX, 2007). A assinatura espectral é a representação gráfica da reflectância de superfície em diferentes intervalos espetrais. Ponzoni, Shimabukuro e Kuplich (2012) descreveram detalhadamente sobre as características do espectro de reflectância de superfície da vegetação, desde a análise de reflectância de folhas isoladas até o comportamento espectral de dosséis complexos. Os mesmos autores resumiram o comportamento espectral da vegetação em três regiões distintas do espectro eletromagnético óptico, ou seja, dentro de 400nm a 2400nm, conforme mostrado na Tabela 1.1.

Tabela 1.1. Comportamento Espectral da Vegetação em diferentes regiões espetrais.

Regiões Espectrais	Comportamento da vegetação
Região do Visível (400 nm – 720 nm)	Pigmentos (Clorofila, carotenos e xantofilas) da folha dominam a reflectância São determinantes para a forma da curva dos fatores Os pigmentos dominantes absorvem a radiação na região do azul, porém somente a clorofila absorve na região do vermelho. A maioria das plantas são moderadamente transparente na região do verde.
Região do Infravermelho Próximo (720 nm -1100 nm)	Ocorre pequena absorção da radiação e espalhamento interno Absorção da agua é geralmente baixa e a reflectância é quase constante Quanto maior for lacunosa for a estrutura foliar maior será o espalhamento interno e consequentemente maiores serão os valores dos fatores de reflectância
Região do infravermelho Médio (1100 nm -32000 nm)	A absorção da decorrente da agua afeta a reflectância das folhas A medida que a folha fica mais seca aumenta os valores do fator de reflectância direcional-hemisférica

Adaptado: Ponzoni, Shimabukuro e Kuplich, (2012).

Para Epiphanio et al. (1996) a vegetação é um dos alvos mais complexos para o sensoriamento remoto, pois apresenta várias propriedades de reflexão foliar em uma mesma planta, ou seja, múltiplos componentes refletores. Logo, para facilitar os estudos da vegetação, índices de vegetação (IV) são amplamente utilizados. Segundo Watzlawick (2003) estas operações aritméticas, também conhecidas como IV têm como objetivo realçar as contribuições espetrais da vegetação e diminuir efeitos de topografia, de geometria solar, de contribuição de alvos adjacentes, etc.

O IV mais difundido é o Índice de Vegetação da Diferença Normalizada (NDVI) descrito por Rouse et al. (1973). Os IVs normalmente utilizam as bandas espetrais do vermelho e do infravermelho próximo para a geração de uma imagem contraste, evidenciando, portanto, áreas com presença de vegetação. Ponzoni e Shimabukuro (2007) salientam que, além do NDVI, índices como: Índice de vegetação de Razão Simples (RVI) e Índice de Vegetação Ajustado para o Solo (SAVI), também são amplamente aplicados.

Diversos estudos desenvolvidos por Adami, Gomes e Coutinho (2015) utilizaram IVs para extrair informações sobre a cobertura da terra no estado do Pará. Huang, Davis e Townshen (2002) utilizaram IVs fusionados com dados de micro-ondas para estimativas de parâmetros da vegetação. Sano, Marra e Meneses (2009) e Picoli (2012) utilizaram a combinação de imagens de sensores micro-ondas e ópticos para classificação de áreas agrícolas. A escassez de produtos utilizando dados micro-ondas ainda é grande ao se comparar com trabalhos utilizando dados de sensores ópticos. Logo, a aplicação combinada de dados micro-ondas e ópticos ainda é incipiente. Assim, destaca-se a importância de estudos utilizando dados de ambos sensores.

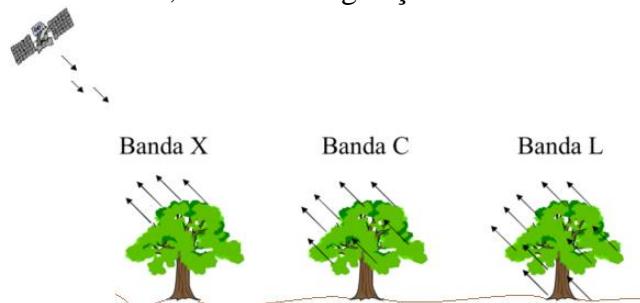
1.2.3. Imageamento Micro-ondas

Como exposto nos tópicos anteriores, os dados ópticos analisam as propriedades microscópicas dos alvos. No entanto, os dados de micro-ondas, ou seja, Radar (i.e., Radio Detection and Ranging – Detecção e localização por meio de ondas de rádio), são influenciados pela textura, estrutura vertical e horizontal dos alvos, além de suas propriedades dielétricas. Estes são capazes de atravessar nuvens, e dependendo da banda utilizada, uma maior ou menor penetrabilidade em dosséis e solos (PEREIRA, 2012 e PONZONI et al. 2012). Oliveira (2012) afirmou que desde a década de 1960 os sensores de micro-ondas vêm sendo utilizados para aplicações militares. Porém, somente nos anos 1970 foi possibilitado o uso civil, como realizações de projetos em grandes extensões territoriais, a exemplo do RADAMBRASIL. Este projeto recomendou, ao final, a criação 35.200.000 ha de unidades de conservação de proteção integral e mais 71.500.000 ha de uso sustentável na Amazônia.

Kasischke, Melack e Dobson (1997) afirmaram que por falta de experiência de interpretação dos dados de micro-ondas, estes não foram amplamente estudados como os dados de imagens ópticas. Após duas décadas, estudos estão sendo realizados, o que tem apontado um crescimento acelerado no desenvolvimento e nos produtos *Synthetic Aperture Radar* (SAR) (PONZONI, SHIMABUKURO e KUPLICH, 2012).

Os sensores SAR são sistemas ativos, que emitem a própria REM e podem operar em diferentes comprimentos de onda. A Figura 1.5 mostra o mecanismo de interação das bandas dos sensores micro-ondas com a vegetação. Estes tem visada lateral e emitem pulso de radiação ao longo da linha de voo, registrando a energia espalhada que é conhecida como retroespalhamento (i.e., se refere à radiação que foi espalhada na direção do micro-ondas após retornar da superfície terrestre) (WARING et al. 1995; OLIVEIRA, 2012; PONZONI, SHIMABUKURO e KUPLICH, 2012).

Figura 1.5. Interação das bandas X, C e L em vegetação.



Fonte: Adaptado de Leckie e Ranson (1998).

De acordo com Meneses e Almeida (2012), as bandas representam frequências que são identificadas por letras. Conforme apresentado na Tabela 1.2, a interação da REM na região do micro-ondas com os objetos na superfície terrestre depende da banda utilizada, sendo que a penetrabilidade das ondas nos objetos aumenta com o comprimento da onda das mesmas.

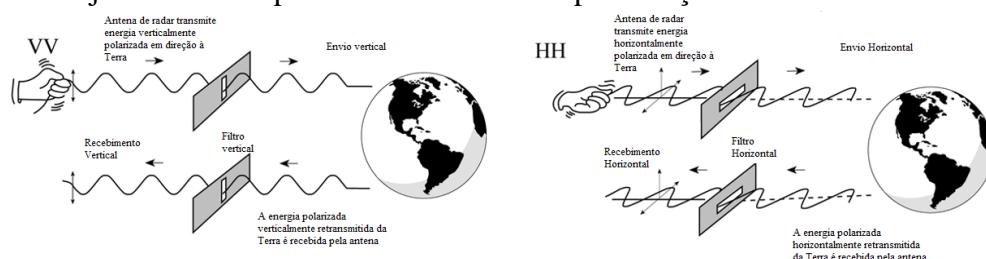
Tabela 1.2. Descrição de Bandas Utilizadas por Sensores SAR.

Banda micro-ondas	Comprimento de onda (cm)	Frequência (MHz)
P	136-77	220-390
UHF	100-30	300-1.000
L	30-15	1.000-2.000
S	15-7,50	2.000-4.000
C	7,50-3,75	4.000-8.000
X	3,75-2,40	8.000-12.500
Ku	2,40-1,67	12.500-18.000
K	1,67-1,18	18.000-26.500
Ka	1,18-0,75	26.500-40.000

Fonte: Adaptado por Ponzoni, Shimabukuro e Kuplich (2012).

Geralmente os sistemas SAR são compostos por transmissor, receptor, modulador e processador de antena, onde as principais características e propriedades da antenas são: polarização, radiação, largura do feixe, lóbulos laterais, direatividade, ganho, impedância de entrada, área de recepção e reciprocidade (DUTRA et al., 2003). As definições dos produtos gerados pelos sensores SAR são: a polarização, que é definida pela trajetória do campo elétrico em um plano, podendo ser horizontal, (i.e., se desloca paralelamente ao eixo de propagação) ou vertical (i.e., que se descola perpendicularmente ao eixo) (PONZONI, SHIMABUKURO e KUPLICH, 2012 e MENESSES e ALMEIDA, 2012). De acordo com Henderson e Lewis (1998), a polarização HH, indica que a energia é transmitida e recebida horizontalmente. VV é transmitida e recebida verticalmente. HV ou VH a polarização é denominada polarização cruzada (Figura 1.6).

Figura 1.6. Trajetória do campo elétrico referente as polarizações



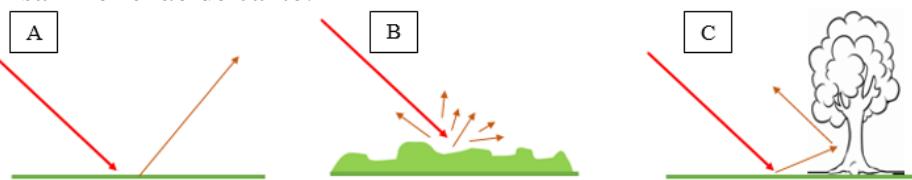
Fonte: adaptado de Jensen (2009).

De acordo com Ponzoni, Shimabukuro e Kuplich (2012), além da amplitude, os sistemas SAR registram a fase do retroespalhamento, que tem ampla utilização em técnicas como polarimetria e interferometria. Estes lidam com a natureza vetorial do campo elétrico da REM. Sendo, a polarimetria, a diferença de fase entre duas medidas SAR e interferometria, diferenças em posições ligeiramente diferentes do sensor. Portanto, através destes diversos recursos, estudos vêm sendo realizados principalmente em áreas de agricultura e floresta. Ulaby (1975) realizou um dos primeiros estudos empregando dados SAR em vegetação. Desde então, Le Toan et al. (1992) e Ponzoni, Shimabukuro e Kuplich (2012), Main et al. (2016), Chen et al. (2016) vêm desenvolvendo técnicas para o aprimoramento das mesmas em sensores micro-ondas e/ou ópticos, sendo eles integrados ou fusionados.

Estudos de vegetação normalmente têm como principal objetivo, a discriminação das formações vegetais, a verificação de sua extensão territorial, e estimativas de suas propriedades biofísicas e bioquímicas. Em geral, a diferenciação de áreas florestais e não florestais é facilitada com o uso de dados SAR quando utilizada em mais de uma banda, polarização e ângulo de incidência (HYYPAA et al., 2000; RANSON et al., 2001; BOYD e DANSON, 2005).

Ponzoni, Shimabukuro e Kuplich (2012) relataram que os objetos da superfície terrestre interagem com as micro-ondas segundo as suas propriedades geométricas. Como mostrado na Figura 1.7, os objetos apresentam relações diretas com a polarização incidente. Todavia, os objetos podem ter a capacidade de despolarizar a onda incidente, gerando o mecanismo denominado como retroespalhamento. Deste modo, podem ocorrer três tipos de mecanismos de retroespalhamento, sendo eles: especular (i.e., quando a superfície é lisa em relação ao comprimento de onda e o espalhamento ocorre na direção oposta ao sistema SAR), difuso (i.e., quando a superfície é rugosa e o espalhamento ocorrem em várias direções) e reflexão de canto (i.e., quando o espalhamento ocorre em duas ou mais superfícies lisas adjacentes).

Figura 1.7. Tipos de superfície e espalhamento associados: (A) lisa - Especular, (B) rugosa-Difuso, (C) lisa - Reflexão de canto.



Fonte: adaptado de Ponzoni, Shimabukuro e Kuplich (2012).

Igualmente como imageadores ópticos, são necessárias correções e parametrizações para utilização dos dados SAR.

1.2.3.1. Processamento imagens SAR

Antes de qualquer processamento, envolvendo dados SAR, faz-se necessário realizar calibração radiométrica, minimizar distorções geométricas e ruídos inerentes. O “Speckle” é um fenômeno muito comum em dados Laser, acústicos e SAR, sendo um fato causado pelas interferências aleatórias construtivas e destrutivas no sinal (CSRE, 2017 e JENSEN, 2009). Para Gagnon e Jouan (1997) e Pohl e Genderen (1998) o ruído é um efeito indesejável, e a realização do procedimento de redução é de suma importância. Conforme observado na Figura 1.8, notam-se a presença de ruídos quando comparados visualmente com a imagem original. No presente caso, a presença de ruído “Sal e pimenta” refere-se para sensores ópticos enquanto efeito “Speckle” para dados adquiridos por sensores SAR.

Figura 1.8. Imagem sob efeito de diferentes tipos de ruídos.



Fonte: Paula Júnior (2017).

O ruído “Speckle”, resulta nos dados SAR, pontos muito claros (interferência construtiva) ou muito escuros (interferência destrutiva), apresentando assim variações bruscas na intensidade da cena (GOODMAN, 1976). Este processo tem caráter multiplicativo, portanto, onde o sinal é maior, o ruído é mais intenso (SANT’ ANNA, 1995). Diversos filtros foram desenvolvidos para a redução do efeito “Speckle” visando uma melhor separabilidade entre os objetos da superfície terrestre (DUTRA et al., 2003 e PARIZZI e BRCIC, 2011). Uma das maneiras de se reduzir o efeito “Speckle” em dados SAR pode ser feito por meio do uso de filtros e/ou processamento *multilooking* (SANT’ ANNA, 1995).

O *Multilooking* é um procedimento essencial para a melhoria da cena de um sensor SAR, visando a redução do “Speckle”. Este procedimento torna o formato retangular do elemento de resolução para um formato próximo ao quadrangular. No entanto, esta operação é realizada partindo da hipótese de que há homogeneidade da média dos elementos (SARMAP, 2012). De acordo com Sant’ Anna (1995) outro exemplo de diminuição do efeito “Speckle” é o filtro Lee (LEE, 1983), que se baseia na média e variância de um pixel que é igual à média e variância

local baseada nos pixels de uma vizinhança fixa. Além dessas técnicas mais frequentemente utilizadas, a filtragem em especial, é uma técnica que realça a imagem, transformando pixel a pixel e levando em consideração os pixels vizinhos. Sant' Anna (1995) descreveu diferentes filtros como: Filtro da Média, de ordem, Modelo de Degradação, LLMMSE (Lee e Kuan-Nathan), MMSE (Frost), robustos e Estimativa de Máximo a posteriori.

Posteriormente a minimização dos ruídos indesejáveis nos sensores micro-ondas, Dutra et al., (2003) afirmaram que nestes sensores há disponibilidade de geração de diversos produtos com o desenvolvimento de técnicas, como por exemplo as decomposições dos alvos, que são comumente utilizadas para as avaliações de recursos naturais.

1.2.3.2. Decomposições polarimétricas

Os dados oriundos de sensores SAR consistem em informações da amplitude e fase. A amplitude se relaciona a intensidade da reflexão da onda eletromagnética na cena, ou seja, fornece dados oriundos da refletividade dos alvos, enquanto a fase é relativa ao trajeto entre a antena e a cena e ao mecanismo de reflexão que altera a fase da onda incidente. As técnicas de decomposição polarimétrica se desenvolveram durante a década de 1990 e foram fundamentadas em modelos de matrizes de coerência e covariância (KUMAR, 2016). Portanto, existem diferentes técnicas de decomposição, sendo eles: decomposição coerente, baseados em modelos de espalhamento e decomposição incoerente fundamentado em autovalores e autovetores (LEE e POTTIER, 2009).

1.2.3.2.1. Decomposições coerentes

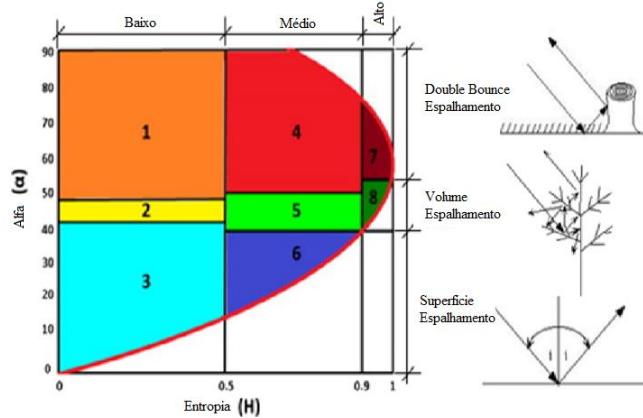
As decomposições coerentes são aquelas que utilizam a matriz de Sinclair proposta por Sinclair (1950), que têm por objetivo decompor tal matriz em outras n matrizes, correspondendo a um tipo de mecanismo de espalhamento (CLOUDE e POTTIER, 1997; LEE e POTTIER, 2009). Neste caso, a matriz de espelhamento recebida é expressa com a soma coerente das matrizes elementares.

De acordo com Ulaby e Elachi (1990) a decomposição coerente é complexa, e expressa como o espalhador transforma o campo transmitido (E^t) no recebido (E^r), assim caracterizando o comportamento do espalhador. As decomposições mais empregadas para utilizações em áreas agrícolas, são as propostas por Freeman e Durden, Lewis e Henderson e Cloude Pottier.

A decomposição de Freeman e Durden (1998) possui três componentes sendo eles: Espalhamento Superficial, Volumétrico e *double-bounce*. Outra decomposição coerente é a

diferença de Fase proposta por Lewis e Henderson (1998), no qual indica a estrutura e a quantidade de biomassa. A decomposição de Cloude e Pottier (1997) supõe que há um mecanismo dominante médio em cada célula, o qual pode ser representado por três parâmetros invariantes à rotação da plataforma: Ângulo alfa médio, Entropia e Anisotropia. Esta decomposição pode-se considerar em um espaço de classificação bidimensional, como apresentado na Figura 1.9, que demonstra as zonas para cada mecanismo de retroespalhamento (*Double Bounce*, Volume e superfície).

Figura 1.9. Plano Bidimensional de Entropia (H) e Ângulo Alpha (α).



Fonte: Cloude e Pottier (1997).

Segundo Cloude e Pottier (1997) são nove diferentes zonas válidas, que representam diferentes mecanismos de retroespalhamento. Uma descrição de cada zona pode ser encontrada na Tabela 1.3. Logo, a complexidade da estrutura de uma floresta, pode ser explicada por apresentar alta entropia, consequentemente uma baixa anisotropia e valor alfa perto de 45° (CLOUDE e POTTIER, 1997).

Tabela 1.3. Mecanismos de Espalhamento das Zonas Bidimensional.

Zonas	Mecanismo de Espalhamento	Representação (Exemplos)
1	Espalhamento múltiplo com baixa entropia. Esta zona é caracterizada por $\bar{\alpha} > 44,5$	Espalhadores diédricos metálicos isolados
2	Espalhamento dipolo com baixa entropia: nesta zona ocorrem mecanismos fortemente correlacionados. Um espalhador dipolo isolado e espalhamento da vegetação com orientação fortemente correlacionada de elementos com espalhamento anisotrópico	Vegetação com elementos de espalhamento anisotrópicos altamente correlacionados
3	Espalhamento superficial com baixa entropia: correspondem aos valores de $\bar{\alpha} < 42,5^\circ$, incluem espalhamento superficial e especular	Água nas bandas P e L, oceano na banda L e superfícies bem suaves
4	Espalhamento múltiplo com média entropia	Double-bounce em áreas urbanas e em florestas
5	Espalhamento da vegetação com média entropia: moderada entropia com mecanismos de espalhamento do tipo dipolo dominante	Superfícies vegetadas com espalhadores anisotrópicos e moderada correlação das orientações do espalhador
6	Espalhamento superficial com média entropia: esta zona reflete o aumento na entropia devido às mudanças na rugosidade da superfície e devido aos efeitos de propagação da copa	Propagação no dossel ou elevação da rugosidade da superfície
7	Alta entropia com espalhamento múltiplo	Camadas de vegetação com estrutura geométrica bem desenvolvida
8	Espalhamento da vegetação com alta entropia: o espalhamento volumétrico aumenta quando $\bar{\alpha} = 45^\circ$ e $H > 0,9$	Espalhamento de copas de florestas e espalhamento de alguns tipos de superfícies vegetadas
9	Não faz parte de uma região válida, pois é impossível distinguir espalhamento superficial com $H > 0,9$	-

Fonte: Cloude e Pottier (1997). A numeração corresponde aos intervalos ilustrados na Figura 1.9.

1.2.3.2.2. Decomposições Incoerentes

As decomposições incoerentes são informações que podem ser geradas a partir dos coeficientes de retroespalhamento dos dados SAR. Henderson e Lewis (1998) propuseram, para uma dada polarização, o coeficiente de retroespalhamento denominado de sigma zero (σ^0 ; dB). O sigma zero depende de diversos parâmetros do alvo (tamanho, forma, propriedades dielétricas, orientação, rugosidade).

A Taxa de Co-Polarização, proposta por Henderson e Lewis (1998), é a razão entre os dois retroespalhamentos obtidos na vertical e horizontal, respectivamente. Logo, através da taxa

de Co-Polarização destacam-se diferentes orientações verticais e horizontais dos alvos, quais são derivadas dos aspectos estruturais da vegetação. Também, proposta pelos mesmos autores, a Relação de Polarização Cruzada, é a razão entre o retroespalhamento da polarização cruzada pelo retroespalhamento horizontal. Esta razão, é sensível à dispersão volumétrica de uma floresta e reduz os efeitos topográficos no retroespalhamento.

Pope, Rey-Benayas e Paris (1994) propuseram três índices: Índice de Biomassa (utilizado para diferenciar áreas vegetadas e não vegetadas, como também, entre terrenos inclinados e planos), Índice de estrutura de copa (aplicado para diferenciação de florestas de pântano) e o Índice de espalhamento de volume (empregado para distinguir entre florestas de terra firme e em regeneração). A Potência Total sugerida por Boerner et al. (1991) é a soma da potência de todas as polarizações espalhadas de um alvo de volta ao SAR. O Índice de Degradação de Florestas sugerida por Saatchi et al. (2010) e Mitchard et al. (2012) é a relação da diferença do retroespalhamento horizontal e o cruzado pela somatória do retroespalhamento horizontal e o cruzado. Este índice avalia a força do mecanismo de double-bounce, logo, é utilizado para diferenciação de grupos vegetações.

1.2.4. Métodos De Classificações para Geração de Mapas Temáticos

O sensoriamento remoto, têm como uns dos principais objetivos, as extrações de informações que possam ser lidas em formato de tabelas, gráficos e mapas. Sendo do mesmo modo, para classificação de dados provenientes do sensoriamento remoto, requerem-se significativos recursos computacionais (FIGUEIREDO, 2005; MENESSES e ALMEIDA, 2012). Logo, *Softwares* foram desenvolvidos para aplicar os algoritmos de classificações de uso e cobertura da terra. Os principais *softwares* desenvolvidos são o SPRING, ENVI, ArcMAP, SNAP, ERDAS, EnMAP Toolbox entre outros.

De acordo com Meneses e Almeida (2012) há diversos métodos de classificação que podem serem divididos em: Classificação Paramétrica e Não-Paramétrica, classificação espectral e espacial, e classificação supervisionada ou não-supervisionada. Ainda, é possível usar classificadores por pixel, por regiões ou por objeto. De forma geral, os métodos supervisionados são aqueles que buscam encontrar características semelhantes àquelas previamente amostradas para cada classe pré-estabelecida, tendo como objetivo agrupar *pixels* com a mesma resposta espectral ou um mesmo conjunto de critérios (NAVULUR, 2006).

Atualmente existem muitos métodos de classificação que visam procurar informações dos componentes de uma cena. Portanto, os níveis de acurácia pelos diferentes métodos tendem

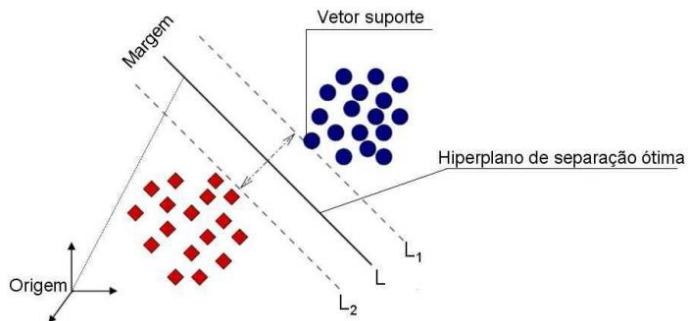
a variar de acordo com o algoritmo empregado e a configuração do classificador. Os classificadores não paramétricos, não seguem ao parâmetros da curva normal, pois não assumem suposições sobre a distribuição de probabilidade e geralmente são reforçados por trabalhar com ampla variedade de distribuição de classe (SEMOLINI, 2002; ANDRADE, FRANCISCO e ALMEIDA, 2014). O emprego do algoritmo correto, para fins de classificação de uso e cobertura da terra. É fundamental para direcionar um bom planejamento apropriado.

As técnicas de classificadores não-paramétricos que mais se destacam são: *Support Vector Machine* (SVM) e o *Random Forest* (RF). O SVM é conhecido pelo alto grau de generalização e por ser fundamentado na probabilidade estatística. Estudos demonstraram que o SVM tem proporcionado resultados superiores a outros métodos. Pesquisadores como Huang, Davis e Townshen (2002), Foody e Mathur (2004) e Pal e Mather (2005) compararam o SVM com outros classificadores (e.g., redes neurais e árvores de decisão) e demonstraram que o classificador se igualou ou mesmo superou tais métodos. Já o RF, é um algoritmo que utiliza múltiplas árvores de decisões para parametrizar modelos e assim estimar as variáveis categóricas ou contínuas (BREIMAN, 2001).

1.2.4.1. Support Vector Machine (SVM)

O SVM foi introduzido por meio da teoria estatística de aprendizagem, assumindo como apoio a superfície de decisão, conhecida como hiperplano ideal, o qual eleva ao máximo os limites de separação de classes (VAPNIK, 1999; NASCIMENTO et al., 2009). Como demonstrado na Figura 1.10, os pontos que ficam próximos a superfície são denominados como, *Support vectors* (Vetores de suporte), e isto ocorre durante a fase de treinamento (BURGES, 1998). De acordo com Tso e Mather (2009) o usuário pode controlar a dualidade entre o rigor do hiperplano, aceitando ou rejeitando os erros de treinamento. Este é um classificador supervisionado e não-paramétrico, com principal objetivo de determinar limites de decisão que permitam a separação ótima entre classes e minimização dos erros (NASCIMENTO et al., 2009).

Figura 1.10. Esquema de classificação do Support Vector Machine (SVM).



Fonte: Nascimento et al. (2009).

O SVM além de realizar a classificação linear, desempenha também a classificação não-linear usando as função-núcleos, ou seja, *Kernel*. De acordo com Nascimento et al. (2009) o SVM possui quatro funções principais (i.e., linear, quadrática, polinomial e base radial), os quais Lorena e Carvalho (2007) descrevem detalhadamente em seu trabalho. Steinwart e Christmann (2008) apontam três motivos para o sucesso do algoritmo: capacidade de aprender com um número reduzido de parâmetros de entrada; robustez diante de vários tipos de violações de modelo e da diversidade de modelos e; eficácia computacional em comparação com outros algoritmos.

Os SVM é um classificador que gera normalmente bons resultados para dados ruidosos e complexos. Em geral os algoritmos orientados em SVM extraem os parâmetros que permitem a generalização. Estes armazenam os ruídos tolerando o reconhecimento das amostras não observados durante a fase de treinamento (ALIXANDRINI, 2009). Além disso, revela-se que tem grande capacidade de percepção para dados com alta entropia, definida como a quantidade de incerteza associada (aleatoriedade), ou seja, quanto maior a entropia, mais atípica (fora de padrão) será a cena, acontecimentos nos quais outros classificadores encontram problemas (LORENA e CARVALHO, 2007; ANDRADE, FRANCISCO e ALMEIDA, 2014).

1.2.4.2. Random Forest (RF)

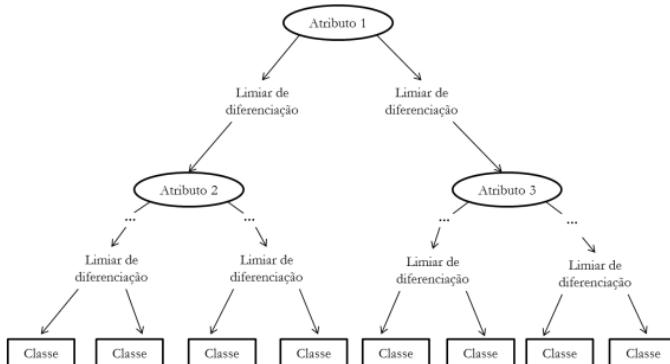
Random Forest (RF) é um algoritmo proposto por Breiman (2001) e é fundamentado em árvores de decisões, “Conjunto” (ensemble), que de acordo com Kuncheva (2004) são métodos que geram muitas classificações e combinam seus resultados, priorizando os melhores. Portanto, o RF constrói muitas árvores de decisão, os quais são utilizadas para maximizar a acurácia da classificação. O RF contempla dados de referência contínuas ou categóricas, além de analisar a cena em duas etapas, constituída por parametrização de modelo e por aplicação do

modelo, ou seja, classificação. Além disso, apresenta precisão *out-of-bag* (OOB) e importância da variável (WASKE et al., 2012).

OOB é muito utilizado para estimar erros de classificação, portanto o RF coleta dois terços (2/3) de treinamento com substituição, enquanto os demais são deixados OOB e estas amostras são atribuídas a árvores que ainda não foram utilizadas, sendo que a diferença é usada entre a classe esperada e a real, igualmente avaliando a exatidão da classificação (PRASAD, IVERSON e LIAW, 2006).

RF implementa uma abordagem que utiliza múltiplas árvores de decisão para parametrizar modelos e assim estimar as variáveis categóricas ou contínuas. De acordo com Polikar (2006) este procedimento parte do princípio que diferentes preditores independentes preveem incorretamente em diferentes áreas, sendo do mesmo modo, ao combinar a previsão é possível melhorar a precisão geral. Segundo Piazza (2014) as árvores são demonstradas no fluxograma (Figura 1.11), onde o nó interno denota um teste com o atributo, o ramo representa o resultado do teste, e o nó externo (folha) exibe a classe.

Figura 1.11. Esquema estrutural do resultado de uma árvore de decisão.



Fonte: Piazza (2014).

De acordo com Neto (2014) este classificador tem por objetivo efetuar a criação de diversas árvores de decisão usando um subconjunto de características selecionadas aleatoriamente de acordo com o conjunto original, chamado de *bootstrap*. Segundo Diniz, Neto e Fontes, (2013) esta técnica consiste em associação de classificadores do tipo “árvore de decisão”, no qual para determinar a classe é realizado por um mecanismo de votação, sendo assim a classificação final é dada pela classe que recebeu maior números de votos. Segundo Caruana, Karampatziakis e Yessenalina (2008) RF é uma das técnicas consideradas mais precisas quando comparada as demais técnicas Não-Paramétricas.

Diferentes métodos de classificação são utilizados para determinados grupo de dados, porém todos são avaliados através de sua acurácia e exatidão da mesma forma para que se consiga comparar e analisar qual o melhor método para determinado estudo.

1.2.4.3. Avaliação De Mapas Temáticos

Sobre a verificação de qualidade de mapas temáticos, a partir de um processo de classificação, Novo (2010), propõe que se avalie a exatidão, utilizando dados amostrados com os resultados obtidos. Congalton e Green (1999) definiram que a exatidão temática está relacionada à acurácia das classes do mapa em relação à realidade, de tal modo, a forma mais comum de avaliação é a matriz de confusão. Essa é representada por uma matriz com números de colunas e linhas iguais, como demonstrado na Tabela 1.4, em que as colunas representam os dados de referências e as linhas indicam o classificado; já os elementos na diagonal da matriz indicam o nível de acerto entre os conjuntos de dados (CONGALTON e GREEN, 1999; LILLESAND, KIEFER e CHIPMAN, 2004).

Tabela 1.4. Representação da Matriz de Confusão.

Classes/Classes	1	2	...	K	Total da Linha X_{i+}
1	X_{11}		...	X_{1K}	X_{1+}
2	X_{21}	X_{22}	...	X_{2K}	X_{2+}
...
K	X_{K1}	X_{K2}	...	X_{KK}	X_{K+}
Total da Coluna X_{+i}			...	X_{+K}	

Através desta matriz, pode-se calcular medidas descritivas (e.g., exatidão geral, exatidão do usuário e exatidão do produtor). A Exatidão Geral (eg), de acordo com a Equação 4 (Eq.4), é dada pela razão entre a soma de todos os elementos classificados corretamente pelo total de amostra. Entretanto, de acordo com Cadena (2011) este tende a superestimar o resultado da classificação, pois considera apenas a proporção de concordância correta.

$$eg = \frac{\sum_{i=1}^r x_{ii}}{N} \quad (\text{Eq.3})$$

Onde: r: número de classes; X_{ii} : elementos da diagonal principal e; N: número total de amostras.

A Exatidão do Usuário (eu), conforme a (Eq. 5) é dada pela razão do número de classificação correta de determinada classe pelo número total de amostras classificadas na mesma. De acordo com Lillesand, Kiefer e Chipman (2004), esta exatidão reflete os erros de

comissão, ou seja, indica a probabilidade de um elemento classificado em uma determinada classe realmente pertencer à mesma.

$$eu = \frac{x_{ii}}{x_{i+}} \quad (\text{Eq.4})$$

Onde: X_{ii} : elementos classificados corretamente em determinada classe; X_{i+} : total de elementos classificados para uma classe determinada.

A razão entre o número de elementos classificados corretamente de uma classe pelo número de elementos de validação amostrados na mesma classe é denominada como Exatidão do Produtor (ep), apresentado na equação 6 (Eq. 6). Este indica os erros de omissão da classificação (LILLESAND, KIEFER e CHIPMAN, 2004).

$$ep = \frac{x_{ii}}{x_{i+}} \quad (\text{Eq.5})$$

Onde: X_{ii} : elementos classificados corretamente e; X_{i+} : total de elementos de validação para uma classe determinada.

Além destas métricas de avaliações de exatidão, é habitualmente aplicado o índice Kappa, entretanto, tal índice é questionado por vários autores de acordo com o trabalho realizado por Powers (2012). Segundo Meneses e Almeida (2012) o índice Kappa (K) é um método muito utilizado na avaliação dos resultados de classificação, sendo introduzido por Cohen (1960). O Índice Kappa varia de 0 (concordância nula) a 1 (concordância perfeita). Esta é uma medida largamente utilizada da acurácia de classificação em vários estudos (LEE, et al., 2001; LIM et al., 2007).

$$K = \frac{\frac{\sum_{i=1}^c x_{ii}}{n} - \frac{\sum_{i=1}^c x_{i+} + x_{+i}}{n^2}}{1 - \frac{\sum_{i=1}^c x_{i+}}{n}} \quad (\text{Eq.6})$$

Onde: C: número total de classes; X_{ii} : é o número de observações na linha i coluna i, respectivamente; X_i : é a soma da linha i da matriz de confusão; X_{+i} : é a soma da coluna i da matriz de confusão e; N= número total de amostras

Para qualificar o índice Kappa, apresentado na Tabela 1.5., Landis e Koch (1977) sugeriram uma métrica que categoriza a classificação associada a este.

Tabela 1.5. Categorização de uma classificação a partir do Índice Kappa.

Estatística Kappa	Classificação
<0	Pobre
0,00-0,20	Ruim
0,21-0,40	Razoável
0,41-0,60	Boa
0,61-0,80	Muito Boa
0,81-1,00	Excelente

Fonte: Landis e Koch (1977).

Além do Índice Kappa convencional existe o Kappa condicional, proposto por Rosenfield e Fitzpatrick-Lins (1986), o qual permite avaliar a acurácia de uma determinada classe. De acordo com Andrade, Francisco e Almeida (2014), esse é calculado com o mesmo princípio que o Kappa, no entanto avaliando os erros de comissão (Eq.8) e o erros de omissão (Eq.9).

$$K = \frac{(n * x_{ii}) - (x_i * x_{+i})}{(n * x_{i+}) - (x_{i+} * x_{+i})} \quad (\text{Eq.7})$$

$$K = \frac{(n * x_{ii}) - (x_i * x_{+i})}{(n * x_{+i}) - (x_{i+} * x_{+i})} \quad (\text{Eq.8})$$

Onde: n: é o numero total de elementos de referencia; xii: é o numero total de elementos classificados corretamente; xi+: é o total de elementos classificados para uma classe i; e x+i: é o total de elementos de referencia amostrado para uma classe i.

Para Congalton e Green (1999) existem meios para testar a significância estatística do índice Kappa. Para a comparação de duas classificações é utilizado o Teste-Z. Esta testa a significância, como apresentado na equação 10 (Eq. 10).

$$z = \frac{|K1 - K2|}{\sqrt{\text{var}(K1) + \text{var}(K2)}} \quad (\text{Eq.9})$$

Onde: K1 e K2 são as estimativas da estatística de Kappa para a matriz de erro #1 e #2 e var (K1) e var (K2) são correspondentes as estimativas de variância.

1.2.5. Aplicação do Sensoriamento Remoto na Ciência Florestal

Desde o início dos lançamentos dos satélites de observação da Terra, comprehendia-se que o Sensoriamento Remoto orbital abrangeia grandes áreas e repetitividade temporal, o qual teria um papel insubstituível na geração de informações para países de grandes extensões territoriais (SHIMABUKURO, MAEDA e FORMAGGIO, 2009). Moreira (2011) afirma que a utilização de geotecnologias para recursos naturais é uma ferramenta poderosa para geração

de dados que, de acordo com Nunes (2008), são ideais para organizar informações, planejar e criar cenários distintos para diferentes análises.

Para a aplicação do sensoriamento remoto em estudos correlatos, escolher produtos que atendam à necessidade da aplicação pretendida é de suma importância. Com por exemplo os autores Assad e Sano (1998) que para discriminarem áreas cobertas com florestas, desmatadas e queimadas, utilizaram o sensor *Advanced Very High Resolution Radiometer* (AVHRR/NOAA) que apresenta uma resolução espacial grosseira (1,1 km). Dória, Amaral e Monteiro (2014) pesquisaram o crescimento populacional da região metropolitana de São Paulo através do sensor *Visible/Infrared Imaging Radiometer Suite* (VIIRS), com resolução espacial de 500 m, e afirmaram que os dados foram consistentes com o censo demográfico levantado no ano de 2010.

Adami, Gomes e Coutinho (2015) realizaram classificações com o sensor *MODerate resolution Imaging Spectroradiometer* (MODIS) do ano de 2008 a 2012 a fim de verificar a mudança de uso e ocupação da terra no Estado do Pará. Estes autores notaram que houve diferenças significativas ao se comparar áreas que eram de florestas em 2008 e que, nos anos seguintes, passaram a ser pastagem e agricultura. De acordo com Boyd e Danson (2005) os estudos utilizando imagens de sensoriamento remoto para aplicações em áreas com recursos florestais, fornecem três níveis de informações: a extensão da cobertura florestal; tipo de floresta e propriedades biofísicas e bioquímicas da floresta.

Como supracitado nos tópicos anteriores, os dados ópticos tendem a explorar propriedade físicas e químicas da superfície terrestre. Entretanto, Florenzano (2002) afirmou que os dados ópticos são influenciados pela presença de nuvens, impedindo que a energia refletida chegue ao sensor. Logo, os sensores SAR vêm suprir esta lacuna, pois normalmente penetram as nuvens, chuva, aerossóis ou nevoeiros, permitindo assim cenas independentemente de situações extremas.

Estudos de classificação de uso e cobertura da terra estão sendo aplicados com a utilização de sensores SAR, como a pesquisa de Muthukumarasamy, Shanmugam e Kolanuvada (2017), que utilizaram o sensor na banda L para classificações, ou como Mishra et al. (2017) que empregaram sensor na banda C. Além desta utilização, os sensores SAR oferecem parâmetros estrutural dos objetos da superfície terrestre. Assim, Viet Nguyen et al. (2016b) utilizaram a banda L para caracterizar as estruturas de florestas tropicais. Martins et al. (2016) pesquisaram a sensibilidade do sensor ALOS/PALSAR na banda L para detecção de áreas degradadas com fogo no norte da Amazônia.

É notório que diversos estudos estão sendo realizados a fim de identificação, classificação do uso e cobertura da terra, tanto com sensores SAR e Ópticos; entretanto, ainda é vago estudo com aplicação da combinação desses sensores, tanto para classificações quanto para extração de propriedade biofísicas.

1.3. CONCLUSÕES

Logo, pode-se concluir que o sensoriamento remoto vem contribuído com diversas informações de suma importância para a ciências agrárias. Foi possível notar, por meio da revisão, que a utilização de sensores ópticos são bastante difundidos, tanto no meio acadêmico, quanto no mercado de trabalho. Já os sensores SAR estão expandindo no mercado e no meio acadêmico, e vêm apresentando resultados promissores.

As tecnologias provenientes do sensoriamento remoto estão assumindo espaços para tomadas de decisões e planejamento em diversas áreas do conhecimento. A combinação de tecnologias como o sensoriamento remoto e sistemas de informações geográficas, tem garantido papel importante para diversas análises e monitoramento em ambientes agrícolas e florestais.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABER, J.D. e MELILLO, J.M. **Terrestrial ecosystems**. San Diego: Academic Press. 2001.
- ADAMI, M.; GOMES, A. R.; COUTINHO, A. C. **Dinâmica do uso e cobertura da terra no estado do Pará entre os anos de 2008 a 2012**. Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR. Anais...João Pessoa -PB: Instituto Nacional de Pesquisas Espaciais, 2015.
- ALIXANDRINI, M. **Classificação de imagens Landsat 7 ETM multiespectrais utilizando SVM**. (INPE, Ed.)Anais do XIV Simpósio Brasileiro de Sensoriamento Remoto. Anais...Natal - BR: Instituto Nacionalde Pesquisas Espaciais, 2009.
- ALVES, M. O. et al. **O Uso de Imagens do “ Drone Agrícola ” para Identificação de Falhas no Plantio de Cana-de- Açúcar**. VII Simpósio Tecnologia de Produção de Cana-de-Açúcar. p.1-4. Piracicaba-SP. 2015.
- ALVARES et al. **Uso de sensoriamento remoto na estimativa do índice de área foliar em Eucalyptus**. XVII Simpósio Brasileiro de Sensoriamento Remoto. p. 6429-6436. 2015.
- ANDRADE, A. C.; FRANCISCO, C. N.; ALMEIDA, C. M. DE. Desempenho De Classificadores Paramétrico e Não Paramétrico na Classificação da Fisionomia Vegetal. **Revista Brasileira de Cartografia**, v. 66, n. 2, p. 349–363, 2014.
- ARGÜELLO, F. V. P. et al. Distribuição espacial de plantios de eucalipto no trecho paulista da bacia hidrográfica rio Paraíba do Sul, SP, Brasil. **Revista Ambiente e Agua**, v. 5, n. 3, p. 77–90, 2010.
- ASSAD, E. D.; SANO, E. E. **Sistemas de Informações Geográficas: aplicação na agricultura**. Brasília Embrapa, 1998. Disponível em:
[<http://livimagens.sct.embrapa.br/amostras/00064290.pdf>](http://livimagens.sct.embrapa.br/amostras/00064290.pdf)
- BLASCHKE, T.; KUX, H. **Sensoriamento remoto e SIG avançados: novos sistemas e sensores inovadores**. 2^a Edição ed. São Paulo: Oficina de Textos, 2007.
- BOERNER, W. M. et al. On the Basic Principles of Radar Polarimetry: the Target Characteristic Polarization State Theory of Kennaugh, Huynen's Polarization Fork Concept, and its Extension to the Partially Polarized Case. **IEEE Transactions on Geoscience and Remote Sensing**, v. 79, n. October, p. 1538–1550, 1991.
- BOYD, D. S. e DANSON, F. M. Satellite remote sensing of forest resources: three decades of research development. **Progress in Physical Geography**, v. 29, n. 1, p. 1–26, 2005.
- BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.
- BRYS, L. M. **Página Dinâmica para Aprendizado Do Sensoriamento Remoto**. Porto Alegre: Universidade Federal do Rio Grande do Sul (Mestrado em Sensoriamento Remoto e Meteorologia), 2008.

BURGES, C. J. C. A Tutorial on Support Vector Machine for Pattern Recognition. **Data Mining and Knowledge Discovery**, v. 2, n. 2, p. 121–167, 1998.

CADENA, G. T. **Clasificación de Tipos de Pavimentos de Vías Urbanas a Partir de Imágenes de Alta Resolución Espacial por Medio de Análisis Orientada a Objetos**. (Dissertação apresentada ao Programa de Pós-Graduação em Ciências Cartográficas da Universidade Estadual Paulista), p. 114, 2011.

CAMPBELL, J. B. **Introduction to remote sensing**. Ed. T. Guilford, New York, 1996.

CARUANA, R.; KARAMPATZIAKIS, N.; YESSENALINA, A. **An Empirical Evaluation of Supervised Learning in High Dimensions**(Proceedings of the 25th International Conference on Machine Learning). Finland. 2008.

CHANDER, G.; MARKHAM, B. L.; HELDER, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. **Remote Sensing of Environment**, v. 113, n. 5, p. 893–903, 2009.

CHAVEZ, P. S. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. **Remote Sensing of Environment**, v. 24, n. 3, p. 459–479, 1988.

CHAVEZ JR, P. S. Radiometric calibration of Landsat Thematic Mapper multispectral images. **Photogrammetric Engineering & Remote Sensing**, v. 55, n. September 1989, p. 1285–1294, 1989.

CHEN, B. et al. Mapping tropical forests and deciduous rubber plantations in Hainan Island, China by integrating PALSAR 25-m and multi-temporal Landsat images. **International Journal of Applied Earth Observation and Geoinformation**, v. 50, p. 117–130, 2016.

CLOUDE, S. R.; POTTIER, E. An entropy based classification scheme for land applications of polarimetric SAR. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 1, p. 68–78, 1997.

COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37–46, 1960.

CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: Principles and practices**. Lewis Publ ed. New York: Lewis Publishers, 1999.

CSRE (CENTRE OF STUDIES IN RESOURCES ENGINEERING). **Module 7 - Multi-looking**. Disponível em: <http://www.csre.iitb.ac.in/~ysrao/ankit/module_7/module_7.pdf>. Acesso em: 30 maio. 2017.

DINIZ, F. A.; NETO, F. M. M.; FONTES, L. M. DE O. RedFace: um sistema de reconhecimento facial baseado em técnicas de análise de componentes principais e autofaces: comparação com diferentes classificadores. **Revista Brasileira de Computação Aplicada**, n. 1, p. 2176–6649, 2013.

DÓRIA, V. E. M.; AMARAL, S.; MONTEIRO, A. M. V. Avaliação do potencial das imagens de luzes noturnas para estativa e distribuição espacial de população urbana: um estudo para a região metropolitana de São Paulo, Brasil, com o sensor **Visible Infrared Imaging radiometer Suite (VIIRS)**. XIX Encontro Nacional de Estudos Populacionais - São Pedro/SP, p. 21, 2014.

DUTRA, L. V. et al. **Processamento de Imagens de Radar de Abertura Sintética – Princípios e Aplicações**. Anais do IV Workshop em Tratamento de Imagens. Anais...Minas Gerais, 2003.

EPIPHANIO, J. C. N. et al. Índices de vegetação no sensoriamento remoto da cultura do feijão. **Pesquisa Agropecuária Brasileira**, v. 31, n. 6, p. 445–454, 1996.

FELDE, G. W. et al. Analysis of Hyperion data with the FLAASH atmospheric correction algorithm. **IEEE International Geoscience and Remote Sensing Symposium. Proceedings**, v. 1, n. C, p. 90–92, 2003.

FIGUEIREDO, D. **Conceitos Básicos de Sensoriamento Remoto** Companhia Nacional de Abastecimento-CONAB. Brasília - DF.2005 Disponível em:
[<http://www.ufscar.br/~debe/geo/paginas/tutoriais/pdf/sensoriamento/Conceitos_de_Sensoriamento_Remoto.pdf>](http://www.ufscar.br/~debe/geo/paginas/tutoriais/pdf/sensoriamento/Conceitos_de_Sensoriamento_Remoto.pdf).

FLORENZANO, TERESA GALLOTTI. Imagens de Satélite para Estudos Ambientais. **Oficina de Textos**, p. 1–97, 2002.

FOODY, G. M.; MATHUR, A. A relative evaluation of multiclass image classification by support vector machine. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, n. 6, p. 1335–1343, 2004.

FREITAS, C. DA C. et al. **Utilização De Imagens De Radar De Abertura Sintética Na Classificação De Uso E Ocupação Do Solo**. São José dos Campos, 2007.

FREEMAN, A. e DURDEN, S. A three-component scattering model for polarimetric SAR data. **IEEE Transactions on Geoscience and Remote Sensing**. v.36, p. 963-973, 1998.

GAO, T. et al. Timber production assessment of a plantation forest: An integrated framework with field-based inventory, multi-source remote sensing data and forest management history. **International Journal of Applied Earth Observation and Geoinformation**, v. 52, p. 155–165, 2016.

GAGNON, L.; JOUAN, A. Speckle filtering of SAR images: a comparative study between complex-wavelet-based and standard filters. **Proceedings of SPIE**. v. 3169. p. 80-91. 1997.

GOODMAN, J. W. Some fundamental properties of speckle*. **Journal of the Optical Society of America**, v. 66, n. 11, p. 1145, 1976.

GOU, S.; LI, X.; YANG, X. Coastal zone classification with fully polarimetric SAR imagery. **IEEE Geoscience and Remote Sensing Letters**, v. 13, n. 11, p. 1616–1620, 2016.

HENDERSON, F. .; LEWIS, A. J. **Manual of remote sensing: principles and applications of imaging radars.** 3. ed. New York: John Wiley and Sons, 1998.

HUANG, C.; DAVIS, L. S.; TOWNSHEN, J. R. G. An assessment of support vector machine for land cover classification. **International Journal of Remote Sensing**, v. 23, n. 4, p. 725–749, 2002.

HYYPPA, J. et al. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. **Forest Ecology and Management**, v. 128, n. 1–2, p. 109–120, 2000.

JANOTH, J. et al. **Procedimentos Baseados em Segmentação para a Análise de Mudanças e Classificação Florestais com Dados de Satélite de Alta Resolução.** cap. 9. In: BLASCHKE, T.; KUX, H. Sensoriamento Remoto e SIG Avançados. 2. ed.- São Paulo: Oficina de textos. 2007.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres.** São José dos Campos: Parêntese, 2009.

KASISCHKE, E. S.; MELACK, J. M.; DOBSON, M. C. The Use of Imaging Radars for Ecological Applications A Review. **Remote Sens Environ**, v. 4257, n. 96, p. 141–156, 1997.

KUMAR, S. **Advances in SAR Polarimetry.** (P. Tutorial, Ed.)SPIE APRS symposium. **Anais...pre-symposium tutorial**, 2016.

KUNCHEVA, L. I. **Combining pattern classifiers: methods and algorithms.** New Jersey: John Wiley & Sons, 2004.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **International Biometrics Society**, v. 33, n. 1, p. 159–174, 1977.

LAVREAU, J. De-Hazing Landsat Thematic Mapper Images. **Photogrammetric Engineering and Remote Sensing**, v. 57, n. 10, p. 1297–1302, 1991.

LE TOAN, T. et al. Relating Forest Biomass to SAR Data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 30, n. 2, p. 403–411, 1992.

LECKIE, D. G. e RANSON, K. J. Principles and Applications of Imaging Radar. In: SONS, J. W. AND (Ed.). **Principles and Applications of Imaging Radar.** New York: John Wiley and Sons. p. 435–509, 1998.

LEE, J. A simple speckle Smoothing algorithm for synthetic aperture radar images. **IEEE Transactions on Systems, Man and Cybernetics**, n. 1, p. 85–89, 1983.

LEE, J. SEN; GRUNES, M. R.; POTTIER, E. Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 11, p. 2343–2351, 2001.

LEE, J. SEN; POTTIER, E. **Polarimetric Radar Imaging: From Basics to Applications.** Taylor & F ed. New York: CRC Press : Taylor & Francis Group, 2009.

- LILLESAND, T.; KIEFER, R. **Remote Sensing and Image Interpretation**. 2008.
- LILLESAND, T. M.; KIEFER, R. W.; CHIPMAN, J. W. **Remote sensing and image interpretation**. New York: John Wiley and Sons, 2004.
- LIM, H. S. et al. **High Spatial Resolution Land Cover Mapping Using ALOS Data Over Kedah, Malaysia**. First Joint PI Symposium of ALOS DataNodes for ALOS Science Program Kyoto. *Anais...*2007.
- LORENA, A. C.; DE CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machine. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.
- LUIZ, A. J. B.; EPIPHANIO, J. C. N. **Amostragem por pontos em imagens de sensoriamento remoto para estimativa de área plantada por município**. X Simpósio Brasileiro de Sensoriamento Remoto. *Anais...*Foz do Igualu- PR: 2001.
- MAIN, R. et al. Hyper-temporal C-band SAR for baseline woody structural assessments in deciduous savannas. **Remote Sensing**, v. 8, n. 8, p. 1–19, 2016.
- MARTINS, F. DA S. R. V. et al. Sensitivity of ALOS/PALSAR imagery to forest degradation by fire in northern Amazon. **International Journal of Applied Earth Observation and Geoinformation**, v. 49, p. 163–174, 2016.
- MELO, D. H. C. T. B. **Uso de Dados Ikonos II na Análise Urbana : Testes Operacionais na Zona Leste de São Paulo**. (Dissertação de Mestrado em Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais, São José dos Campos. p.146, 2002.
- MENESES, P. R.; ALMEIDA, T. DE. **Introdução ao Processamento de Imagens de Sensoriamento Remoto**. p. 266, 2012.
- MENESES, P. R.; MADEIRA NETTO, J. S. **Sensoriamento Remoto: Reflectância dos alvos naturais**. Brasília - DF: Embrapa Cerrados, 2001.
- MISHRA, V. N. et al. Dual-polarimetric C-band SAR data for land use/land cover classification by incorporating textural information. **Environmental Earth Sciences**, v. 76, n. 1, p. 1–16, 2017.
- MITCHARD, E. T. A. et al. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: Overcoming problems of high biomass and persistent cloud. **Biogeosciences**, v. 9, n. 1, p. 179–191, 2012.
- MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. 4^a ed. Viçosa- MG: Editora UFV, 2011.
- MUTHUKUMARASAMY, I.; SHANMUGAM, R. S.; KOLANUVADA, S. R. SAR polarimetric decomposition with ALOS PALSAR-1 for agricultural land and other land use/cover classification: case study in Rajasthan, India. **Environmental Earth Sciences**, v. 76, n. 13, p. 1–13, 2017.

NASCIMENTO, R. F. F. et al. **O algoritmo Support Vector Machine (SVM): avaliação da separação ótima de classes em imagens CCD-CBERS-2.** (INPE, Ed.) Anais XIV Simpósio Brasileiro de Sensoriamento Remoto. Anais...Natal - BR: 2009.

NAVULUR, K. **Multiespectral Image Analysis Using the Object-Oriented Paradigm.** Taylor & F. ed. Broken Sound Parkway, NW. p. 206. 2006.

NETO, C. D. G. **Potencial de técnicas de mineração de dados para o mapeamento de áreas cafeeiras.** (Monografia referente a disciplina Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais. São José dos Campos, p.50, 2014.

NOVO, E. M. L. DE M. **Sensoriamento Remoto: Princípios e Aplicações.** São Paulo: Edgard Blucher, 2010.

NUNES, G. M. **Sensoriamento remoto aplicado na análise da cobertura vegetal das Reservas de Desenvolvimento Sustentável Amanã e Mamirauá.** Simpósio de Pós-Graduação em Engenharia Urbana. Universidade Estadual de Campinas, 2008.

OLIVEIRA, F. S. DE. **Uso De Imagens Do Satélite Alos Para Estimativa De Parâmetros Dendométricos de Plantios De Eucalipto.** Viçosa- MG: Universidade Federal de Viçosa (Tese Doutorado). p.146, 2012.

OLIVER,C.J e QUEGAN. Understanding Synthetic Aperture Radar Images. **Artech House.** p.480,1998.

PAL, M.; MATHER, P. M. Support vector machine for classification in remote sensing. **International Journal of Remote Sensing**, v. 26, n. 5, p. 1007–1011, 2005.

PARIZZI, A.; BRCIC, R. Adaptive InSAR stack multilooking exploiting amplitude statistics: A comparison between different techniques and practical results. **IEEE Geoscience and Remote Sensing Letters**, v. 8, n. 3, p. 441–445, 2011.

PAULA JÚNIOR, I. C. DE. **Capítulo 5: Técnicas de Processamento Digital de Imagens com Java.** Disponível em: <<http://docplayer.com.br/7926431-Tecnicas-de-processamento-digital-de-imagens-com-java.html>>. Acesso em: 20 jul. 2017.

PEREIRA, L. DE O. **Avaliação de Métodos de Integração de Imagens Ópticas e de Radar para Classificação do Uso e Cobertura da Terra na Região Amazônica.** Dissertação (Mestrado em Sensoriamento Remoto)- Instituto Nacional de Pesquisa Espaciais, São José dos Campos, 2012.

PIAZZA, G. A. **Processamento Digital de Imagens de Alta Resolução Espacial com Enfoque na Classificação dos Estágios Sucessionais Iniciais da Floresta Ombrófila Densa em Santa Catarina.** Dissertação (Mestrado em Engenharia Ambiental)- Centro de Ciências Tecnológicas . Fundação Universidade Regional de Blumenau , Blumenau, 2014.

PICOLI, M. C. A. **Utilização de dados palsar/als (banda I) no monitoramento da cana-de-açúcar na Banda L.** Tese (Doutorado em Engenharia Agricola. Universidade Estadual de Campinas. p. 109, 2012.

- POHL, C.; GENDEREN, J. L. VAN. Multisensor image fusion in remote sensing : Concepts , methods and applications. **Remote Sensing**, v. 19, p. 823–854, 1998.
- POLIKAR, R. Ensemble based systems in decision making. **Circuits and Systems Magazine, IEEE**, v. 6, n. 3, p. 21–45, 2006.
- PONZONI, F. J. et al. **Calibração de Sensores Orbitais**. Oficina de ed. São Paulo: Oficina de Textos, 2015.
- PONZONI, F. J.; SHIMABUKURO, Y. E. **Sensoriamento Remoto no estudo da vegetação**. 1. ed. Parêntese . São José dos Campos, p. 144, 2007.
- PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento Remoto da Vegetação**. 2. ed. São Paulo: Oficina De Texto, 2012.
- POPE, K. O.; REY-BENAYAS, J. M.; PARIS, J. F. Radar remote sensing of forest and wetland ecosystems in the Central American Tropics. **Remote Sensing of Environment**, v. 48, n. 2, p. 205–219, 1994.
- POWERS, D. M. W. **The problem with kappa**. Conference of the European Chapter of the Association for Computational Linguistics. **Anais...**França: Association for Computational Linguistics, 2012
- PRASAD, A. M.; IVERSON, L. R.; LIAW, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. **Ecosystems**, v. 9, n. 2, p. 181–199, 2006.
- RAMESH, K. N. et al. Detection of Rows in Agricultural Crop Images Acquired by Remote Sensing from a UAV. **International Journal of Image, Graphics and Signal Processing**, v. 8, n. 11, p. 25–31, 2016.
- RANSON, K. et al. Characterization of forests in Western Sayani mountains, Siberia from SIR-C SAR data. **Remote Sensing of Environment**. v.75.p.188-200. 2001
- ROSENFIELD, G. H.; FITZPATRICK-LINS, K. A coefficient of agreement as a measure of thematic classification accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 52, n. 1979, p. 223–227, 1986.
- ROUSE, J. W. et al. **Monitoring vegetation systems in the great plains with ERTS**. Third Earth Resources Technology Satellite (ERTS) symposium. **Anais...**Washington: 1973. Disponível em: <<https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf>>
- SAATCHI, S. S. et al. **Estimation of Forest Biomass Change from Fusin of Radar and Lidar Measurements.2010**. Disponível em: <<https://pt.slideshare.net/grssieee/estimation-of-forest-biomass>>. Acesso em: 1 set. 2017.
- SANO, E. E.; MARRA, E.; MENESSES, P. R. Análise de imagens do satélite Alos Palsar para o mapeamento de uso e cobertura da terra do Distrito Federal. **Geociências**, v. 28, p. 441–451, 2009.

SANT' ANNA, S. J. S. Comparação do desempenho de filtros redutores de “speckle”.
Dissertação (Mestrado em Sensoriamento Remoto/ Processamento de Imagens)- Instituto Nacional de Pesquisa Espaciais, p. 2009, 1995.

SARMAP. SARscape®. Disponível em:
[<http://www.sarmap.ch/pdf/SARscapeTechnical.pdf>](http://www.sarmap.ch/pdf/SARscapeTechnical.pdf). Acesso em: 30 maio. 2017.

SEMOLINI, R. Support Vector Machine, Inferência Transdutiva e o Problema de Classificação. Campinas -SP. 2002.

SHIMABUKURO, Y. E.; MAEDA, E. E.; FORMAGGIO, A. R. Sensoriamento Remoto e Sistemas de Informações Geográficas aplicados ao estudo dos recursos agronômicos e florestais. **Revista Ceres**, v. 56, n. 4, p. 399–409, 2009.

SINCLAIR, G. The Transmission and Reception of Elliptically Polarized Waves. **Proceedings of the IRE**, v. 38, n. 2, p. 148–151, 1950.

SIQUEIRA, R. V. **Estimativa da Temperatura de Superfície na Região Metropolitana de Goiânia Por Meio de Imagens Landsat e Previsão De Temperaturas para Periodos Posteriores.** Programa de Pós- graduação Stricto Sensu em Engenharia do Meio Ambiente da Escola de Engenharia Civil. Universidade Federal de Goiás. p. 1110, 2015.

SONG, C. et al. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? **Remote Sensing of Environment**, v. 75, n. 2, p. 230–244, 2001.

STEFFEN, A. C. **Introdução ao Sensoriamento Remoto.** Disponível em:
[<http://www.inpe.br/unidades/cep/atividadescep/educasere/apostila.htm>](http://www.inpe.br/unidades/cep/atividadescep/educasere/apostila.htm). Acesso em: 2 nov. 2016.

STEINWART, I.; CHRISTMANN, A. **Support Vector Machine.** New York: Springer-Verlag, 2008.

TOUTIN, T. Review article: Geometric processing of remote sensing images: models, algorithms and methods. **International Journal of Remote Sensing**, v. 25, n. 10, p. 1893–1924, 2004.

TSO, B., MATHER, P. **Classification Methods for Remotely Sensed Data**, 2nd ed. CRC Press, p. 376. 2009.

ULABY, F.; ELACHI, C. **Radar Polarimetry for Geoscience Applications.** Artech Hou ed. 1990.

ULABY, F. T. Radar Response to Vegetation. **IEEE Transactions on Antennas and Propagation**, v. 32, n. 1, p. 36–45, 1975.

VAPNIK, V. N. **The nature of statistical learning theory.** Springer- ed. New York. v. 8, 1995

VERMOTE, E. et al. Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An Overview. **IEEE TR**, v. 35, n. 3, p. 675–686, 1997.

VIET NGUYEN, L. et al. Estimation of Tropical Forest Structural Characteristics Using ALOS-2 SAR Data. **Advances in Remote Sensing**, v. 5, n. 2, p. 131–144, 2016a.

VIET NGUYEN, L. et al. Mapping Tropical Forest Biomass by Combining ALOS-2, Landsat 8, and Field Plots Data. **Land**, v. 5, n. 4, p. 31, 2016b.

WARING, R. H. et al. Imaging radar for ecosystem studies. **American Institute of Biological Sciences**, v. 45, n. 10, p. 715–723, 1995.

WASKE, B. et al. imageRF – A user-oriented implementation for remote sensing image analysis with Random Forests. **Environmental Modelling & Software**, v. 35, p. 192–193, 2012.

WATZLAWICK, L. F. **Estimativa de biomassa e carbono em floresta ombrófila mista e plantações florestais a partir de dados de imagens do satélite Ikonos II**. Curitiba: UFPR, Tese (Doutorado em Ciências Florestais) – Setor de Ciências Agrárias, Universidade Federal do Paraná.p.120, 2003.

CAPÍTULO 2

Avaliação de algoritmos paramétricos e não-paramétricos para classificação de uso e cobertura da terra no Planalto Sul Catarinense

RESUMO

O objetivo deste trabalho foi avaliar a performance das classificações supervisionadas oriundas dos dados ópticos (Landsat-5/TM) em diferentes tamanhos amostrais, ou seja, proporção entre dados de treinamento e validação. Foram avaliados 10% para fins de treinamento e 90% de validação, e ainda, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10, respectivamente. A área de estudo está localizada em uma área rural no planalto sul catarinense. Foram consideradas seis classes de uso e cobertura da terra (i.e. Água, Solo Exposto, Área Urbanizada, Campo Nativo, Floresta Nativa e Reflorestamento). Cada classe foi constituída por mil pixels, totalizando assim 6.000 pixels. Os métodos de classificação supervisados utilizados foram *Support Vector Machine* (SVM), *Random Forest* (RF) e *Maximum Likelihood* (ML). A qualidade dos mapas gerados foi avaliada usando a acurácia global e o índice Kappa. As significâncias dos mapas foram avaliadas por meio do teste Z. Os resultados obtidos mostram que o conjunto de dados 90/10, 60/40 e 40/60 obtiveram os melhores índices Kappa, em ordem para RF, SVM e ML, respectivamente. Houve diferenças significativas entre os algoritmos de classificação empregados. Apenas o conjunto 50/50 foram significativos para RF e SVM com teste Z. O classificador ML, ou seja, classificador paramétrico, assemelhou-se mais com o classificador SVM (não-paramétrico). Assim, para as condições analisadas, destaca-se que RF foi o classificador que requereu o menor conjunto de validação. Por outro lado, SVM e ML foram os algoritmos de classificação que requereram ao menos que 40% das amostras fossem destinadas a validação. Portanto, é recomendado a utilização de classificador não-paramétricos para realizações de classificação de uso e cobertura da terra em cenas de sensores ópticos, quais são comumente utilizados algoritmos paramétricos.

Palavras-chave: Sensoriamento Remoto, Amostras de treinamento e Validação, Classificação pixel a pixel, Algoritmos paramétricos e não-paramétricos.

CHAPTER 2

Evaluation of Parametric and Non-Parametric Algorithms for Classification of Land Use and Land Cover in the Santa Santa Catarina State Southern Plateau

ABSTRACT

The main goal of this study was to evaluate the performance of supervised classification of optical data (Landsat-5/TM) exploring different training and validation data size. The study area is in a rural area of the Santa Catarina State Southern Plateau. We consider six land use and land cover classes (LULC) (i.e. Water, bare soil, urban area, grasslands, native forest and planted forest). Each LULC class has 1,000pixels, totaling 6,000 pixels. We evaluate 10% of the pixels for the training dataset, whereas the remaining 90% for the validation dataset. Thus, we explore also the proportions of 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10, respectivelly. The choosen supervised classification methods were Support Vector Machine (SVM), Random Forest (RF) and Maximum Likelihood (ML). The quality of the generated maps was evaluated using both overall accuracy and Kappa index. The significance level was evaluated using Z-test. The results show that the datasets 90/10, 60/40 and 40/60 reached the highest Kappa indices for RF, SVM and ML, respectively. We found significative differences among the methods, although the 50/50 dataset was significant for both RF and SVM methods. The performance using ML was similar to those reached by SVM. RF was the classifier that required the smallest validation dataset. On the other hand, SVM and ML required at least 40% of the samples for validation. Therefore, we recommend the use of non-parametric classification methods such as RF for the LULC classification purposes using optical scenes.

Keywords: Remote Sensing; Training Samples and Validation; Pixel-to-pixel Classification; Parametric and Non-parametric Algorithms.

2.1. INTRODUÇÃO

Figueiredo (2005) afirmou que um dos princípios do Sensoriamento Remoto é a geração/produção de mapas temáticos, tendo como a finalidade específica de quantificar e/ou qualificar o uso e a cobertura da terra de um determinado local. Jensen (2009) afirmou que os dados de sensoriamento remoto podem ser usados para estudar diferentes variáveis ambientais (i.e. Solo, água, vegetação e etc.). Para a geração de mapas temáticos é comum o uso de sensores ópticos. Estes sensores estão limitados pela região do espectro eletromagnético óptico (EEM) que abrange o intervalo de 400nm a 2400nm (MENESES e ALMEIDA, 2012).

Uma boa classificação requer uma boa confiabilidade dos dados amostrais. As amostras podem ser coletadas em campo por meio de receptor GNSS ou diretamente sobre imagens orbitais ou aerotransportadas. Estas em geral, de maior resolução espacial aquela que se deseja classificar. Ainda, faz-se necessário o conhecimento do sensor quanto as suas diferentes resoluções (i.e., resolução espectral, espacial, radiométrica e temporal) (MELO, 2002; NOVO, 2010; STEFFEN, 2017). Toutin (2004) afirmou que cenas de sensores orbitais, apresentam distorções, e estas têm que serem minimizadas, pois podem interferir no produto final. Técnicas de pré-processamento tendem a minimizar duas grandes categorias de distorções como os observados (atmosfera e Terra) ou referente ao sistema de aquisição (plataforma, sensor e outros instrumentos de medições). Sendo assim, técnicas de pré-processamento são indispensáveis para não comprometer o resultado final.

Após as diferentes etapas de pré-processamento, pode-se fazer uso das imagens, a exemplo da realização de classificações. Estas permitem a extração de informações que podem ser lidas em formato de tabelas, gráficos e mapas (FIGUEIREDO, 2005). De acordo com Meneses e Almeida (2012) há diversos métodos de classificação de imagens. Estes podem ser divididos em: classificação paramétrica e não-paramétrica, classificação espectral e espacial, e classificação supervisionada ou não-supervisionada. Existe até a possibilidade de decompor os classificadores por pixel, por regiões ou por objeto.

A utilização do melhor classificador para o mapeamento do uso da terra, realizando assim mapas temáticos e métricas qualitativas, é de fundamental importância para direcionar um bom planejamento. Consequentemente também para a definição de prioridades por parte do poder público e privado envolvidos na gestão dos recursos naturais.

Além do classificador, o tamanho de amostras de entrada e validação para o mapeamento, podem apresentar diferenças, assim permitindo superestimar ou subestimar os valores qualitativos de uma classificação. Deste modo, Dicks e Lo (1990) afirmaram que quanto

maior o tamanho da amostra, maior será a confiabilidade nas avaliações. Entretanto, a obtenção de amostras em campo envolve recursos financeiros com o pessoal de apoio, bem como o tempo requerido para a coleta. Ainda, não existe na literatura um padrão de amostra ideal.

Portanto, o objetivo geral deste capítulo foi avaliar a performance dos classificadores supervisionados amplamente utilizados na literatura, a saber, *Random Forest* (RF), *Support Vector Machine* (SVM) e *Maximum Likelihood* (ML) para fins de classificação de diferentes classes de uso e cobertura da terra. Foram avaliados para cada um dos algoritmos de classificação diferentes proporções de amostras de treinamento e validação. As análises foram desenvolvidas sobre dados do sensor TM/Landsat-5.

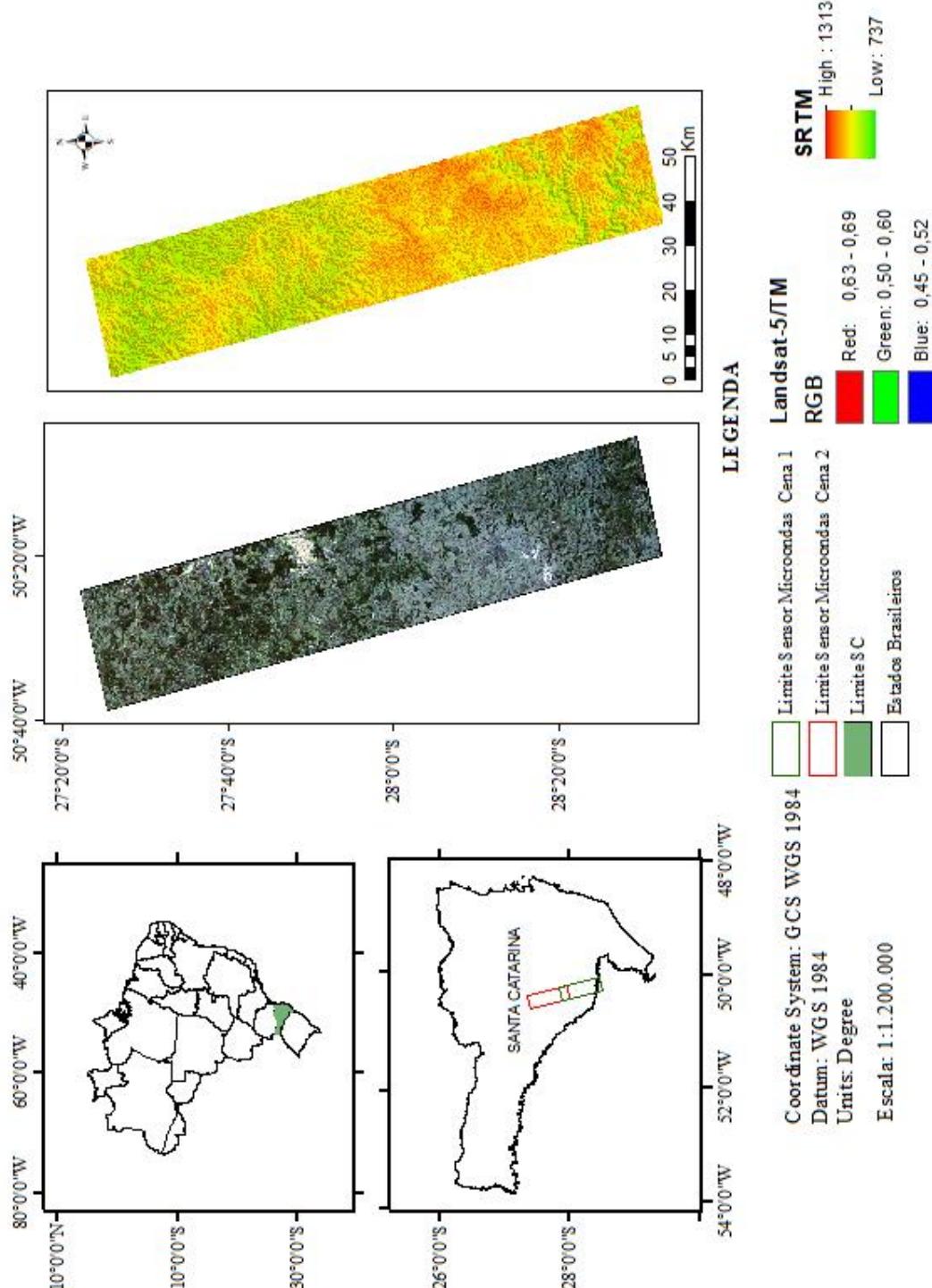
2.2. MATERIAL E MÉTODOS

2.2.1. Descrição da área de estudo

O estudo foi desenvolvido em sete municípios do planalto sul catarinense (Lages, Curitibanos, Capão Alto, São José do Cerrito, Correia Pinto, Ponte Alta e São Joaquim). A região de estudo está localizada na longitude $50^{\circ} 39' 40,61''$ O. e latitude $28^{\circ} 28' 59,52''$ S., abrangendo uma extensão territorial de 3.085,6612 km² (Figura 2.1).

Segundo Alvares et al. (2013) o clima é classificado como Cfb (temperado subtropical). Estas áreas possuem um forte perfil agrícola, somando-se a expressividade na produção florestal, fator decisivo para alavancagem e consolidação dos segmentos de celulose e papel, madeireiro e moveleiro da macrorregião (ABRAF, 2013).

Figura 2.1. Localização da área de estudo.



Fonte: O autor (2017).

2.2.2. Pré-processamento de imagens

No presente estudo, foram obtidas cenas do satélite TM/Landsat-5, o qual é administrado pela NASA e gerenciado pelo *U.S. Geological Survey* (USGS's), no *Center for Earth Resources Observation and Science*. Embora este sensor esteja desativado desde 2013, os dados históricos da série Landsat se constituem como uma importante fonte de dados para a observação da terra. As características do Landsat-5 são apresentadas na Tabela 2.1. Foram utilizadas duas cenas da órbita 221, referente ao dia 20 de março de 2009, sendo uma do ponto 79 e outra do 80. As cenas foram adquiridas pelo Earth Explorer (site: www.earthexplorer.usgs.gov).

Tabela 2.1. Características espetrais e espacial do satélite Landsat 5.

Sensor	Região do Espectro	Intervalo (μm)	Res. Espacial (m)
TM (Thematic Mapper)	Azul	0,45 -0,52	30
	Verde	0,50 -0,60	
	Vermelho	0,63-0,69	
	Infravermelho Próximo	0,76-0,90	
	Infravermelho Médio	1,55 -1,75	
	Infravermelho Termal	10,4-12,5	120
	Infravermelho Médio	2,06-2,35	30

Fonte: Novo (2010).

Foram realizadas diversas etapas de pré-processamento das cenas. Portanto, realizou-se a conversão de níveis digitais para radiância ($\text{Wm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$) e posteriormente a conversão para reflectância de superfície, considerando a minimização dos efeitos atmosféricos através do “*fast line of sight atmospheric analysis of spectral hypercubes*” (Flaash) proposto por Felde et al. (2003).

A cena do Landsat-5 apresenta 185x185 km e para o presente estudo, foi realizado um recorte que envolve uma área coincidente com as cenas do sensor ALOS/PALSAR, qual é utilizado nos demais capítulos.

2.2.3. Classificações dos dados remotamente situados

Foi utilizado um aplicativo GIS para a coleta das amostras. Este procedimento foi realizado por meio de interpretação visual. Para a coleta dos dados, foi utilizado como

referência a cena do Landsat-5. Para averiguação dos resultados foram utilizadas cenas disponibilizadas pelo Google Earth Pro além de cenas ortorretificadas do levantamento aéreo, disponibilizadas pelo Estado de Santa Catarina. As amostras foram coletadas em formato *shapefile*. Foram coletadas seis classes de uso e cobertura da terra, conforme a descrição na Tabela 2.2.

Tabela 2.2. Descrição das Classes de Uso e Cobertura da terra.

Floresta Nativa	Áreas cobertas por vegetação de porte florestal nativa.
Reflorestamento	Áreas cobertas por vegetações de porte florestal plantado (reflorestamentos de eucalipto e pinus).
Solo Exposto	Áreas em que houve corte raso da vegetação natural e/ou colheita florestal e/ou agronômica.
Água	Áreas como rios, lagos, represas e entre outras.
Área Urbanizada	Áreas com feições e características de ocupações urbanas observadas nas imagens.
Campo Nativo	Áreas cobertas por vegetações rasteira nativa e/ou plantada (manejada).

Fonte: O autor (2017).

Para a determinação do tamanho de amostras de treinamento e validação, foram geradas amostras aleatórias simples, sendo eles 10% para treinamento e 90% para validação (10/90). Em outras palavras, de 1.000 pixels de cada classe, 100 pixels foram selecionados para treinamento e os 900 pixels remanescentes para validação. Portanto, foram geradas amostras para 10/90, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10, respectivamente.

Para a classificação das imagens foram utilizados os seguintes algoritmos de classificação, *Support Vector Machine* (SVM), *Random Forest* (RF) e *Maximum Likelihood* (ML). Maiores detalhes sobre os algoritmos podem ser encontrados em trabalhos como de Breiman (2001) que especifica o RF e Mountrakis, Im e Ogole (2011) que realizaram uma revisão sobre o SVM e sobre o *Maximum Likelihood*, Harris e Stöcker (1998) descreveram detalhadamente o método.

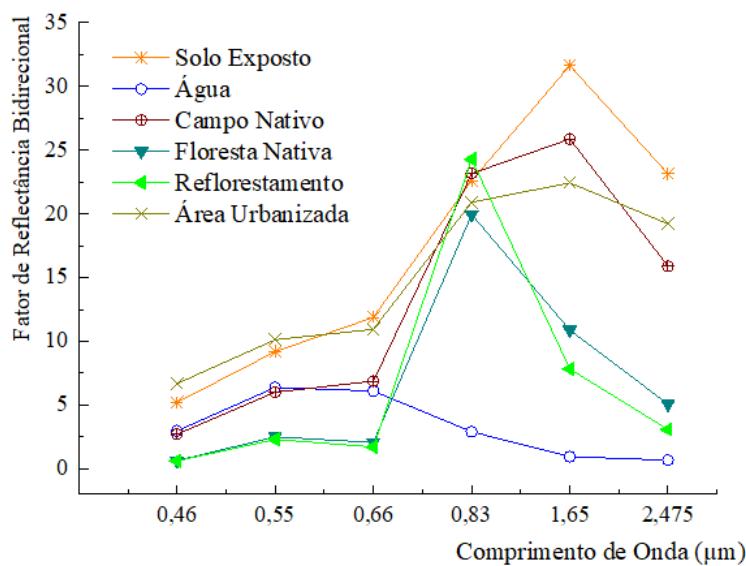
As classificações SVM e RF foram geradas no aplicativo EnMAPToolBox (VAN DER LINDEN et al., 2015) enquanto que ML no aplicativo ENVI (SULSOFT, 2017). Posteriormente, as avaliações de acurácia global foram realizadas para as classificações obtidas pelos diferentes algoritmos e conjuntos amostrais. As avaliações dos mapas temáticos foram realizadas por meio de matrizes de confusão e do índice Kappa, sugerido por Cohen (1960) e Congalton (1991). A significância das acuráncias obtidas pelos algoritmos de classificação e conjuntos amostrais foi realizada pelo Teste-Z, cujo valor crítico para 95% e 99% é de 1,96 e 2,57, respectivamente. A diferença entre os algoritmos de classificação, de maneira geral, foi

realizada no aplicativo Rstudio (RSTUDIO TEAM, 2016) aplicando o teste Tukey, visto que, para a análise de Variância (ANOVA) as repetições foram os valores do índice Kappa, e os tratamentos foram os algoritmos de classificação (i.e. RF, SVM e ML).

2.3. RESULTADOS E DISCUSSÃO

A seleção de 1.000 pixels amostrados para cada uma das seis classes de uso e cobertura da terra, para fins de geração de amostras de treinamento e validação, foram submetidas inicialmente a análise de perfis espectrais (Figura 2.2). As classes solo exposto, campo nativo e área urbanizada obtiveram os maiores valores de reflectância de superfície no infravermelho médio (1,55-1,75 μm), seguidos pela floresta nativa e reflorestamento no infravermelho próximo (0,76-0,90 μm). A classe água mostrou em geral baixos valores de reflectância de superfície, com os maiores valores ocorrendo na região do visível (400-690 nm) e decrescendo gradualmente para o infravermelho médio e próximo, respectivamente. Os perfis espectrais corroboram com os exemplos explanados para vários alvos por Ponzoni, Shimabukuro e Kuplich (2012).

Figura 2.2. Assinatura Espectral das classes amostradas para classificação de uso e Cobertura da terra.



Nota: Foram utilizados o conjunto total de dados.

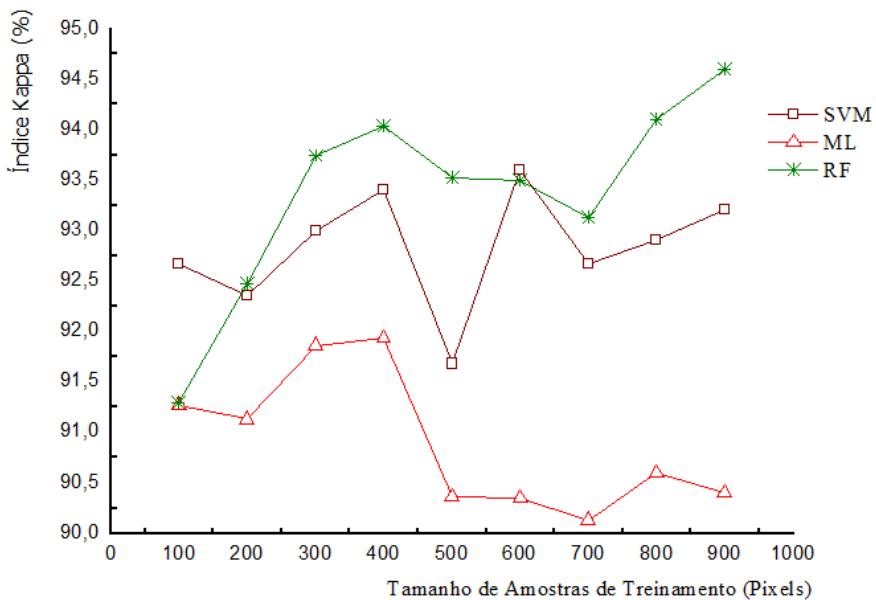
Fonte: O autor (2017).

De acordo com as respostas espectrais do presente estudo, pode se observar que os intervalos espectrais do infravermelho próximo e médio, evidenciaram uma melhor discriminação entre as diferentes classes selecionadas. Estes resultados são similares como na

pesquisa de Weiss et al. (2013), porém, encontraram dificuldades em distinguir duas classes na banda do infravermelho médio.

Os valores dos índices Kappa para os diferentes conjuntos amostrais, mostraram que o SVM e RF obtiveram um incremento do índice de acordo com o aumento das amostras de treinamento (Figura 2.3). De acordo com a categorização proposta por Landis e Koch (1977), todos os índices Kappa, para todos os algoritmos, resultaram em “Excelente”. Entretanto, o classificador ML apresentou os menores valores de índice Kappa.

Figura 2.3. Valores do Coeficiente Kappa e tamanho de amostras de treinamento/Validação.



Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

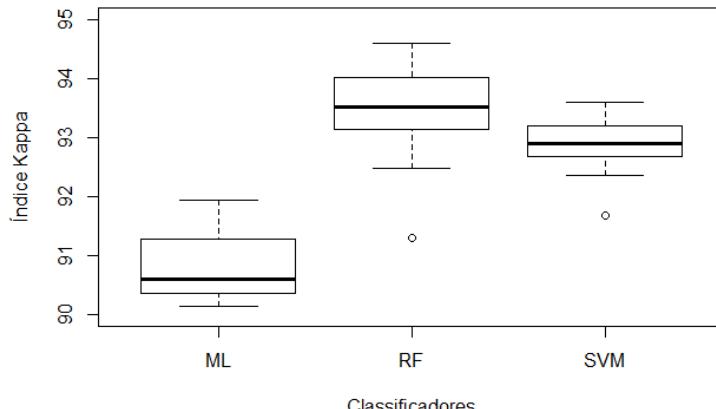
Fonte: O autor (2017).

Garofalo et al. (2015) pesquisaram o classificador SVM, segmentando áreas com 100 polígonos (ROI's), em cenas do Landsat-8 e constataram um índice Kappa de 92%. Logo, o presente estudo apresentou similaridade ao tamanho de amostra de 100 pixels para treinamento e 900 para validação (10/90). Gaiad et al. (2017) avaliaram algoritmos baseados em máquina de aprendizagem, e sem esboçar o valor do tamanho da amostra de treinamento e validação, chegaram a conclusão que o SVM foi melhor que o RF, diferenciando a esta pesquisa em que somente dois momentos, 10/90 e 60/40, o SVM apresentou índice maior que o RF.

Ao avaliar os algoritmos de classificação entre si, ou seja, algoritmos foram os tratamentos (i.e. SVM, RF e ML) e os índices Kappa (%) as repetições, foi possível notar que houve uma diferença significativa a 5% de probabilidade entre eles. A Figura 2.4 mostra os níveis máximos e mínimos, que evidência a amplitude interquartílica e a mediana dos índices

Kappa para cada classificador. Foi possível notar para o SVM e RF, apresentam valores atípicos (outliers), onde o conjunto amostral 10/90 para o RF e 50/50 para o SVM. Nos três algoritmos de classificação exibiram assimetria positiva.

Figura 2.4. Bloxplot- Relação índice Kappa com algoritmos de classificação.



Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

Como apresentado na Tabela 2.3. valores médios do índice Kappa em diferentes algoritmos de classificação, mostram que existem diferenças significativas de 5% entre os algoritmos. Entretanto, os classificadores não-paramétricos (SVM e RF) não apresentaram diferenças significativas, ou seja, foram similares entre si. Já o ML, um classificador paramétrico, comumente empregado em trabalhos envolvendo classificação de uso e cobertura da terra com sensores óptico, apresentou o menor valor médio, qual diferenciou estatisticamente dos demais algoritmos de classificação.

Tabela 2.3. Valores de Índice Kappa obtidos para os diferentes algoritmos de classificação.

Classificador	SVM	RF	ML
Média	92,83a	93,38a	90,89b

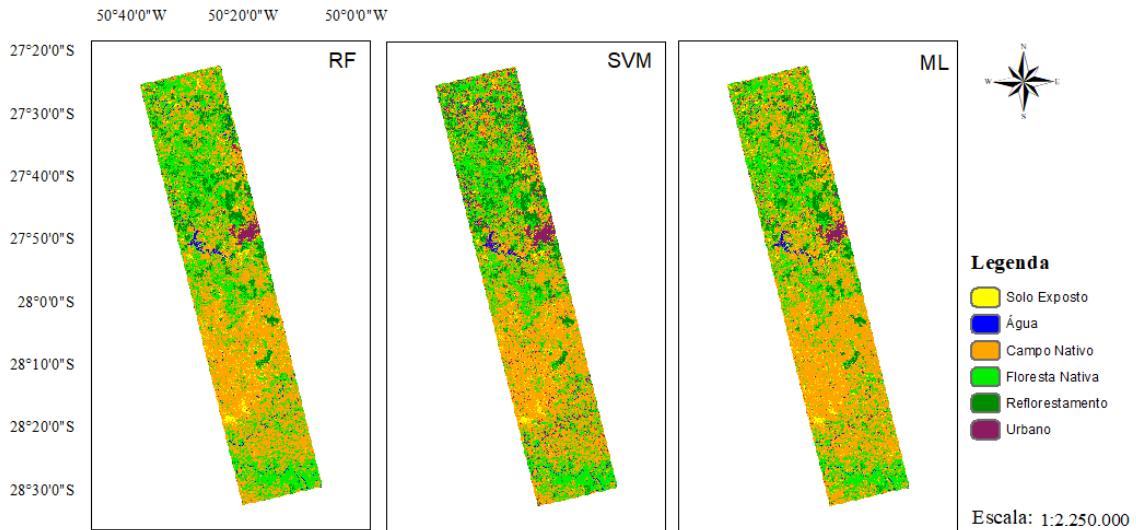
Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

Bolfe et al. (2004) avaliaram o classificador ML com o sensor TM/Landsat-5 para discriminação de florestas naturais, eucalipto e pinus, e encontraram valor de Kappa de 84,90%. Contudo, os autores não descreveram o tamanho da amostra de treinamento e validação. Todavia, o valor encontrado é “execente”, de acordo com a classificação qualitativa, ao reportado neste estudo. As classificações obtidas são apresentadas na Figura 2.5. Foi possível

notar, que para os três classificadores existe uma predominância para as classes floresta nativa e campo nativo. A representatividade territorial de cada classe é apresentada na Tabela 2.4.

Figura 2.5. Cenas Classificadas com diferentes algoritmos de classificação.



Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O Autor (2017).

Conforme apresentado na Tabela 2.4, os valores de área de cada classificador referente aos melhores índices Kappa, ou seja, para o RF amostra 90/10, para o SVM amostra 60/40 e para o ML amostra 40/60. Foi percebido diferenças nos valores das áreas classificadas como Área Urbanizada, onde o SVM apresentou o maior valor. O classificador ML apresentou menores valores que o classificador RF, entretanto, foi superior somente na classe Campo Nativo.

Tabela 2.4. Representatividade das diferentes classes de uso e cobertura da terra, para a área de estudo, considerando três distintos classificadores e conjunto amostral.

Classe/Classificador	RF 90/10 (km ²)	SVM 60/40(km ²)	ML 40/60(km ²)
Solo Exposto	111,30	93,60	134,59
Água	40,77	30,38	28,19
Campo Nativo	1594,14	1467,48	1676,28
Floresta Nativa	1038,35	998,56	910,06
Reflorestamento	231,04	219,55	221,83
Área Urbanizada	69,21	275,24	113,87

Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

A Tabela 2.5. com valores obtidos através do Teste-Z, demonstra que para o classificador RF existem diferenças significativas entre os tamanhos de amostras, principalmente quando comparado com o conjunto amostral de 10/90, entretanto, também houve diferenças entre os conjuntos amostrais de 20/80 com 30/70 e 40/60. De maneira geral, o classificador RF apresentou 27,77% dos dados significativos, sendo 19,44% para a 1% e 8,33% para 5% de significância.

Logo, para o SVM, houve diferenças entre os tamanhos de amostras 10/90 com 30/70 e 40/60 e entre os 50/50 com 40/60 e 60/40. Assim totalizando 13,88% de dados significativos, sendo 5,55% para 1% e 8,33% para 5%. Já para ML, foi possível notar que houve diferença para amostras de 50/50 com 30/70 e 40/60. Portanto, apenas 5,56 das amostras foram significativas, sendo 2,77% tanto para 1% e 5% de probabilidade de erro.

Tabela 2.5. Valores do Teste Z para diferentes conjuntos amostrais avaliados para os classificadores (a) Random Forest, (b) Support Vector Machine e (c) Maximum Likelihood.

(a)									
RF	10	20	30	40	50	60	70	80	90
10	-	1,988*	4,154**	4,521**	3,423**	3,169**	2,360*	3,267**	3,009**
20	-	-	2,156*	2,578**	1,610 NS	1,474 NS	0,844 NS	1,893 NS	1,934 NS
30	-	-	-	0,485 NS	0,346 NS	0,352 NS	0,786 NS	0,418 NS	0,782 NS
40	-	-	-	-	0,779 NS	0,757 NS	1,143 NS	0,077 NS	0,513 NS
50	-	-	-	-	-	0,027 NS	0,470 NS	0,646 NS	0,956 NS
60	-	-	-	-	-	-	0,427 NS	0,644 NS	0,951 NS
70	-	-	-	-	-	-	-	0,970 NS	1,213 NS
80	-	-	-	-	-	-	-	-	0,396 NS
90	-	-	-	-	-	-	-	-	-

(b)									
SVM	10	20	30	40	50	60	70	80	90
10	-	1,920 NS	2,960**	3,530**	0,690 NS	3,450*	1,830 NS	1,840 NS	1,650 NS
20	-	-	1,070 NS	1,690 NS	0,960 NS	1,800 NS	0,397 NS	0,598 NS	0,703 NS
30	-	-	-	0,636 NS	1,871 NS	0,856 NS	0,414 NS	0,108 NS	0,165 NS
40	-	-	-	-	2,392*	0,280 NS	0,898 NS	0,534 NS	0,164 NS
50	-	-	-	-	-	2,454*	1,126 NS	1,234 NS	1,205 NS
60	-	-	-	-	-	-	1,070 NS	0,710 NS	0,318 NS
70	-	-	-	-	-	-	-	0,220 NS	0,404 NS
80	-	-	-	-	-	-	-	-	0,215 NS
90	-	-	-	-	-	-	-	-	-

(c)									
ML	10	20	30	40	50	60	70	80	90
10	-	0,228 NS	0,938 NS	1,016 NS	1,241 NS	1,162 NS	1,281 NS	0,652 NS	0,622 NS
20	-	-	1,127 NS	1,197 NS	1,023 NS	0,963 NS	1,103 NS	0,508 NS	0,517 NS
30	-	-	-	0,112 NS	1,982**	1,856 NS	1,905 NS	1,209 NS	1,036 NS
40	-	-	-	-	2,0,24*	1,902 NS	1,949 NS	1,263 NS	1,081 NS
50	-	-	-	-	-	0,011 NS	0,232 NS	0,217 NS	1,028 NS
60	-	-	-	-	-	-	0,212 NS	0,219 NS	0,034 NS
70	-	-	-	-	-	-	-	0,383 NS	0,173 NS
80	-	-	-	-	-	-	-	-	0,123 NS
90	-	-	-	-	-	-	-	-	-

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; NS Teste não significativo.

Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

A Tabela 2.6. apresenta as diferenças significativas entre os dois algoritmos não-paramétricos. E ao investigar estes, ou seja, SVM e RF, pode se constatar que 50/50 foi o único que apresentou diferença significativa a 0,05, ou seja, 5% de probabilidade de erro. Portanto, através dos resultados obtidos, pode afirmar que os dois algoritmos classificam de maneira semelhante. O tamanho de amostra 60/40 foi o qual apresentou menor diferença entre os algoritmos de classificação estudados.

Tabela 2.6. Valores do Teste Z comparando os classificadores com os respectivos tamanhos de amostras.

SVM/RF	10	20	30	40	50	60	70	80	90
10	1,148 ^{NS}	-	-	-	-	-	-	-	-
20	-	0,210 ^{NS}	-	-	-	-	-	-	-
30	-	-	1,242 ^{NS}	-	-	-	-	-	-
40	-	-	-	1,005 ^{NS}	-	-	-	-	-
50	-	-	-	-	2,471*	-	-	-	-
60	-	-	-	-	-	0,128 ^{NS}	-	-	-
70	-	-	-	-	-	-	0,495 ^{NS}	-	-
80	-	-	-	-	-	-	-	1,083 ^{NS}	-
90	-	-	-	-	-	-	-	-	0,921 ^{NS}

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; ^{NS} Teste não significativo.

Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

A Tabela 2.7. apresenta os valores obtidos através do Teste-Z comparando dois algoritmos (RF e ML). Ao comparar estes, pode se constatar que o conjunto amostral 10/90 foi o único que não apresentou diferença significativa. O que determina que RF e ML, realizaram classificações diferentes entre os tamanhos de amostras.

Tabela 2.7. Valores do Teste Z comparando os classificadores com os respectivos tamanhos de amostras de treinamento.

RF/ML	10	20	30	40	50	60	70	80	90
10	0,037 ^{NS}	-	-	-	-	-	-	-	-
20	-	2,186*	-	-	-	-	-	-	-
30	-	-	3,035**	-	-	-	-	-	-
40	-	-	-	3,168**	-	-	-	-	-
50	-	-	-	-	4,082**	-	-	-	-
60	-	-	-	-	-	3,637**	-	-	-
70	-	-	-	-	-	-	2,950**	-	-
80	-	-	-	-	-	-	-	2,931**	-
90	-	-	-	-	-	-	-	-	2,513**

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; ^{NS} Teste não significativo.

Nota. RF é abreviação de *Random Forest* e ML é para *Maximum Likelihood*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

A Tabela 2.8. mostra a comparação dos algoritmos SVM (não-paramétrico) e ML (paramétrico) através do Teste-Z. Pode se constatar que os conjuntos 10/90, 30/70, 50/50, 80/20 e 90/10 não apresentaram diferenças significativas. O que determina que SVM e ML, realizaram classificações diferentes entre os tamanhos de amostras, todavia apresentam mais classificações semelhantes que o RF comparado ao ML.

Tabela 2.8. Valores do Teste Z comparando os classificadores com os respectivos tamanhos de amostras de treinamento.

ML/SVM	10	20	30	40	50	60	70	80	90
10	0,111 ^{NS}	-	-	-	-	-	-	-	-
20	-	1,976*	-	-	-	-	-	-	-
30	-	-	1,796 ^{NS}	-	-	-	-	-	-
40	-	-	-	2,166*	-	-	-	-	-
50	-	-	-	-	1,621 ^{NS}	-	-	-	-
60	-	-	-	-	-	3,764**	-	-	-
70	-	-	-	-	-	-	2,459*	-	-
80	-	-	-	-	-	-	-	1,858 ^{NS}	-
90	-	-	-	-	-	-	-	-	1,605 ^{NS}

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; ^{NS} Teste não significativo.

Fonte: O autor (2017).

Diante destas considerações, recomenda a utilização do RF para fins de classificação, requerendo poucos, ou seja, 10% dos dados para validação. O RF, uma vez que o mesmo apresentou melhores resultados, quando comparados com o ML (paramétrico) e SVM (não-paramétrico). Através deste presente trabalho, foi possível notar que o tamanho da amostra influência os algoritmos de classificação, principalmente quando comparado com classificadores paramétricos e não-paramétricos.

2.4 CONCLUSÕES

Diante das condições e particularidades da área de estudo, ou seja, um ambiente tipicamente rural, pode ser observado, que o Random Forest (RF) é o melhor classificador. Ainda, o RF requereu o menor conjunto de validação. Para este classificador, o maior valor de índice Kappa foi encontrado para o conjunto amostral 90/10. Já para o SVM e ML os maiores valores de índice Kappa foram 60/40 e 40/60, respectivamente.

A ANOVA mostrou diferença significativa entre os classificadores. Todavia, o teste Z, mostrou que para o RF e SVM, apenas o conjunto amostral 50/50 apresentou diferença significativa. O conjunto amostral 60/40 apresentou a menor diferença de classificação entre os algoritmos.

A classificação supervisionada proporcionou notar quais são os tamanhos ideais de amostras de treinamento e validação para cada classificador, assim podendo recomendar valores para maximização dos resultados através dos tamanhos de amostras. Portanto, pode ser recomendado pesquisas que envolvam esta temática de tamanho de amostra de treinamento e validação, pois cada área de estudo e a diferenças de sensores, podem gerar novos resultados e interpretações. Assim para a recomendação e criação de parâmetros na literatura, é de suma importância realizar estas categorizações de tamanhos de amostras ideais.

Deste modo, pode-se concluir nesta pesquisa, que o classificador RF apresentou as melhores classificações quando comparado aos demais algoritmos testados (SVM e ML). Portanto, é recomendado a utilização de classificador não-paramétricos para realizações de classificação de uso e cobertura da terra em cenas de sensores ópticos, quais são comumente utilizados algoritmos paramétricos.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAF. Anuário Estatístico-Associação Brasileira de Produtores de Florestas Plantadas. **Anuaário Estatístico ABRAF**, p. 146, 2013.

ALVARES, C. A. et al. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.

BOLFE, É. L. et al. Avaliação da classificação digital de povoamentos florestais em imagens de satélite através de índices de acurácia. **Revista Árvore**, v. 28, n. 1, p. 85–90, 2004.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37–46, 1960.

CONGALTON, R. G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. **Remote Sens Environ**, v. 46, n. October 1990, p. 35–46, 1991.

DICKS, S. E.; LO, T. H. C. Evaluation of thematic map accuracy in a land-use and land-cover mapping program. **Photogrammetric Engineering & Remote Sensing**, v. 56, n. 9, p. 1247–1252, 1990.

FELDE, G. W. et al. Analysis of Hyperion Data with the FLAASH Atmospheric Correction Algorithm. Internationalgeoscienceand Remote Sensing Symposium. **Anais...** New York: 2003.

FIGUEIREDO, D. **Conceitos Básicos de Sensoriamento Remoto** Companhia Nacional de Abastecimento-CONAB. Brasília - DF: [s.n.]. Disponível em:
[<http://www.ufscar.br/~debe/geo/paginas/tutoriais/pdf/sensoriamento/Conceitos de Sensoriamento Remoto.pdf>](http://www.ufscar.br/~debe/geo/paginas/tutoriais/pdf/sensoriamento/Conceitos de Sensoriamento Remoto.pdf).

GAIAD, N. P. et al. Uso E Cobertura Da Terra Apoiados Em Algoritmos Baseados Em Aprendizado De Máquina: O Caso De Mariana -Mg. **Enciclopédia Biosfera**, v. 14, n. 24, p. 1211, 2017.

GAROFALO, D. F. T. et al. Análise comparativa de classificadores digitais em imagens do Landsat-8 aplicados ao mapeamento temático. **Pesquisa Agropecuária Brasileira**, v. 50, n. 7, p. 593–604, 2015.

HARRIS, J. W.; STÖCKER, H. **Handbook of mathematics and computational science**. New York, 1998.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres**. São José dos Campos: Parêntese, 2009.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **International Biometrics Society**, v. 33, n. 1, p. 159–174, 1977.

MELO, D. H. C. T. B. Uso De Dados Ikonos II Na Análise Urbana : Testes Operacionais Na Zona Leste De São Paulo.(Dissertação de Mestrado em Sensoriamento Remoto)
Instituto Nacional de Pesquisas Espaciais. São José dos Campos, p.146, 2002.

MENESES, P. R.; ALMEIDA, T. de. **Introdução ao processamento de imagem de sensoriamento remoto.** p. 266, 2012.

MOUNTRAKIS, G.; IM, J.; OGOLE, C. Support vector machine in remote sensing: A review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 66, n. 3, p. 247–259, 2011.

NOVO, E. M. L. DE M. **Sensoriamento Remoto: Princípios e Aplicações.** São Paulo: Edgard Blucher, 2010.

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento Remoto da Vegetação.** 2. ed. São Paulo: Oficina De Texto, 2012.

RSTUDIO TEAM, 00. **RStudio: Integrated Development for RBostonRStudio,** , 2016.
Disponível em: <<http://www.rstudio.com/>>

STEFFEN, A. C. **Introdução ao Sensoriamento Remoto.** Disponível em:
<<http://www.inpe.br/unidades/cep/atividadescep/educasere/apostila.htm>>. Acesso em: 2 nov. 2016.

SULSOFT. **Envi 5.4: Descrição Técnica**SulSoft Alimentando sua Criatividade, , 2017.
Disponível em: <http://www.envi.com.br/pdfs/envi_descricao-tecnica-funcionalidades.pdf>

TOUTIN, T. Review article: Geometric processing of remote sensing images: models, algorithms and methods. **International Journal of Remote Sensing**, v. 25, n. 10, p. 1893–1924, 2004.

VAN DER LINDEN, S. et al. The EnMAP-box-A toolbox and application programming interface for EnMAP data processing. **Remote Sensing**, v. 7, n. 9, p. 11249–11266, 2015.

WEISS, C. V. et al. **Mapeamento do uso e ocupação do solo utilizando imagens de satélite do sensor TM/Landsat 5 no litoral sul do Rio Grande do Sul, Brasil.** Anais do XVI Simpósio Brasileiro de Sensoriamento Remoto -SBSR. **Anais....** 2013

CAPÍTULO 3

Avaliação da performance de dados SAR polarimétricos e ópticos para fins de mapeamento no planalto sul catarinense

RESUMO

Dados remotamente situados são comumente utilizados para a confecção de mapas de uso e cobertura da terra. Todavia, a quase totalidade destes produtos são gerados a partir de imagens coletadas por sensores ópticos. No entanto, um fator determinante, que é a cobertura de nuvens, pode limitar o número de cenas disponíveis ao longo de um ano. Sendo assim, os sensores de micro-ondas podem suprir esta demanda limitante, justamente por oferecer cenas “livres” de nuvens. Portanto, o objetivo deste trabalho foi avaliar a performance de dados ópticos (Landsat-5/TM) e dados SAR polarimétricos (ALOS/PALSAR), de forma isolada e combinada, para o mapeamento de uma área tipicamente rural no planalto sul catarinense. Para ambas as cenas foram realizadas diferentes etapas de pré-processamentos. A combinação de ambas fontes de dados permitiu avaliar 13 cenários de classificações usando o *Support Vector Machine* (SVM) e o *Random Forest* (RF). Os resultados mostraram que o índice Kappa máximo encontrado somente para dados oriundos do sensor SAR foi de 86,45%, já para o sensor óptico foi de 94,66% e a integração dos dois sensores atingiu um índice Kappa de 96,75%, ou seja, evidenciando que a combinação dos dados ópticos com os dados SAR, trazem melhorias para fins de classificação do uso e cobertura da terra. Ao se comparar os algoritmos de classificação foi notado que, RF quanto SVM, possuem similaridade no índice Kappa para os diferentes cenários. Entretanto, a ANOVA mostrou que existem diferenças significativas. Com o teste Z, apenas cinco cenários diferenciaram-se significativamente, incluindo a combinação dos dados ópticos e SAR. Sendo assim, é recomendada a utilização de dados SAR para classificações de uso e cobertura da terra no planalto sul catarinense, pois, há dificuldade de se obter imagens de sensores ópticos, livre de nuvens.

Palavras-chave: Sensoriamento Remoto, Decomposição de alvos, SAR, Classificação supervisionada.

CHAPTER 3

Evaluating the Performance of Polarimetric SAR and Optical for Mapping Purposes in the Santa Santa Catarina State Southern Plateau

ABSTRACT

Remote Sensing data are commonly used for producing land use and land cover (LULC) maps. However, most of these products are generated by optical images. A reasonable number of cloud free images are difficult to be acquired in certain tropical regions. Therefore, microwave sensors such as Synthetic Aperture Radar (SAR) may attend this demand just by allowing data acquisition regardless of climate conditions. The main goal of this study was to evaluate the performance of optical images from Landsat-5/TM and polarimetric SAR from ALOS/PALSAR, combined and alone, for mapping purposes in a rural area of the Santa Catarina State Southern Plateau. We performed several preprocessing steps for both scenes. A total of 13 data input scenarios were evaluated using both Support Vector Machine (SVM) and Random Forest (RF) classification methods. The results showed that the use of PALSAR data alone reached a Kappa index of 86.45%, whereas the Landsat 94.66%. The combined use of both Landsat and PALSAR achieved a Kappa index of 96.75%. These results demonstrate that LULC classification using Landsat benefits from PALSAR integration. However, in absence of cloud free Landsat images, very common in the selected region, the use of PALSAR is advised. Both RF and SVM performed similarly in terms of classification performance in the different scenarios. However, Z-test reveals that five scenarios showed significant differences, including the combination of Landsat and PALSAR. We recommend the use of SAR for LULC approaches since the region is frequently covered by clouds.

Keywords: Remote Sensing; Target Decomposition; SAR; Supervised classification.

3.1. INTRODUÇÃO

Informações geradas a partir de imagens e dados remotamente situados, vêm se sobressaindo em diferentes áreas do conhecimento. Estas, na forma de mapas, têm se destacado nas tomadas de decisões (FLORENZANO, 2002). A utilização de sensores remotos para extração de informações sobre ambientes com vegetação, seja tanto para fins de classificação (JANOTH et al., 2007), quanto para recuperação de parâmetros biofísicos (MENESES e ALMEIDA, 2012), têm se desenvolvido nos últimos tempos.

De acordo com Epiphanio et al. (1996) a vegetação é um dos alvos mais complexos para o sensoriamento remoto, pois apresenta várias propriedades de reflexão foliar em uma mesma planta, ou seja, múltiplos componentes refletores. Sendo assim, para o estudo, índices de vegetação são amplamente utilizados. Estes índices são razões matemáticas que para Watzlawick (2003), têm como objetivo realçar as contribuições espectrais da vegetação e diminuir contribuições dos demais espectros.

Para a confecção de mapas temáticos são utilizadas, em geral, imagens advindas de sensores ópticos, por apresentarem maior número de sensores em órbita e por fornecerem imagens, pelas quais o usuário consegue visualizar os objetos na superfície terrestre. Todavia, imagens coletadas por sensores ópticos, apresentam como limitação a obtenção de imagens livres de nuvens. Assim, a utilização de sensores operando na região do micro-ondas surge como alternativa. Estes sensores foram rotineiramente utilizados em épocas militares (MENESES e ALMEIDA, 2012). Atualmente dados de sensores Radares podem gerar informações complementares em aplicações envolvendo classificação e recuperação de parâmetros biofísicos. Os sensores Radares, são sistemas ativos que emitem a própria radiação e podem operar em diferentes frequências (PONZONI, SHIMABUKURO e KUPLICH, 2012).

Estudos envolvendo vegetação têm como objetivo principal obter a discriminação das formações vegetais e verificação de sua extensão territorial. Além das estimativas de suas propriedades biofísicas e bioquímicas. Em geral a diferenciação de áreas com e sem floresta é facilitada com o uso de dados Radar, principalmente quando são utilizados mais de uma banda, polarização e ângulos de incidência (HYYPAA et al., 2000; RANSON et al., 2001; BOYD e DANSON, 2005).

Estudos de classificação de uso e cobertura da terra estão sendo aplicados com a utilização de sensores Radar, como na pesquisa de Muthukumarasamy, Shanmugam e Kolanuvada (2017), que utilizaram o sensor na banda L, e também os autores Mishra et al. (2017), que utilizaram sensor na banda C, para a geração de mapas de uso da terra. Viet Nguyen

et al. (2016) utilizaram a banda L para caracterizar as estruturas de uma floresta tropical. Já Martins et al. (2016) pesquisaram a sensibilidade do sensor ALOS/PALSAR, também na banda L, para detecção de áreas degradadas com fogo no norte da Amazônia. Além disso, os sensores Radar podem oferecer ainda parâmetros estruturais dos objetos da superfície terrestre.

Atualmente, diversos estudos estão sendo realizados com a finalidade de identificação, classificação do uso e cobertura da terra, tanto com dados de sensores Radar quanto Ópticos. Entretanto, ainda são vagos os estudos utilizando ambos sensores. Portanto, o objetivo principal desta pesquisa foi avaliar a performance dos dados ópticos, bem como de dados SAR polarimétricos, de forma individual e em diferentes combinações para fins de mapeamento de ambientes rurais no planalto sul catarinense.

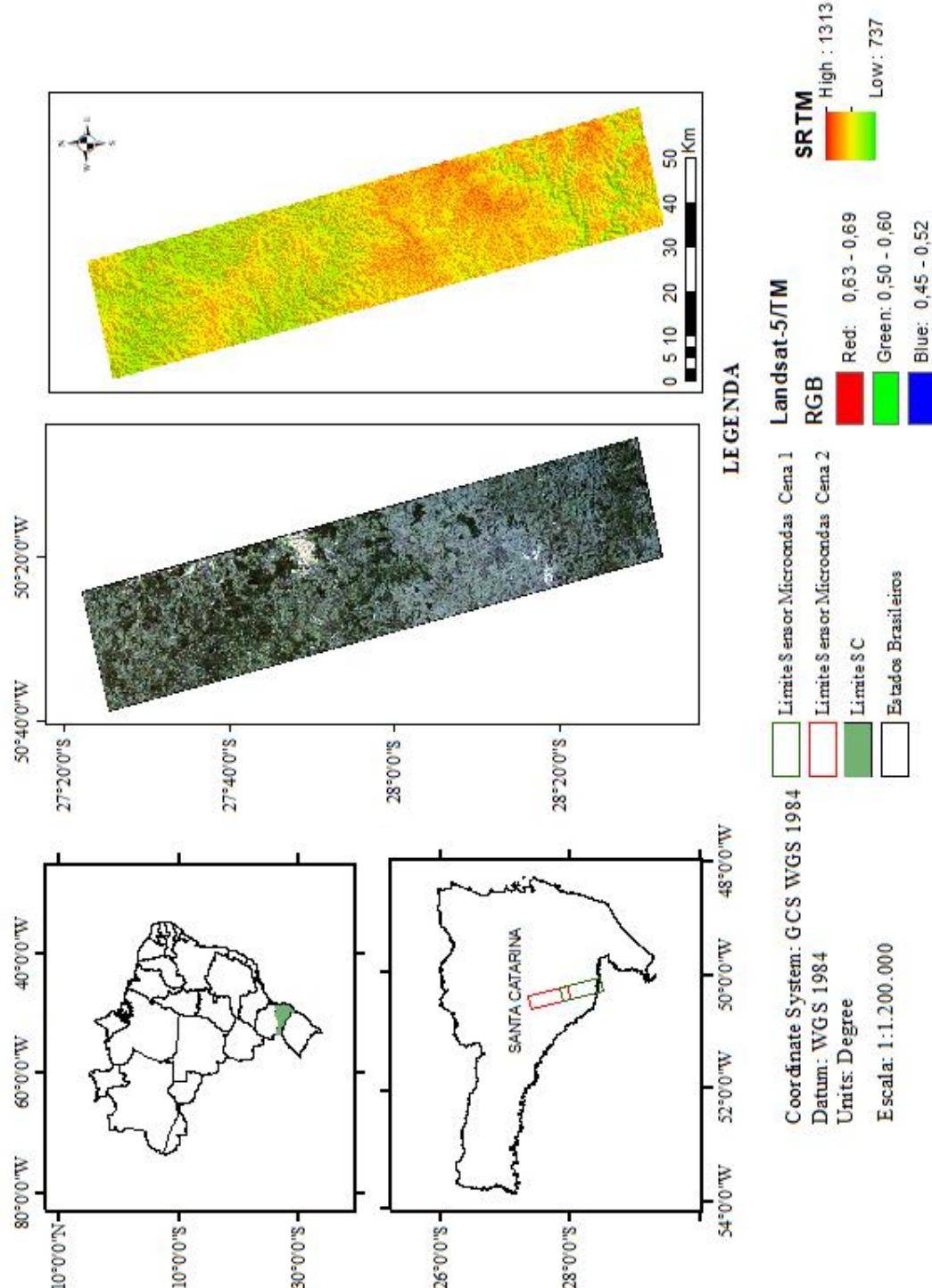
3.2. MATERIAL E MÉTODOS

3.2.1. Descrição da área de estudo

O estudo envolveu sete municípios da serra catarinense (Lages, Curitibanos, Capão Alto, São José do Cerrito, Correia Pinto, Ponte Alta e São Joaquim). A região de estudo está localizada na longitude 50° 39' 40,61'' O. e latitude 28°28' 59,52'' S., abrangendo 3.085,6612 km², conforme apresentado na Figura 3.1.

Segundo Alvares et al. (2013) o clima é classificado como Cfb (temperado subtropical). Estas áreas possuem um forte perfil agrícola, somando-se a expressividade na produção florestal, fator decisivo para alavancagem e consolidação dos segmentos de celulose e papel, madeireiro e moveleiro da macrorregião (ABRAF, 2013).

Figura 3.1. Localização da área de estudo.



Fonte: O autor (2017).

3.2.2. Sensores Orbitais

Foram utilizados dados de dois sensores, sendo um óptico (Landsat-5/TM) e outro de micro-ondas (ALOS/PALSAR). Os sensores estudados apresentam as seguintes características:

- **Landsat-5/TM:** É administrado pela NASA e gerenciado pelo *U.S. Geological Survey* (USGS's), no *Center for Earth Resources Observation and Science*. Embora o sensor esteja desativado desde 2013, os dados históricos da série Landsat são constituídos como uma importante fonte de dados. O Landsat-5 apresenta características específicas conforme a Tabela 3.1. Foram utilizadas duas cenas da órbita 221, referente ao dia 20 de março de 2009, sendo uma do ponto 79 e outra do 80. As cenas foram adquiridas pelo site Earth Explorer (site: www.earthexplorer.usgs.gov).

A cena do Landsat-5 apresenta 185x185 km e para o presente estudo, foi realizado um recorte que envolve uma área coincidente com as cenas do sensor ALOS/PALSAR.

Tabela 3.1. Resoluções espectrais e espacial do Satélite Landsat -5.

Sensor	Região do Espectro	Intervalo (μm)	Res. Espacial (m)
TM (Thematic Mapper)	Azul	0,45 -0,52	30
	Verde	0,50 -0,60	
	Vermelho	0,63-0,69	
	Infravermelho Próximo	0,76-0,90	
	Infravermelho Médio	1,55 -1,75	
	Infravermelho Termal	10,4-12,5	120
	Infravermelho Médio	2,06-2,35	30

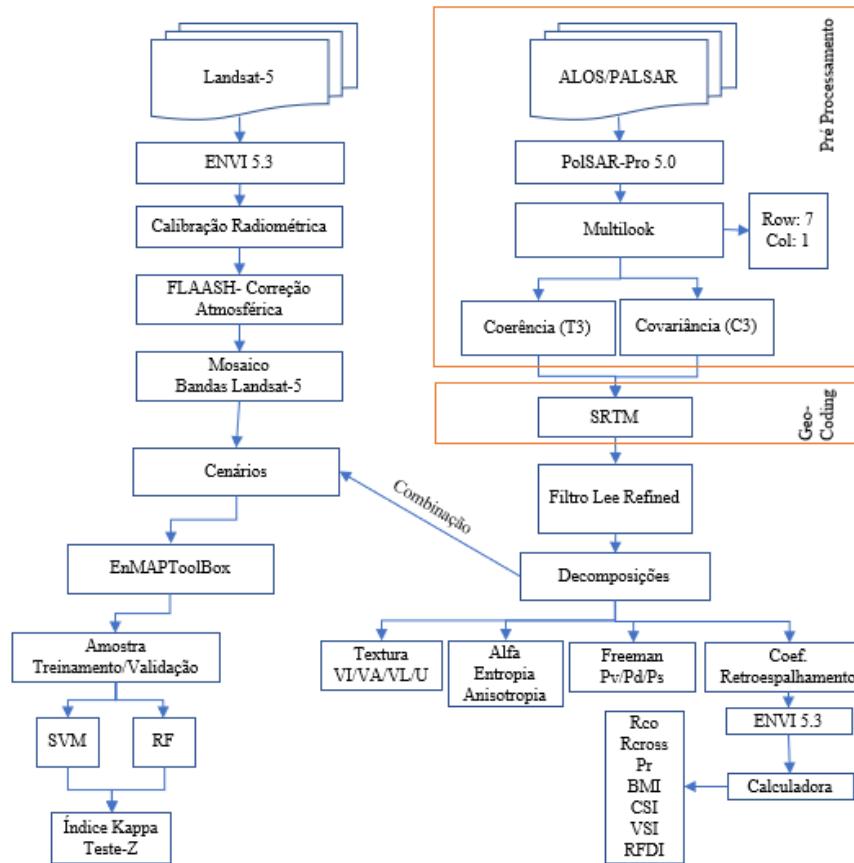
Fonte: Novo (2010).

- **ALOS/PALSAR:** Sensor radar de abertura sintética (SAR) na banda-L, proporciona observação detalhada, bem como interferometria de repetição. Foi desenvolvido para contribuir nos campos de mapeamento, observação regional de cobertura da terra, monitoramento de desastres e levantamento de recursos naturais. O ALOS é uma missão da Agência de Exploração Aeroespacial do Japão (JAXA). As cenas utilizadas neste estudo apresentam resolução espacial reamostradas de 30 metros e com quatro polarizações (HH, VV, HV e VH) adquiridas de modo experimental na região sul do Brasil. Foram utilizadas duas cenas do dia três de novembro de 2009, as quais foram adquiridas em formato *Slant-range Singlelook Complex* (SLC), no nível de processamento 1.1.

3.2.3. Pré-processamento das cenas

Foram realizadas diversas etapas de pré-processamento das cenas, a priori para a cena do Landsat-5/TM, foi realizada a conversão de níveis digitais para radiância ($\text{Wm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$) e posteriormente a conversão para reflectância de superfície, considerando a minimização dos efeitos atmosféricos através do “*fast line of sight atmospheric analysis of spectral hypercubes*” (Flaash) proposto por Felde et al. (2003). Logo, para o ALOS/PALSAR nível 1.1, foram realizados procedimentos para redução do efeito *speckle*, ou seja, diminuição da largura da distribuição na resolução espacial pelo mesmo fator, e foi realizada ainda a decomposição polarimétrica, conforme apresentado na Figura 3.2, onde é explanado um esquema do processamento realizado para ambos sensores.

Figura 3.2. Processamento realizado para processamento das cenas dos sensores utilizados.



Fonte: O autor (2017).

De posse dos dados de amplitude e fase polarimétrica foram calculadas: incoerência, coerência e textura. A decomposição apresenta como objetivo o mecanismo médio de espalhamento como uma soma de elementos independentes, os quais foram associados a um

mecanismo físico. Maiores informações sobre os processamentos dos dados podem ser encontradas nas referências apresentadas nas tabelas a seguir. Na Tabela 3.2 são evidenciadas as decomposições incoerentes, as quais representam dados gerados a partir dos coeficientes de retroespalhamento.

Henderson e Lewis (1998) propuseram, para uma dada polarização, o coeficiente de retroespalhamento denominado de sigma zero (σ^0 ; dB). O sigma zero depende de diversos parâmetros do alvo (tamanho, forma, propriedades dielétricas, orientação, rugosidade).

A Taxa de Co-Polarização, proposta por Henderson e Lewis (1998), é a razão entre os dois retroespalhamentos obtidos na vertical e horizontal, respectivamente. Logo, através da taxa de Co-Polarização destacam-se diferentes orientações verticais e horizontais dos alvos, quais são derivadas dos aspectos estruturais da vegetação. Também, proposta pelos mesmos autores, a Relação de Polarização Cruzada, é a razão entre o retroespalhamento da polarização cruzada pelo retroespalhamento horizontal. Esta razão, é sensível à dispersão volumétrica de uma floresta e reduz os efeitos topográficos no retroespalhamento.

Pope, Rey-Benayas e Paris (1994) propuseram três índices: Índice de Biomassa (utilizado para diferenciar áreas vegetadas e não vegetadas, como também, entre terrenos inclinados e planos), Índice de estrutura de copa (aplicado para diferenciação de florestas de pântano) e o Índice de espalhamento de volume (empregado para distinguir entre florestas de terra firme e em regeneração). A Potência Total sugerida por Boerner et al. (1991) é a soma da potência de todas as polarizações espalhadas de um alvo de volta ao SAR. O Índice de Degradação de Florestas sugerida por Saatchi et al. (2010) e Mitchard et al. (2012) é a relação da diferença do retroespalhamento horizontal e o cruzado pela somatória do retroespalhamento horizontal e o cruzado. Este índice avalia a força do mecanismo de double-bounce, logo, é utilizado para diferenciação de grupos vegetações.

Tabela 3.2. Atributos Incoerentes

Incoerente	Equação	Descrição	Autores
Coeficiente Retroespalhamento $\sigma_{HH}^0, \sigma_{HV}^0,$ σ_{VH}^0 e σ_{VV}^0	$\sigma_{ij}^0 = 10 \frac{\sigma_{ij}^0 (\text{dB})}{10}, \text{ onde}$ $\sigma_{ij}^{0 (\text{dB})} = 10 \times \log_{10} (I^2 + Q^2) + FC - A$	Indica a orientação dos componentes da floresta (i.e., mais troncos produzem maior retroespalhamento vertical)	Henderson e Lewis (1998) Schimada et al. (2009)
Relação de Co-Polarização	$R_{CO} = \frac{\sigma_{VV}^0}{\sigma_{HH}^0}$	Destaca diferentes orientações verticais e horizontais derivadas dos aspectos estruturais da vegetação	Henderson e Lewis (1998)
Relação de Polarização Cruzada	$R_{Cross} = \frac{\sigma_{HV}^0}{\sigma_{HH}^0}$	Sensível à dispersão volumétrica da floresta para apoiar a classificação e reduzir os efeitos topográficos no retroespalhamento	Henderson e Lewis (1998)
Potência Total	$Pt = \sigma_{HH}^0 + \sigma_{VV}^0 + 2 \times \sigma_{HV}^0$	Energia total de retroespalhamento derivada da cobertura vegetal	Boerner et al. (1991)
Índice de Biomassa	$BMI = \frac{\sigma_{HH}^0 + \sigma_{VV}^0}{2}$	Quantidade relativa de componente lenhoso em comparação com biomassa foliar	Pope et al. (1994)
Índice de estrutura de Copa	$CSI = \frac{\sigma_{VV}^0}{\sigma_{HH}^0 + \sigma_{VV}^0}$	Presença relativa de dispersores verticais (troncos e caules) de vegetação	Pope et al. (1994)
Índice espalhamento de Volume	$VSI = \frac{\sigma_{HV}^0}{\sigma_{HV}^0 + BMI}$	Indicador de espessura do dossel ou densidade relativa	Pope et al. (1994)
Índice degradação de florestas	$RFDI = \frac{\sigma_{HH}^0 - \sigma_{HV}^0}{\sigma_{HH}^0 + \sigma_{HV}^0}$	Relação projetada para avaliar a força do mecanismo de duplo salto, o que é útil para diferenciar a vegetação	Saatchi et al. (2010); Mitchard et al. (2012)

Nota: Fc = fator de calibração (-83); A= fator de conversão (32). Os valores obtidos em escala logarítmica foram convertidos em escala linear para os cálculos dos atributos.

Fonte: O autor (2017).

As decomposições coerentes consistem em um espalhamento determinístico, sendo utilizando normalmente com a matriz de Sinclair [S], o qual caracteriza os estados de polarização (SINCLAIR, 1950; ULABY e ELACHI, 1990 ; CLOUDE, 2009; LEE e POTTIER, 2009). Conforme apresentado na Tabela 3.3, as equações se relacionam com a onda espalhada ou recebida à onda incidente ou transmitida (ULABY e ELACHI, 1990).

De acordo com Ulaby e Elachi (1990) a decomposição coerente é complexa, e expressa como o espalhador transforma o campo transmitido (Et) no recebido (Er), assim caracterizando

o comportamento do espalhador. As decomposições mais empregadas para utilizações em áreas agrícolas, são as propostas por Freeman e Durden, Lewis e Henderson e Cloude Pottier.

A decomposição de Freeman e Durden (1998) possui três componentes sendo eles: Espalhamento Superficial, Volumétrico e double-bounce. Outra decomposição coerente é a diferença de Fase proposta por Lewis e Henderson (1998), no qual indica a estrutura e a quantidade de biomassa. A decomposição de Cloude e Pottier (1997) supõe que há um mecanismo dominante médio em cada célula, o qual pode ser representado por três parâmetros invariantes à rotação da plataforma: Ângulo alfa médio, Entropia e Anisotropia. Esta decomposição pode-se considerar em um espaço de classificação bidimensional, como apresentado na Figura 1.9, que demonstra as zonas para cada mecanismo de retroespalhamento (Double Bounce, Volume e superficie).

Segundo Cloude e Pottier (1997) são nove diferentes zonas válidas, que representam diferentes mecanismos de retroespalhamento. Uma descrição de cada zona pode ser encontrada na Tabela 1.3. Logo, a complexidade da estrutura de uma floresta, pode ser explicada por apresentar alta entropia, consequentemente uma baixa anisotropia e valor alfa perto de 45° (CLOUDE e POTTIER, 1997).

Tabela 3.3. Atributos Coerentes

Coerente	Equação	Descrição	Autores
Diferença de Fase	$\Delta\varphi_{HH} - VV = \arg(S_{HH}S_{VV}^*)$ $\Delta\varphi_{HV} - HV = \arg(S_{HH}S_{HV}^*)$ $\Delta\varphi_{HV} - VV = \arg(S_{HV}S_{VV}^*)$	Indicação da estrutura e quantidade de biomassa	Henderson e Lewis (1998)
Entropia	$H = - \sum_{i=1}^3 p_i \log_3(p_i); p_i = \frac{\lambda_i}{\sum_{j=1}^3 \lambda_j}$	Relacionado à complexidade da estrutura da floresta. Floresta mais complexa e diversificada apresenta alta H, baixa A e perto de 45°.	Cloude e Pottier (1997)
	$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3}$		
	$\alpha = \sum_{i=1}^3 p_i \alpha_i$		
Contribuição de dispersão de volume para potência total	$P_V = \frac{8 f_V}{3}$	Proporção da retroespalhamento volumétrica associada à estrutura da floresta e ao conteúdo de biomassa	Freeman e Durden (1998)
Contribuição de dispersão de Double-bounce para a potência total	$P_d = f_d(1 + \beta ^2)$	Indicação da abertura do dossel, densidade e número de árvores (troncos)	Freeman e Durden (1998)
Contribuição de dispersão de Surface potência total	$P_s = f_s(1 + \beta ^2)$	Relacionado à abertura do dossel	Freeman e Durden (1998)

Nota: arg, S, * e () denotam função de argumento, amplitude de dispersão complexa, conjugado complexo e média espacial, respectivamente. $\lambda_1 > \lambda_2 > \lambda_3 > 0$ são autovalores reais da matriz de coerência [T]. $i =$ é um parâmetro derivado dos autovetores de [T] que descreve o mecanismo de dispersão dominante.

Fonte: O autor (2017).

Como apresentado na Tabela 3.4, os atributos de texturas, trabalham com a intensidade do sinal, assim sendo distinguido em diferentes momentos. De acordo com Wijaya et al. (2010), Gallardo-Cruz et al. (2012) e Laurin et al. (2013) a utilização de informações de texturas melhoraram o resultado das classificações em ambientes tropicais complexos.

Tabela 3.4. Atributos Texturais calculados a partir dos coeficientes de retroespalhamento dos dados SAR.

Textura	Equação	Descrição	Autores
VI	$VI_{XY} = \frac{\sum_{i,j} (I_{ij})^2 / m}{\left(\sum_{i,j} I_{ij} / m \right)^2} - 1$	Uma proporção da média da intensidade quadrada para a intensidade média ao quadrado.	Oliver and Quegan, 1998
VA	$VA_{XY} = \frac{\sum_{i,j} (A_{ij})^2 / m}{\left(\sum_{i,j} A_{ij} / m \right)^2} - 1$	Uma proporção de intensidade média para a amplitude média quadrática.	Oliver and Quegan, 1998
VL	$VL_{XY} = \left[\sum_{i,j} \ln(I_{ij})^2 / m \right] \cdot \left[\sum_{i,j} \ln(I_{ij}) / m \right]^2$	A diferença entre o valor médio do logaritmo de intensidade quadrado e o quadrado do logaritmo de intensidade média.	Oliver and Quegan, 1998
U	$U_{XY} = \left[\sum_{i,j} \ln(I_{ij}) / m \right] \cdot \ln \left[\sum_{i,j} (I_{ij}) / m \right]$	Uma medida de log normalizada de textura. Calcula a distribuição de K.	Oliver and Quegan, 1998

Observação: XY representa HH, HV, VH ou VV.

Fonte: O autor (2017).

3.2.4. Descrição das classes de uso e cobertura da terra

Em aplicativo GIS foi realizada a coleta das amostras das diferentes classes de uso e cobertura da terra. Este procedimento foi realizado de forma visual. Para a coleta, foi utilizada a cena do Landsat-5 e para a verdade de campo, foram utilizadas imagens do Google Earth e as imagens ortorretificadas do levantamento aéreo, disponibilizadas pelo Estado de Santa Catarina (GOVERNO DO ESTADO DE SANTA CATARINA, 2017). As amostras foram coletadas em formato *shapefile* e posteriormente transformadas em Região de Interesse (ROI). Estas transformações foram realizadas no aplicativo ENVI. As seis classes de uso e cobertura da terra estão descritas na Tabela 3.5.

Tabela 3.5. Descrição das seis classes de uso e cobertura da terra.

Floresta Nativa	Áreas cobertas por vegetação de porte florestal nativa.
Reflorestamento	Área cobertas por vegetação de porte florestal plantado (reflorestamentos de eucalipto e pinus).
Solo Exposto	Área em que houve corte raso da vegetação natural e/ou colheita florestal e/ou agronômica.
Água	Áreas como rios, lagos, represas e entre outras.
Área Urbanizada	Áreas com feições e características de ocupações urbanas observadas nas imagens.
Campo Nativo	Área cobertas por vegetação rasteira nativa e/ou plantada (manejada).

Fonte: O autor (2017).

Foram selecionados 1000 pixels para cada classe. As quais foram divididos em 600 pixels para treinamento e 400 pixels para validação. Este nível de tratamento proporcionou a menor diferença entre os classificadores (e.g. SVM e RF) (resultados apresentados no Capítulo 2 da presente pesquisa). As classificações dos dados remotamente situados foram realizadas empregando os seguintes algoritmos: *Support Vector Machine* (SVM) e *Random Forest* (RF). O algoritmo SVM foi introduzido por meio da teoria estatística de aprendizagem assumindo como apoio a superfície de decisão, conhecida como hiperplano ideal, o qual elevar ao máximo os limites de separação de classes (VAPNIK, 1999; NASCIMENTO et al., 2009). Já o RF é um algoritmo proposto por Breiman (2001) sendo fundamentado em árvores de decisões, “Conjunto” (ensemble), que de acordo com Kuncheva (2004) são métodos que geram muitas classificações e combinam seus resultados, priorizando os melhores resultados.

As classificações foram geradas tanto para o sensor óptico, quanto para o SAR e as diferentes combinações entre ambos. A Tabela 3.6 expõe os diferentes cenários, variáveis e número total de variáveis para cada cenário.

Tabela 3.6. Combinação dos cenários para classificação de uso e cobertura da terra.

Cenário	Descriores	Variáveis	Nº variáveis
S1	Coerentes	ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} , H, A, α , Pv, Pd, Ps, Pedestal, RVI e Autovalores (l1, l2, l3)	14
S2	Coerentes + Texturas	ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} , H, A, α , Pv, Pd, Os, Pedestal, α , RVI, Autovalores (l1, l2, l3), VI _(HH, HV, VH e VV) , VA _(HH, HV, VH e VV) , VL _(HH, HV, VH e VV) e U _(HH, HV, VH e VV)	30
S3	Incoerentes + coerentes	σ_{HH} , σ_{HV} , σ_{VH} , σ_{VV} , R _{co} , P _T , BMI, CSI, VSI, RFDI, ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} , H, A, α , Pv, Pd, Os, Pedestal, α , RVI e Autovalores (l1, l2, l3)	24
S4	Incoerentes	σ_{HH} , σ_{HV} , σ_{VH} , σ_{VV} , R _{co} , P _T , BMI, CSI, VSI e RFDI.	10
S5	Incoerentes + Texturas	S4 + Textura	26
S6	SAR total	S2 + S4	40
S7	SAR+ Óptica	S2 + S4 + S13	46
S8	Óptica+ Incoerentes	S13 + S4	16
S9	Óptica + Incoerentes + Texturas	S13 + S5	32
S10	Óptica + Incoerentes + Coerentes	S13 + S3	30
S11	Óptica + Coerentes + Texturas	S13 + S2	36
S12	Óptica + Coerentes	S13 + S1	20
S13	Dados ópticos	$\rho_{485}, \rho_{560}, \rho_{660}, \rho_{830}, \rho_{1625}, \rho_{2215}$	6

Nota: ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} = Diferença de fase; H= Entropia; A= anisotropia; α =Alfa médio; Pv= Contribuição de dispersão de volume; Pd= Contribuição de dispersão de Double-bounce; Ps= Contribuição de dispersão de Surface; σ = Retroespalhamento; Rco= Relação de Co-Polarização; Rcross= Relação de Polarização Cruzada; Pt= Potência Total; BMI= índice de Biomassa; CSI= índice de estrutura de Copa; VSI= índice de espalhamento de Volume; RFDI= índice de degradação de florestas.

Fonte: O autor (2017).

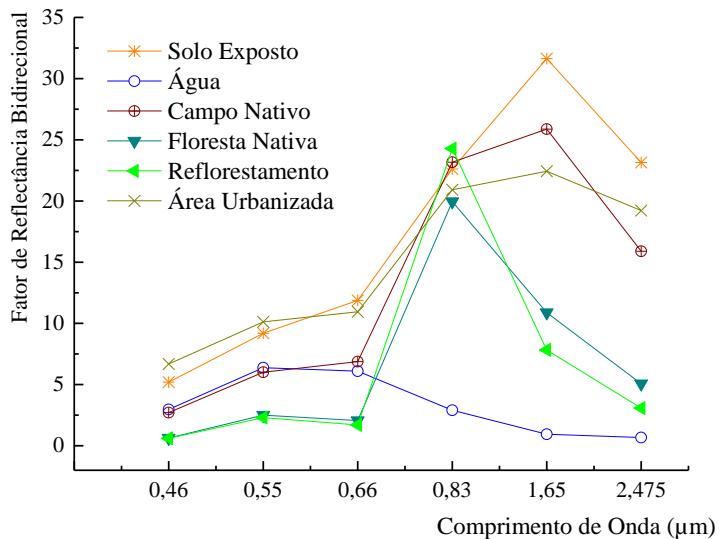
Para a geração das classificações foi utilizado o aplicativo EnMAPToolBox (VAN DER LINDEN et al., 2015). Posteriormente, com as classificações obtidas, foram geradas as matrizes de confusão e também calculado o índice Kappa sugerido por Cohen (1960) e Congalton (1991). As significâncias dos algoritmos de classificação e dos cenários, foram realizadas pelo Teste-Z, cujos valores críticos para 95% e 99% são em ordem de 1,96 e 2,57, respectivamente.

3.3. RESULTADOS E DISCUSSÃO

A seleção de 1.000 pixels amostrados para cada uma das seis classes de uso e cobertura da terra, foram submetidas inicialmente a análise de perfis espectrais através do sensor

TM/Landsat-5 (Figura 3.3). As classes solo exposto, campo nativo e área urbanizada obtiveram os maiores valores de reflectância de superfície no infravermelho médio (1,55-1,75 μm), seguidos pela floresta nativa e plantada no infravermelho próximo (0,76-0,90 μm). A classe água mostrou em geral baixos valores de reflectância de superfície, com os maiores valores ocorrendo na região do visível (400-690nm) e decrescendo gradualmente para o infravermelho médio e próximo, respectivamente. Os perfis espectrais corroboram com os exemplos explanados para vários alvos por Ponzoni, Shimabukuro e Kuplich (2012) como supracitado na capítulo 1.

Figura 3.3. Assinatura Espectral das classes amostradas para classificação de uso e Cobertura da terra referente ao sensor TM/Landsat-5.



Nota: Foram utilizados o conjunto total de dados.

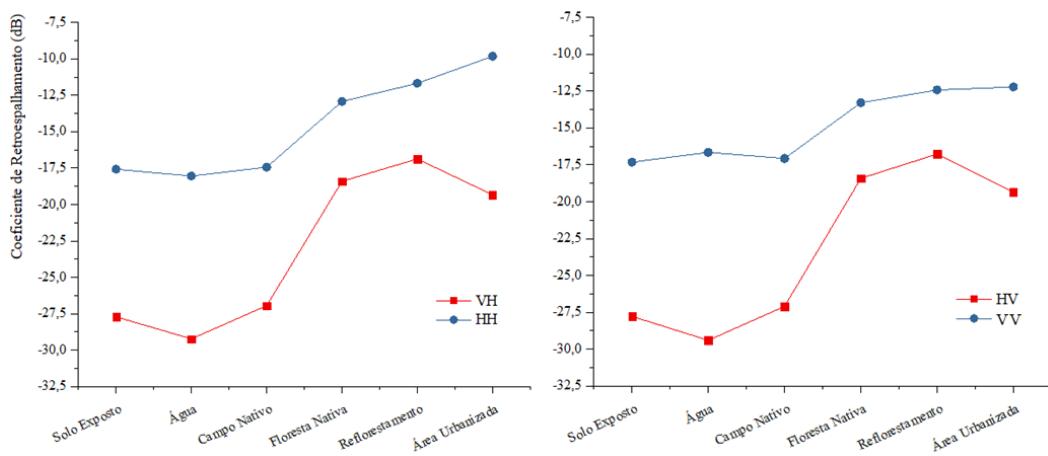
Fonte: O autor (2017).

Na Figura 3.4, é possível observar que as polarizações cruzadas adquiriram os menores valores de retroespalhamento e apresentaram similaridade entre elas. Logo, para HH e VV ocorreu similaridade entre as classes de floresta nativa e reflorestamento. Mishra et al. (2017), ao interpretarem os coeficientes de retroespalhamento para HH e HV, com dados da banda C, encontraram valores inferiores a este estudo.

Wijaya, Marpu e Gloaguen (2010) pesquisaram a discriminação de vegetação em ambientes pantanosos com a utilização de dados do sensor TerraSAR X (HH e VV). Os autores notaram grande sobreposição de suas classes em virtude da heterogeneidade baixa das mesmas representadas. No entanto, no presente estudo, houve sobreposição, porém, de maneira moderadamente baixa, pois as classes interagem de maneira diferenciada com o sinal do ALOS/PALSAR.

Liesenberg et al. (2016) estudaram o efeito da geometria e da umidade, na classificação de florestas tropicais com ALOS/PALSAR (Dual) e encontraram similaridades em suas classes florestais, tais como as classes deste estudo (Figura 3.4). Os mesmos pesquisadores encontraram valores semelhantes da polarização HH para suas classes de florestas secundárias (0 a 15 anos) com as classes de floresta nativa e Reflorestamento deste estudo, conforme apresentado na Figura 3.5.

Figura 3.4. Coeficiente de Retroespalhamento em relação as polarizações

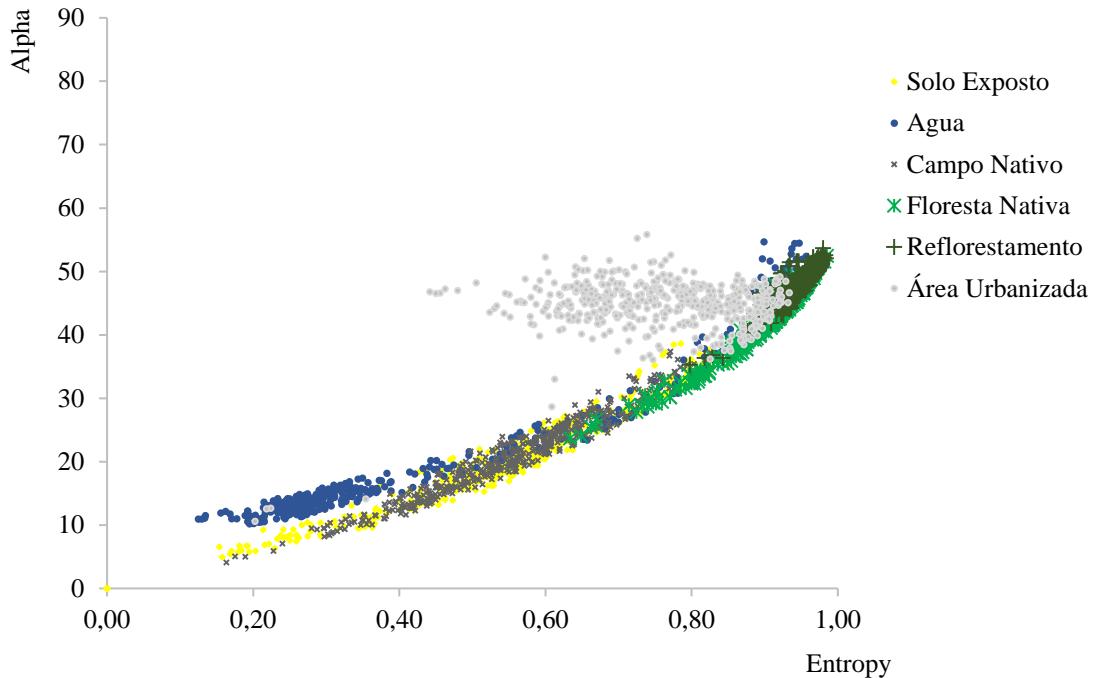


Nota: Foram utilizados o conjunto total de dados SAR.

Fonte: O autor (2017).

O plano bidimensional de entropia e ângulo alpha, proposto por Cloude e Pottier (1997) é apresentado na Figura 3.5. Este demonstra que cada tipo de mecanismo de espalhamento está correspondendo com às descrições das classes propostas neste estudo (i.e. Reflorestamento que se encontra na zona 2, a qual abrange o mecanismo de espalhamento de copas de florestas e de alguns tipos de superfícies vegetadas). As classes amostrais correspondentes a solo exposto, água e campo nativo ficaram claramente abaixo de 40° de alpha, o que corrobora, com os estudos dos autores supracitados, visto que áreas mais planas tendem a apresentar estes valores, indicando predominância em dispersão de superfície.

Figura 3.5. Plano bidimensional alfa vs. entropia proposto por Claude e Pottier (1997) para as classes amostrais selecionadas.



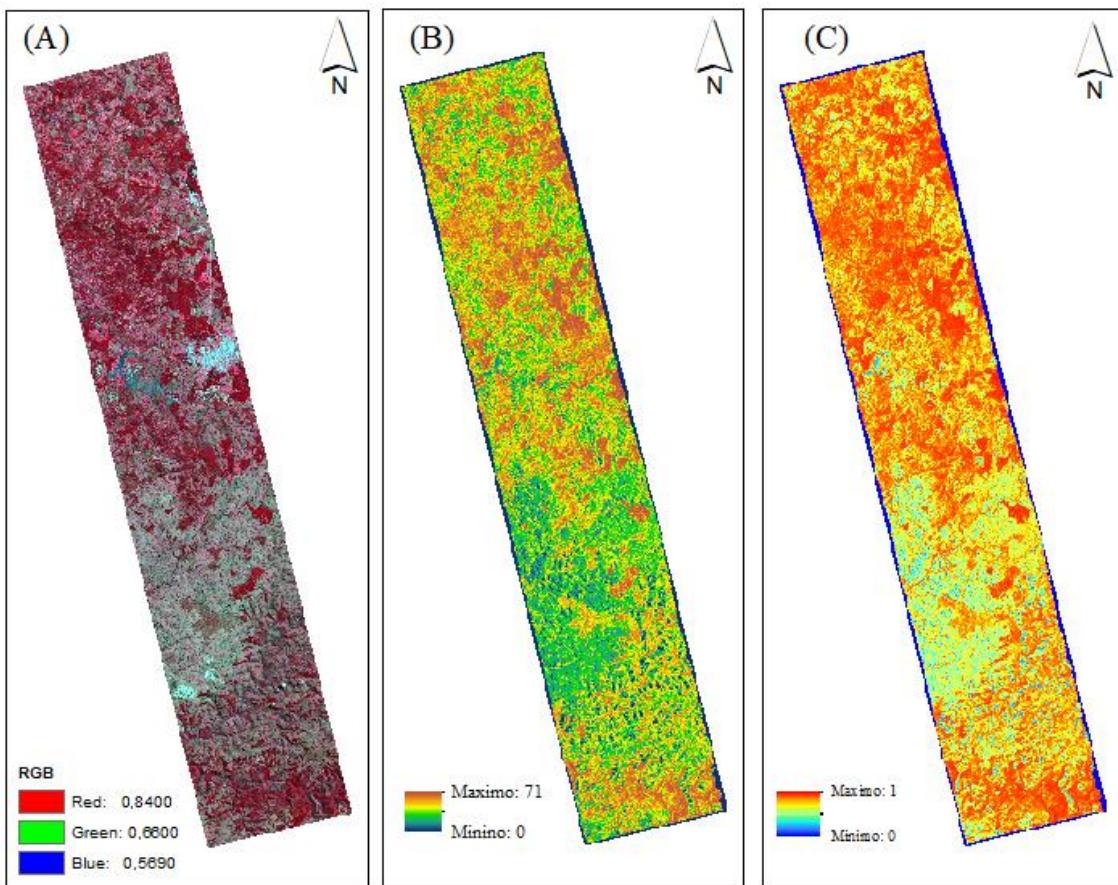
Nota: Foram utilizados o conjunto total de dados.

Fonte: O autor (2017).

Martins et al. (2016), analisaram a degradação da Amazônia brasileira pelo fogo e concluíram, que a entropia diferenciou os valores de florestas com maior altura, a anisotropia distinguiu áreas com e sem presença de florestas, e os atributos coerentes viabilizaram a caracterização de áreas florestais afetadas por fogo.

Analizando a decomposição de Claude e Pottier (1997), como apresentado na Figura 3.6., é perceptível que houve uma predominância entre valores de entropia elevados na área de estudo. De acordo com os autores, há uma presença significativa de florestas, o que também é explicado com os valores máximos encontrados na Figura 3.4. Garcia et al. (2010) estudaram cenas do TERRASAR-X para classificação da cobertura vegetal do sudoeste da Amazônia e concluíram que a combinação de três componentes da matriz de covariância (HH, VV e HH-VV) somando a entropia, apresentou um produto satisfatório na separação de classes.

Figura 3.6. Decomposições de Claude e Pottier (1997). (A) Landsat-5; (B) Angulo médio Alpha; (C) Entropia.



Fonte: O autor (2017).

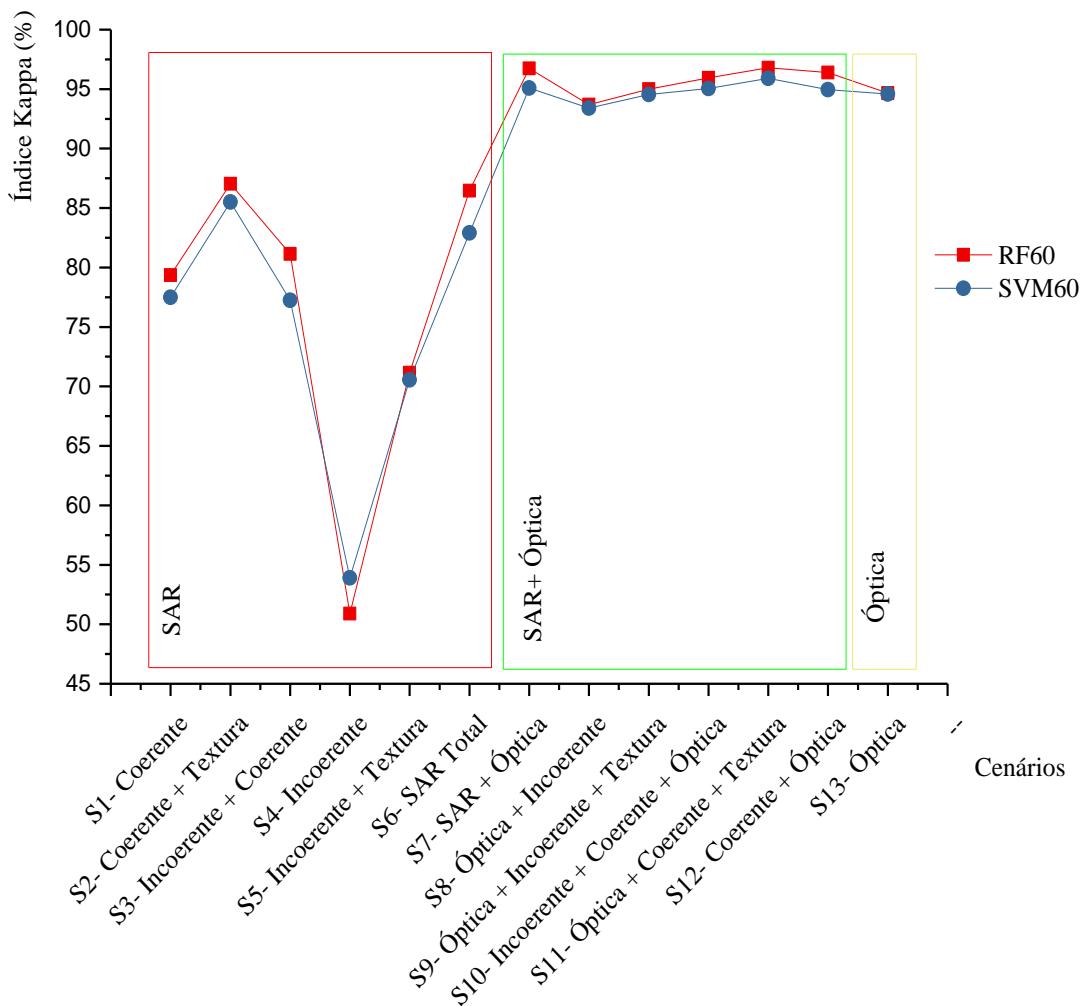
A figura 3.7. mostra os índices Kappa para cada algoritmo de classificação e para os diferentes cenários estudados. Pode ser percebido, que a soma total dos produtos SAR com o sensor óptico (S7) apresentou o maior índice Kappa, sendo, para RF de 96,75% e SVM de 95,1%. A soma total dos produtos SAR apresentou índice de 86,45% para RF e de 82,9% para SVM, já o sensor óptico apresentou um índice equivalente a 94,66% para RF e de 94,58% para SVM. Em outras palavras, a perda do índice Kappa do SAR total (S6) foi de 8,21% e 11,68% em relação ao sensor óptico (S13) no classificador RF e SVM respectivamente. O ganho da integração dos sensores (S7) foi de 2,09% para RF e 0,52% para SVM em relação ao sensor óptico (S13).

As classificações realizadas utilizando somente dados do sensor ALOS/PALSAR, alcançaram valores de índice Kappa variando de 50,9% a 87,05% (i.e., S1, S2, S3, S4, S5). Avtar, Takeuchi e Sawada (2013), para classificação de uso e cobertura da terra, encontraram valores de índice Kappa para os decompositores coerentes (i.e., Freeman e Durden e Cloude e Pottier) de 88% para ambos, ou seja, valores acima a deste estudo, qual foi de 79% para a soma das decomposições coerentes (S1).

Os menores valores foram encontrados para os atributos incoerentes e incoerentes mais combinações (i.e., S4, S5, S6) variando de 50,9% a 71,15%. Por outro lado, dados SAR integrados com dados ópticos (i.e., S7, S8, S9, S10, S11, S12) proporcionaram índices Kappa acima de 93,4%.

Mishra et al. (2017) trabalharam com sensor Radar na banda C para realizar a classificação com o cenário de duas polarizações mais textura, e obtiveram o índice Kappa para SVM e RF de 86,7% e 86,0%, respectivamente. Igualando a esse estudo com os valores para RF de S6 (86,45%) e SVM com S2 (85,5%).

Figura 3.7. Índice Kappa com os diferentes cenários de produtos SAR e óptico.



Nota: SVM é a abreviação para *Support Vector Machine* e RF é abreviação de *Random Forest*.

Fonte: O autor (2017).

Muthukumarasamy, Shanmugam e Kolanuvada (2017), avaliaram ALOS/PALSAR e realizaram classificações separadamente das decomposições coerentes, chegando à conclusão

de que Freeman-Durden e Cloude-Pottier atingiram índices Kappa de 78,0% e 53,0% respectivamente. No entanto, na presente pesquisa Freeman e Cloude foram somados e classificados (S1), resultando no índice de Kappa de 77,5%.

Gou, Li e Yang (2016) compararam dados polarimétricos de ALOS/PALSAR, para classificação de zonas costeiras, utilizando a textura, e concluíram que o índice Kappa obtido foi de 69,04% com o classificador SVM. Já Liesenberg et al. (2016) compararam diferentes cenários de ALOS/PALSAR (dual) e somaram ao sensor Landsat, utilizando o classificador RF, e obtiveram o índice Kappa (S13= 76%, S10= 80%, S8= 77%, S3= 47% e S4= 45%). Estes foram inferiores ao deste estudo, porém com a mesma similaridade em diferentes cenários (Figura 3.7.).

Ao se comparar os cenários (Tabela 3.7), o Teste-Z para o classificador RF não foi significativo para: S1 com S3, S2 com S6, S7 com S10/ S11/ S12, S8 com S9/ S13, S9 com S10, S10 com S11/ S12 e S11 com S12. Foi notado também que as classes S6, S7 e S13 para RF foram significativas a 1%, demonstrando que houve diferenças entre as classificações. Ao avaliar o classificador SVM com os produtos S1 com S3, S7 com S9/ S10/ S11/ S12, S8 com S9/ S13, S9 com S10/ S11/ S12/ S13, S10 com S11/ S12, S11 com S12 e S12 com S13 não foram observadas diferenças significativas. Entretanto, para as classes S6, S7 e S13 as diferenças foram significativas havendo disparidade entre as classificações.

Já para a avaliação individual dos algoritmos, o classificador RF apresentou 83,34% dos cenários significativos a 1% de probabilidade, 2,56% a 5% e 14,10% não foram significativos ao Teste-Z, como apresentado na Tabela 3.9. Na comparação entre os cenários para o classificador SVM, 73,08% foi significativa a 1% de probabilidade, 7,69% a 5% e 19,2% não foram significativos ao Teste-Z, como apresentado na Tabela 3.8. Pode ser percebido que os valores não significativos foram referentes ao cenário S7, SAR+Óptica, em relação ao demais cenários de integração dos dados SAR e ópticos.

Estes resultados apontam diferenças nas combinações, no qual pode apresentar melhorias, como por exemplo, com as decomposições coerentes e textura, ou então tornar “pior” a classificação como os atributos da decomposição incoerente. Entretanto, os dados SAR, quando comparados individualmente aos dados ópticos, apresentam redução de apenas 8 a 10% no índice Kappa da classificação. Isto é um indicativo de que dados SAR podem ser utilizados para fins de classificação em períodos aonde não há imagens ópticas disponíveis. A ocorrência de nuvens em imagens Landsat é recorrente na região do planalto sul catarinense.

Tabela 3.7. Valores do Teste Z comparando os classificadores com seus respectivos cenários.

SVM/RF	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
S1	1,39 ^{NS}	-	-	-	-	-	-	-	-	-	-	-	-
S2	-	1,41 ^{NS}	-	-	-	-	-	-	-	-	-	-	-
S3	-	-	2,98**	-	-	-	-	-	-	-	-	-	-
S4	-	-	-	1,77 ^{NS}	-	-	-	-	-	-	-	-	-
S5	-	-	-	-	0,40 ^{NS}	-	-	-	-	-	-	-	-
S6	-	-	-	-	-	3,07**	-	-	-	-	-	-	-
S7	-	-	-	-	-	-	2,63**	-	-	-	-	-	-
S8	-	-	-	-	-	-	-	0,38 ^{NS}	-	-	-	-	-
S9	-	-	-	-	-	-	-	-	0,64 ^{NS}	-	-	-	-
S10	-	-	-	-	-	-	-	-	-	1,37 ^{NS}	-	-	-
S11	-	-	-	-	-	-	-	-	-	-	1,96*	-	-
S12	-	-	-	-	-	-	-	-	-	-	-	2,25*	-
S13	-	-	-	-	-	-	-	-	-	-	-	-	0,13 ^{NS}

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; ^{NS} Teste não significativo.Nota: SVM é a abreviação para *Support Vector Machine* e RF é abreviação de *Random Forest*.**Tabela 3.8.** Valores do Teste Z, comparando as combinações para o classificador SVM.

SV M	S 1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
S1	-	6,42*	0,18 ^{NS}	15,47*	4,88**	4,21**	16,42*	14,36*	15,74*	16,36*	17,08*	16,23*	14,60*
S2	-	-	6,61*	22,07*	11,31*	2,22*	10,28*	8,11**	9,55**	10,21*	10,98*	10,08*	8,35**
S3	-	-	-	15,28*	4,70**	4,40**	16,61*	14,55*	15,92*	16,54*	17,26*	16,42*	14,78*
S4	-	-	-	-	10,48*	19,80*	32,09*	30,07*	31,42*	32,02*	32,72*	31,90*	30,30*
S5	-	-	-	-	-	9,10**	21,16*	19,15*	20,49*	21,10*	21,79*	20,97*	19,38*
S6	-	-	-	-	-	-	12,41*	10,28*	11,70*	12,34*	13,09*	12,21*	10,52*
S7	-	-	-	-	-	-	-	2,30*	0,78 ^{NS}	0,07 ^{NS}	0,77 ^{NS}	0,22 ^{NS}	2,05*
S8	-	-	-	-	-	-	-	-	1,52 ^{NS}	2,23*	3,06**	2,08*	0,26 ^{NS}
S9	-	-	-	-	-	-	-	-	-	0,71 ^{NS}	1,55 ^{NS}	0,56 ^{NS}	1,27 ^{NS}
S10	-	-	-	-	-	-	-	-	-	-	0,84 ^{NS}	0,14 ^{NS}	1,97*
S11	-	-	-	-	-	-	-	-	-	-	-	0,98 ^{NS}	2,81*
S12	-	-	-	-	-	-	-	-	-	-	-	-	1,83 ^{NS}
S13	-	-	-	-	-	-	-	-	-	-	-	-	-

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; ^{NS} Teste não significativo.Nota: SVM é a abreviação para *Support Vector Machine*.

Tabela 3.9. Valores do Teste Z, comparando as combinações para o classificador RF.

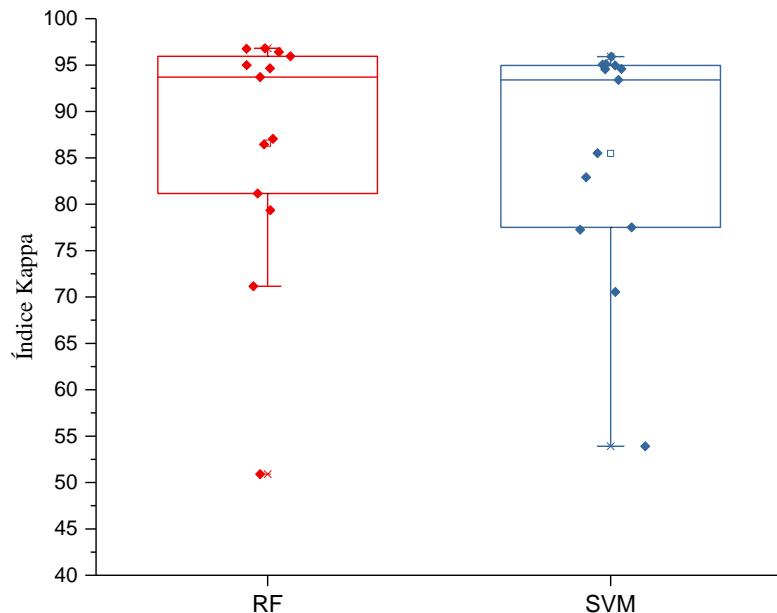
RF	S ₁	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
S1	-	6,43* *	1,40 NS	18,75* *	5,87** *	5,88** *	17,28* *	13,36* *	14,96* *	16,19* *	17,35* *	16,80* *	13,12* *
S2	-	-	5,04* *	25,41* *	12,29* *	0,55 NS	11,31* *	7,11** *	8,80** *	10,12* *	11,39* *	10,78* *	6,86**
S3	-	-	-	20,20* *	7,27** *	4,49** *	15,92* *	12,00* *	13,62* *	14,87* *	16,05* *	15,49* *	11,76* *
S4	-	-	-	-	12,69* *	24,84* *	36,18* *	32,39* *	33,95* *	35,14* *	36,25* *	35,72* *	32,16* *
S5	-	-	-	-	-	11,74* *	22,85* *	19,09* *	20,63* *	21,81* *	22,92* *	22,39* *	18,86* *
S6	-	-	-	-	-	-	11,82* *	7,65** *	9,33** *	10,65* *	11,90* *	11,30* *	7,40**
S7	-	-	-	-	-	-	-	4,51** *	2,78* *	1,35 NS	0,09 NS	0,61 NS	4,76**
S8	-	-	-	-	-	-	-	-	1,77 NS	3,20**	4,60**	3,93**	0,26 NS
S9	-	-	-	-	-	-	-	-	-	1,44 NS	2,86*	2,18*	2,03*
S10	-	-	-	-	-	-	-	-	-	-	1,43 NS	0,74 NS	3,45**
S11	-	-	-	-	-	-	-	-	-	-	-	0,70 NS	4,85**
S12	-	-	-	-	-	-	-	-	-	-	-	-	4,18**
S13	-	-	-	-	-	-	-	-	-	-	-	-	-

Obs.: * Teste Z significativo a 5%; ** Teste Z significativo a 1%; NS Teste não significativo.

Nota: RF é abreviação de *Random Forest*.

Para a comparação dos classificadores, foi notado que apenas os produtos S3, S6, S11 e S12 foram significativamente diferentes. Porém ao avaliar os classificadores de maneira geral, foi observado que existe diferença significativa a 1% de probabilidade. A Figura 3.8. mostra os níveis máximos e mínimos, que evidencia a amplitude interquartílica e a mediana dos índices Kappa para cada classificador e destaca que o RF obtém maior amplitude dos dados. Contudo, as medianas dos ambos classificadores são semelhantes (RF de 93,7% e 93,4 para o SVM). Foi notado ainda, que por meio da mediana a assimetria dos dois algoritmos é negativa, apresentando valores atípicos (outliers).

Figura 3.8. Boxplot- Relação índice Kappa com classificadores.



Nota. SVM é a abreviação para *Support Vector Machine*, RF é abreviação de *Random Forest*. Estes algoritmos foram utilizados para as classificações de uso e cobertura da terra do presente estudo.

Fonte: O autor (2017).

Os dois algoritmos de classificação apresentaram valores atípicos, (Figura 3.8.), sendo que estes valores são referentes aos dados incoerentes (S4). Pode ser percebido também, que há um alto volume de índices Kappa concentrados em valores acima de 90,0% em ambos classificadores. Estes valores são equivalentes a combinação de dados SAR e ópticas.

Diante dos resultados apresentados, é possível entender que os produtos, ou seja, a decomposição polarimétrica, relatam uma melhoria no índice Kappa da classificação de uso e cobertura da terra. Avtar, Takeuchi e Sawada (2013) relataram as mesmas conclusões, que a utilização de dados SAR polarimétricos, mostram melhora na detecção de diferenças estruturais entre dosséis de coberturas vegetais, ajudando assim o mapeamento e gestão de tais recursos.

Ainda, as decomposições incoerentes não apresentaram melhorias no índice Kappa. Sendo assim, os produtos oriundos de sensores Radar, exceto decomposições incoerentes, são valiosos para classificações de uso e cobertura da terra, principalmente quando houver escassez de cenas procedentes de sensores ópticos. Vale salientar que o modo de aquisição quádrupla ainda é experimental. Sendo sugirido, portanto, que sejam avaliados de forma isolada dados de polarização dupla.

3.4. CONCLUSÕES

Com o presente estudo foi possível concluir que os classificadores RF e SVM obtiveram valores similares de índice Kappa para os diferentes cenários analisados. Entretanto, por meio da ANOVA, foi possível notar a existência de diferença significativa entre os classificadores. Todavia, com o teste Z apenas os cenários S3, S6, S7, S11, S12 apresentaram tais diferenças.

O uso do Landsat-5 (S13) gerou um índice Kappa equivalente a 94,6% enquanto o sensor ALOS/PALSAR (S6) gerou um índice de 86,45%, e a integração (S7) dos dois sensores atingiu um índice Kappa de 96,75%, ou seja, evidenciando que a combinação dos dados ópticos com os dados SAR, trazem melhorias para fins de classificação do uso e cobertura da terra. Entretanto, utilizando somente a somatória dos produtos SAR, ocorre uma redução de até 10% no índice Kappa, o que pode ser uma alternativa para períodos sem disponibilidade de dados ópticos. Portanto, com base nos resultados obtidos, pode ser concluído que a combinação dos dados SAR e ópticos permite obter resultados significativos. Assim, é recomendada a utilização de dados SAR para classificações de uso e cobertura da terra no planalto sul catarinense, pois, há dificuldade na obtenção de imagens de sensores ópticos livres de nuvens.

Como sugestão, pode ser realizado estudos futuros nos quais a seleção de atributos dos sensores SAR possa ser aplicada para fim de classificações de uso e cobertura da terra; também podem ser realizadas novas pesquisas envolvendo diferentes métodos de classificação, como por exemplo, classificação orientada ao objeto, que em geral, tendem a apresentar melhores resultados quando comparada aos classificadores pixel a pixel.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAF. Anuário Estatístico-Associação Brasileira de Produtores de Florestas Plantadas. **Anuaário Estatístico ABRAF**, p. 146, 2013.

ALVARES, C. A. et al. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.

AVTAR, R., TAKEUCHI, W., SAWADA, H. Full polarimetric PALSAR-based land cover monitoring in Cambodia for implementation of REDD policies. **International Journal of Digital Earth**, v.6. p. 255-275, 2013.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

BOYD, D. S. e DANSON, F. M. Satellite remote sensing of forest resources: three decades of research development. **Progress in Physical Geography**, v. 29, n. 1, p. 1–26, 2005.

BOERNER, W. M. et al. On the Basic Principles of Radar Polarimetry: the Target Characteristic Polarization State Theory of Kennaugh, Huynen's Polarization Fork Concept, and its Extension to the Partially Polarized Case. **IEEE Trans. Geosci. Remote Sens.**, v. 79, n. October, p. 1538–1550, 1991.

COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37–46, 1960.

CONGALTON, R. G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. **Remote Sensing Environment**, v. 46, n. October 1990, p. 35–46, 1991.

CLOUDE, S. R.; POTTIER, E. An entropy based classification scheme for land applications of polarimetric SAR. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 1, p. 68–78, 1997.

EPIPHANIO, J. C. N. et al. Índices de vegetação no sensoriamento remoto da cultura do feijão. **Pesquisa Agropecuaria Brasileira**, v. 31, n. 6, p. 445–454, 1996.

FELDE, G. W. et al. **Analysis of Hyperion Data with the FLAASH Atmospheric Correction Algorithm**. (IEEE, Ed.) **INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM. Anais...** New York: 2003

FLORENZANO, TERESA GALLOTTI. Imagens de Satélite para Estudos Ambientais. **Oficina de Textos**, p. 1–97, 2002.

FREEMAN, A. e DURDEN, S. A three-component scattering model for polarimetric SAR data. **IEEE Transactions on Geoscience and Remote Sensing**. v.36, p. 963-973, 1998.

GALLARDO-CRUZ, J. A. et al. Predicting tropical dry forest successional attributes from space: Is the key hidden in image texture? **PLoS ONE**, v. 7, n. 2, p. 38–45, 2012.

GARCIA, C. E. et al. **Análise de Imagem STRIPMAP/TERRASAR-X para classificação da cobertura Vegetal na Porção Sudoeste Amazônia**. IX Seminário de Atualização em

Sensoriamento Remoto e Sistemas de Informações Geográficas Aplicados à Engenharia Florestal. **Anais...** p. 201-208, 2010.

GOU, S.; LI, X.; YANG, X. Coastal zone classification with fully polarimetric SAR imagery. **IEEE Geoscience and Remote Sensing Letters**, v. 13, n. 11, p. 1616–1620, 2016.

GOVERNO DO ESTADO DE SANTA CATARINA, O. Sistema de Informações Geográficas de Santa Catarina. Disponível em: <<http://sigsc.sds.sc.gov.br/>>.

HENDERSON, F. e LEWIS, A. J. **Manual of remote sensing: principles and applications of imaging radars.** 3. ed. New York: John Wiley and Sons, 1998.

HYYPPA, J. et al. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. **Forest Ecology and Management**, v. 128, n. 1–2, p. 109–120, 2000.

JANOTH, J. et al. **Procedimentos Baseados em Segmentação para a Análise de Mudanças e Classificação Florestais com Dados de Satélite de Alta Resolução.** cap. 9. In: BLASCHKE, T.; KUX, H. Sensoriamento Remoto e SIG Avançados. 2. ed.- São Paulo: Oficina de textos. 2007.

KUNCHEVA, L. I. **Combining pattern classifiers: methods and algorithms.** New Jersey: John Wiley & Sons, 2004.

LAURIN, G. V. et al. Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa. **International Journal of Applied Earth Observation and Geoinformation**, v. 21, n. 1, p. 7–16, 2013.

LIESENBERG, V. et al. Evaluating Moisture and Geometry Effects on L-Band SAR Classification Performance Over a Tropical Rain Forest Environment. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 9, n. 12, p. 5357–5368, 2016.

MARTINS, F. DA S. R. V. et al. Sensitivity of ALOS/PALSAR imagery to forest degradation by fire in northern Amazon. **International Journal of Applied Earth Observation and Geoinformation**, v. 49, p. 163–174, 2016.

MENESES, P. R.; ALMEIDA, T. DE. **Introdução ao Processamento de Imagens de Sensoriamento Remoto.** CNPq ed. Brásilia: [s.n.].

MISHRA, V. N. et al. Dual-polarimetric C-band SAR data for land use/land cover classification by incorporating textural information. **Environmental Earth Sciences**, v. 76, n. 1, p. 1–16, 2017.

MITCHARD, E. T. A. et al. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: Overcoming problems of high biomass and persistent cloud. **Biogeosciences**, v. 9, n. 1, p. 179–191, 2012.

MUTHUKUMARASAMY, I.; SHANMUGAM, R. S.; KOLANUVADA, S. R. SAR polarimetric decomposition with ALOS PALSAR-1 for agricultural land and other land

use/cover classification: case study in Rajasthan, India. **Environmental Earth Sciences**, v. 76, n. 13, p. 1–13, 2017.

NASCIMENTO, R. F. F. et al. **O algoritmo Support Vector Machine (SVM): avaliação da separação ótima de classes em imagens CCD-CBERS-2.** (INPE, Ed.)Anais XIV Simpósio Brasileiro de Sensoriamento Remoto. Anais...Natal - BR: 2009

NOVO, E. M. L. DE M. **Sensoriamento Remoto: Princípios e Aplicações.** São Paulo: Edgard Blucher, 2010.

OLIVER,C.J e QUEGAN. Understanding Synthetic Aperture Radar Images. **Artech House.** p.480,1998.

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento Remoto da Vegetação.** 2. ed. São Paulo: Oficina De Texto, 2012.

POPE, K. O.; REY-BENAYAS, J. M.; PARIS, J. F. Radar remote sensing of forest and wetland ecosystems in the Central American Tropics. **Remote Sensing of Environment**, v. 48, n. 2, p. 205–219, 1994.

RANSON, K. et al. Characterization of forests in Western Sayani mountains, Siberia from SIR-C SAR data. **Remote Sensing of Environment.** v.75.p.188-200. 2001

SAATCHI, S. S. et al. **Estimation of Forest Biomass Change from Fusin of Radar and Lidar Measurements.2010.** Disponível em: <<https://pt.slideshare.net/grssieee/estimation-of-forest-biomass>>. Acesso em: 1 set. 2017.

SINCLAIR, G. The Transmission and Reception of Elliptically Polarized Waves. **Proceedings of the IRE**, v. 38, n. 2, p. 148–151, 1950.

ULABY, F.; ELACHI, C. **Radar Polarimetry for Geoscience Applications.** Artech Hou ed. [s.l: s.n.].

VAN DER LINDEN, S. et al. The EnMAP-box-A toolbox and application programming interface for EnMAP data processing. **Remote Sensing**, v. 7, n. 9, p. 11249–11266, 2015.

VAPNIK, V. N. **The nature of statistical learning theory.** Springer- ed. New York. v. 8, 1995

NGUYEN, L.V. et al. Estimation of Tropical Forest Structural Characteristics Using ALOS-2 SAR Data. **Advances in Remote Sensing**, v. 5, n. 2, p. 131–144, 2016.

WATZLAWICK, L. F. **Estimativa de biomassa e carbono em floresta ombrófila mista e plantações florestais a partir de dados de imagens do satélite Ikonos II.** Curitiba: UFPR, Tese (Doutorado em Ciências Florestais) – Setor de Ciências Agrárias, Universidade Federal do Paraná.p.120, 2003.

WIJAYA, A.; LIESENBERG, V.; GLOAGUEN, R. Retrieval of forest attributes in complex successional forests of Central Indonesia: Modeling and estimation of bitemporal data. **Forest Ecology and Management**, v. 259, n. 12, p. 2315–2326, 2010.

WIJAYA, A.; MARPU, P. R.; GLOAGUEN, R. Discrimination of peatlands in tropical swamp forests using dual-polarimetric SAR and Landsat ETM data. **International Journal of Image and Data Fusion**, v. 1, n. 3, p. 257–270, 2010.

CAPÍTULO 4

Análise do Tempo de Execução de Dois Algoritmos Não-Paramétricos Para Classificação de Uso e Cobertura da Terra no Planalto Sul Catarinense

RESUMO

O objetivo do presente estudo foi avaliar o tempo computacional de dois classificadores não-paramétricos *Support Vector Machine* (SVM), e o *Random Forest* (RF) para fins de classificação de uso e ocupação da terra, tanto para diferentes tamanhos de amostras, quanto para diferentes números de variáveis de entrada. A área de estudo está localizada no planalto sul catarinense envolvendo expressivamente espaços rurais. Foram consideradas seis classes (i.e. Água, Solo Exposto, Área Urbanizada, Campo Nativo, Floresta Nativa e Reflorestamento). Cada classe foi constituída por mil pixels, totalizando assim 6.000 pixels. Foram utilizados dois sensores, sendo um óptico (Landsat-5/TM) e outro de micro-ondas (ALOS/PALSAR). As avaliações do tempo computacional foram divididas em dois grupos, uma realizada a partir do tamanho de amostra de entrada e outra com diferentes combinações de cenas. A primeira foi realizada utilizando somente o sensor óptico. Logo, diferentes tamanhos de amostras de treinamento e validação foram classificados, sendo elas 10% treinamento e 90% validação, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10. A segunda avaliação foi realizada utilizando os dois sensores, ou seja, diferentes cenários foram realizados, com os produtos de decomposições (Incoerentes, Coerentes e Texturas) da cena do SAR e com as reflectâncias do sensor óptico. Portanto, foram geradas 13 combinações de diferentes variáveis. Para avaliação de qualidade dos mapas gerados, foi utilizado o índice Kappa. Já para a avaliação do tempo de execução computacional, foram anotados o tempo de término do processamento oferecido pelo aplicativo EnMAPToolBox, o qual foi utilizado para a classificação. Sendo assim, pode ser concluído que além da avaliação de qualidade, ou seja, acurácia da classificação, o tamanho da amostra de treinamento interfere no tempo computacional. Por outro lado, o número de variáveis, não intervém no tempo computacional. Os classificadores apresentaram diferenças, sendo que o RF apresentou menores tempos computacionais e maiores índices Kappa, ao ser comparado com o SVM. Assim, pode ser recomendada e a utilização do classificador RF para futuras classificações no planalto sul catarinense.

Palavras-chave: Classificação supervisionada, Demanda computacional, tempo de classificação.

CHAPTER 4

Evaluating Computational Time Demand for Two Non-Parametric Algorithms for Classification of Land Use and Land Cover in the Santa Catarina State Southern Plateau

ABSTRACT

The goal of this study was to evaluate the computational time of both Support Vector Machine (SVM), e o Random Forest (RF) classifiers for land use and land cover classification. We exlore both training dataset size as well as number of variables coming from the remotely sensed data. Our study area is a rural area located in the Santa Catarina Southern Plateau. We consider six land use and land cover (LULC) classes (i.e. Water, bare soil, urban area, grasslands, native forest and planted forest). A total of 6,000 pixels were considered, i.e., 1,000 pixels for each LULC class. We evaluate data from two sensors, one optical (Landsat-5/TM) and one microwave (ALOS/PALSAR). The computacional time demand was divided into two groups, one from the data input size and other with different scenarios of combining scenes and derived products. The first one was performed only with Landsat. Different training dataset sizes were considered, being 10% training and 90% validation. Thus, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 and 90/10, respectively. The second evaluation was performed using 13 different scenarios, with Landsat reflectance and ALOS/PALSAR products (coherent data, incoherent data, and texture) used alone and combined. The quality of the generated maps was evaluated with Kappa index. The computational time demand was considered by reports generated with the EnMAPToolBox. The results showed that the training dataset size interferes significantly in the computational time demand. On the other hand, the number of imagens and products as variables does not interfer in the computational time demand. RF showed the lowest computacional time demand and the highest Kappa indices when compared to SVM. Based on the achieved results, we recommend the use of RF for further LULC approaches in the Santa Catarina State Southern Plateau.

Keywords: Supervised classification; Computational demand; Classification time.

4.1. INTRODUÇÃO

Dados de sensoriamento remoto proporcionam a extração de informações, que podem ser explicadas em diferentes formatos (i.g. tabelas, gráficos e mapas) (FIGUEIREDO, 2005, MENESSES E ALMEIDA, 2012). Costa et al. (1996) afirmaram que as geotecnologias disponibilizam uma série de ferramentas que servem para auxiliar a gestão do uso e ocupação da terra. Estas ferramentas de mapeamento e classificação do uso e ocupação da terra, estão cada vez mais úteis em diferentes análises e finalidades.

De acordo com Meneses e Almeida (2012) há diversos métodos de classificações de dados remotamente situados. Estes podem ser divididos em: método de classificação supervisionada ou não-supervisionada, paramétricas e não-paramétricas. Existem também métodos para classificar cenas por pixel, por regiões ou por objeto. A escolha de classificadores que se adequam aos conjuntos de dados, é fundamental para geração de produtos que possam ser utilizados com confiabilidade e acurácia.

Segundo Novo (2010) para verificar a qualidade do processo de classificação, são avaliadas a exatidão e a acurácia dos resultados. De tal modo, Congalton e Green (1999) definiram que a exatidão está relacionada a acurácia das classes coletadas na imagem em relação à realidade. Esta avaliação é realizada por meio da matriz de confusão, da qual podem ser calculadas medidas descritivas (i.e., exatidão geral, exatidão do usuário e exatidão do produtor). Além destas métricas de exatidão, é habitualmente aplicado o índice Kappa. Entretanto, Powers (2012) realizou uma investigação no qual questiona o valor Kappa. Lee et al., (2001) e Lim et al., (2007) afirmaram que é uma medida largamente utilizada em classificações de diversos estudos.

Outra forma de avaliar o resultado gerado por um classificador, além das medidas de acurácias, é a avaliação do tempo de execução em que o *software* exerce ao gerar o produto final. Meneses e Almeida (2012) afirmaram que dados provenientes do sensoriamento remoto podem requerer significativos recursos computacionais. De acordo com Medina e Ferting (2006) existe uma relação direta entre o tempo de resposta e o tamanho da amostra de entrada, entretanto, existe um fator mais importante que afeta o tempo de execução: a complexidade do algoritmo utilizado. Sendo assim, para serem comparados dois algoritmos de classificação, deve ser utilizado o mesmo sistema de *hardware* e *software*.

Deste modo, o objetivo do trabalho foi avaliar o tempo computacional de dois classificadores não-paramétricos para classificação de uso e cobertura terra para diferentes números de variáveis de entrada.

4.2. MATERIAL E MÉTODOS

4.2.1. Caracterização da área de estudo

O estudo envolveu sete municípios da serra catarinense (Lages, Curitibanos, Capão Alto, São José do Cerrito, Correia Pinto, Ponte Alta e São Joaquim). A região de estudo está localizada na longitude $50^{\circ} 39' 40,61''$ O. e latitude $28^{\circ}28' 59,52''$ S., abrangendo 3.085,6612 km², conforme apresentado na Figura 4.1.

Segundo Alvares et al. (2013) o clima é classificado como Cfb (temperado subtropical). Estas áreas possuem um forte perfil agrícola, somando-se a expressividade na produção florestal, fator decisivo para alavancagem e consolidação dos segmentos de celulose e papel, madeireiro e moveleiro da macrorregião (ABRAF, 2013).

4.2.2. Sensores Orbitais

O estudo foi realizado com dois sensores, sendo um óptico (Landsat-5/TM) e outro de micro-ondas (ALOS/PALSAR). Os sensores estudados apresentam as seguintes características:

- **Landsat-5/TM:** É administrado pela NASA e gerenciado pelo *U.S. Geological Survey* (USGS's), no *Center for Earth Resources Observation and Science*. Embora o sensor esteja desativado desde 2013, os dados históricos da série Landsat são constituídos como uma importante fonte de dados. O Landsat-5 apresenta características específicas conforme a Tabela 4.1. Foram utilizadas duas cenas da órbita 221, referente ao dia 20 de março de 2009, sendo uma do ponto 79 e outra do 80. As cenas foram adquiridas pelo Earth Explorer (site: www.earthexplorer.usgs.gov).

A cena do Landsat-5 apresenta 185x185 km e para o presente estudo, foi realizado um recorte que envolve uma área coincidente com as cenas do sensor ALOS/PALSAR.

Tabela 4.1. Resoluções espetrais e espacial do Satélite Landsat -5.

Sensor	Região do Espectro	Intervalo (μm)	Res. Espacial (m)
TM (Thematic Mapper)	Azul	0,45 -0,52	30
	Verde	0,50 -0,60	
	Vermelho	0,63-0,69	
	Infravermelho Próximo	0,76-0,90	
	Infravermelho Médio	1,55 -1,75	
	Infravermelho Termal	10,4-12,5	120
	Infravermelho Médio	2,06-2,35	30

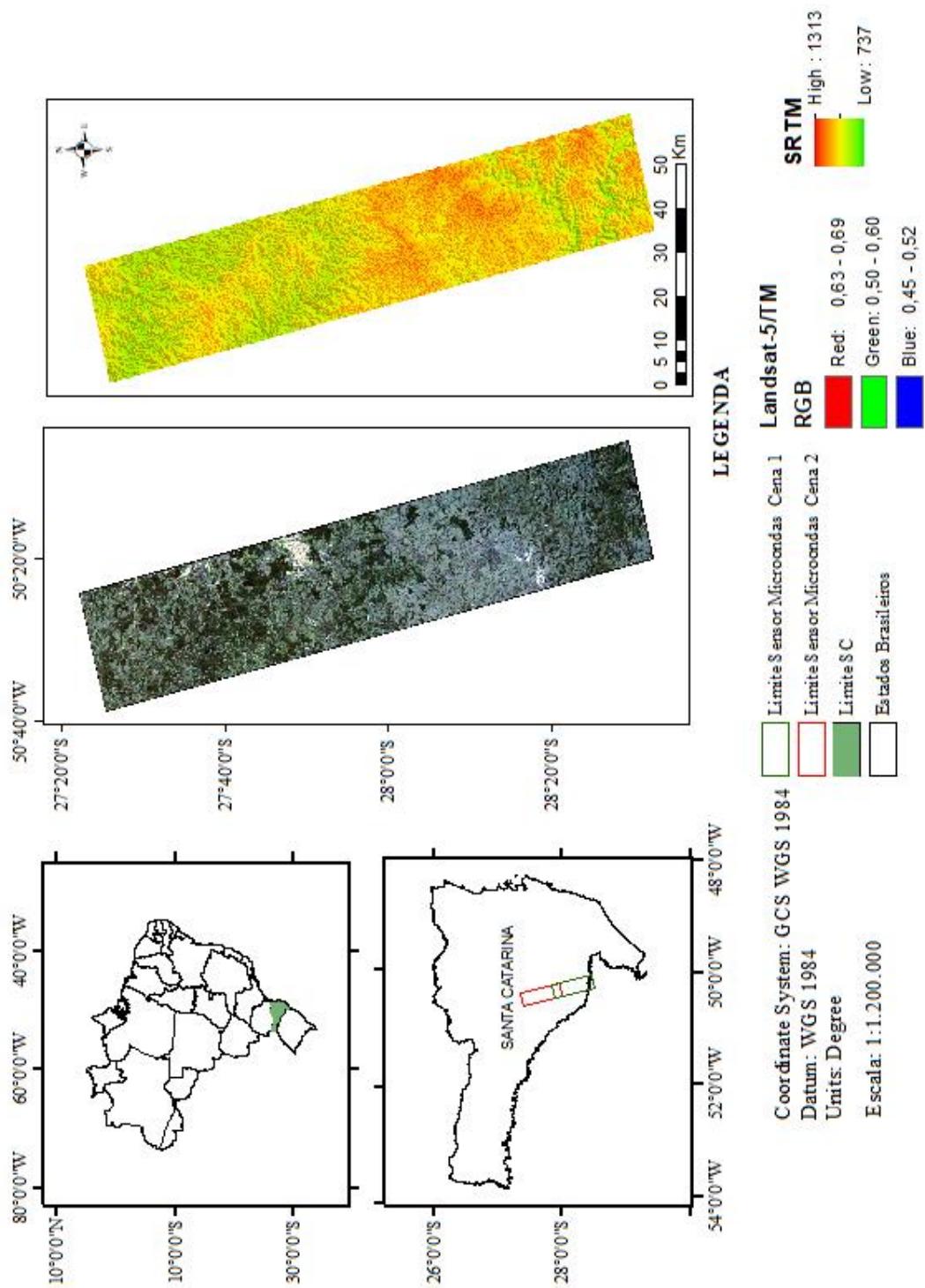
Fonte: Novo (2010).

- **ALOS/PALSAR:** Sensor radar de abertura sintética (SAR) na banda-L, proporciona observação detalhada, bem como interferometria de repetição. Foi desenvolvido para contribuir nos campos de mapeamento, observação regional de cobertura da terra, monitoramento de desastres e levantamento de recursos naturais. O ALOS é uma missão da Agência de Exploração Aeroespacial do Japão (JAXA). As cenas utilizadas neste estudo apresentam resolução espacial de 30 metros e com quatro polarizações (HH, VV, HV e VH) adquiridas de modo experimental na região sul do Brasil. Foram utilizadas duas cenas do dia três de novembro de 2009, as quais foram adquiridas em formato *Slant-range Singlelook Complex* (SLC), no nível de processamento 1.1.

4.2.3. Pré-processamento e Processamento das cenas

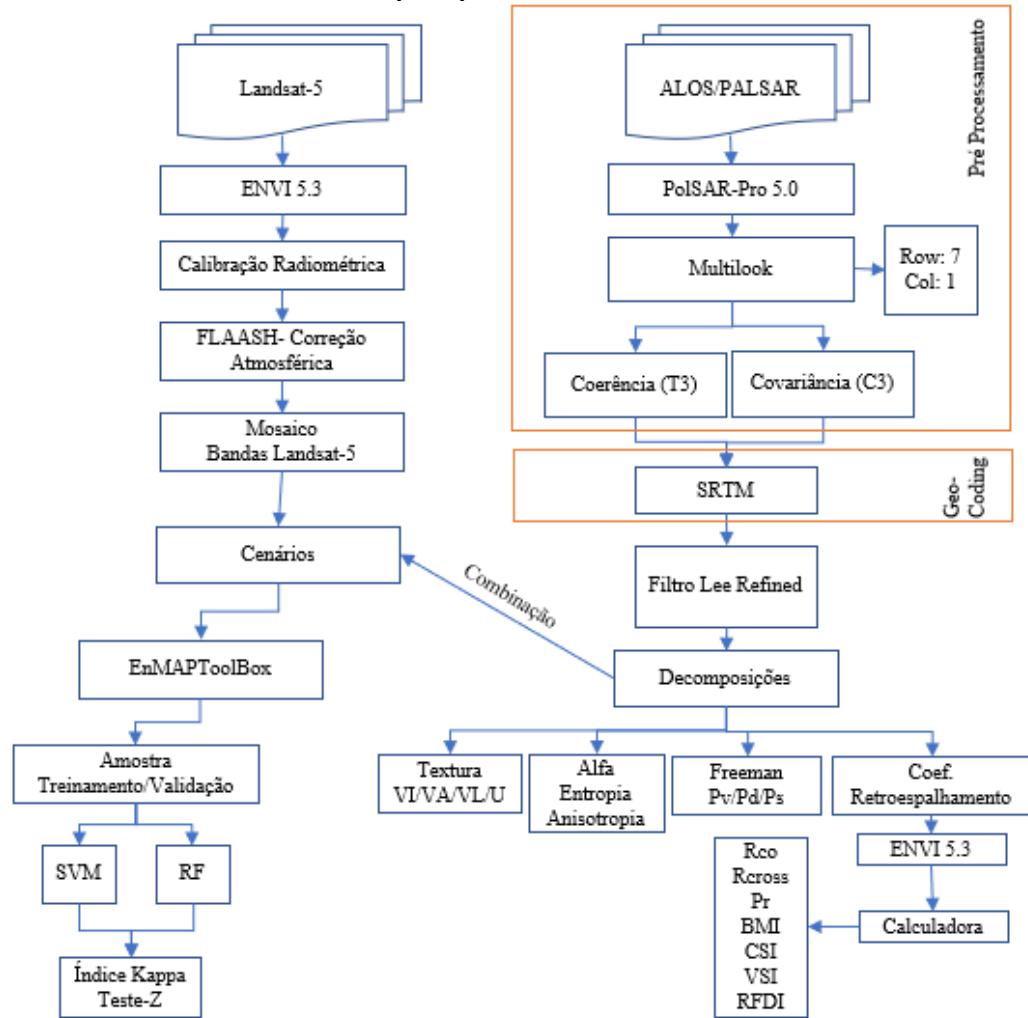
Foram realizadas diversas etapas de pré-processamento das cenas. Priori para a cena do Landsat-5/TM foi realizada a conversão de níveis digitais para radiância ($\text{Wm}^{-2}\text{sr}^{-1}\text{nm}^{-1}$) e posteriormente a conversão para reflectância de superfície, considerando a minimização dos efeitos atmosféricos através do “*fast line of sight atmospheric analysis of spectral hypercubes*” (Flaash) proposto por Felde et al. (2003). Para o ALOS/PALSAR nível 1.1, foram realizados procedimentos para redução do efeito *speckle*, ou seja, diminuição da largura da distribuição na resolução espacial pelo mesmo fator, foi realizada ainda a decomposição polarimétrica, conforme apresentado na Figura 4.2, onde é explanado um esquema do processamento realizado para ambos sensores.

Figura 4.1. Localização da área de estudo.



Fonte: O autor (2017).

Figura 4.2. Processamento realizado para processamento das cenas dos sensores utilizados.



Fonte: O autor (2017).

De posse dos dados de amplitude e fase polarimétrica foram calculadas: incoerência, coerência e textura. A decomposição apresenta como objetivo o mecanismo médio de espalhamento como uma soma de elementos independentes, os quais foram associadas a um mecanismo físico. Maiores informações sobre os processamentos dos dados podem ser encontradas nas referências apresentadas nas tabelas a seguir. Na Tabela 4.2. são evidenciadas as decomposições incoerentes, as quais representam dados gerados a partir dos coeficientes de retroespalhamento.

Tabela 4.2. Atributos Incoerentes

Incoerente	Equação	Descrição	Autores
Coeficiente Retroespalhamento $\sigma_{HH}^0, \sigma_{HV}^0,$ σ_{VH}^0 e σ_{VV}^0	$\sigma_{ij}^0 = 10 \frac{\sigma_{ij}^{0 (\text{dB})}}{10}, \text{ onde}$ $\sigma_{ij}^{0 (\text{dB})} = 10 \times \log_{10} (I^2 + Q^2)$ $+ FC - A$	Indica a orientação dos componentes da floresta (i.e., mais troncos produzem maior retroespalhamento vertical)	Henderson e Lewis (1998) Schimada et al. (2009)
Relação de Co-Polarização	$R_{CO} = \frac{\sigma_{VV}^0}{\sigma_{HH}^0}$	Destaca diferentes orientações verticais e horizontais derivadas dos aspectos estruturais da vegetação	Henderson e Lewis (1998)
Relação de Polarização Cruzada	$R_{Cross} = \frac{\sigma_{HV}^0}{\sigma_{HH}^0}$	Sensível à dispersão volumétrica da floresta para apoiar a classificação e reduzir os efeitos topográficos no retroespalhamento	Henderson e Lewis (1998)
Potência Total	$Pr = \sigma_{HH}^0 + \sigma_{VV}^0 + 2 \times \sigma_{HV}^0$	Energia total de retroespalhamento derivada da cobertura vegetal	Boerner et al. (1991)
Índice de Biomassa	$BMI = \frac{\sigma_{HH}^0 + \sigma_{VV}^0}{2}$	Quantidade relativa de componente lenhoso em comparação com biomassa foliar	Pope et al. (1994)
Índice de estrutura de Copa	$CSI = \frac{\sigma_{VV}^0}{\sigma_{HH}^0 + \sigma_{VV}^0}$	Presença relativa de dispersores verticais (troncos e caules) de vegetação	Pope et al. (1994)
Índice espalhamento de Volume	$VSI = \frac{\sigma_{HV}^0}{\sigma_{HV}^0 + BMI}$	Indicador de espessura do dossel ou densidade relativa	Pope et al. (1994)
Índice degradação de florestas	$RFDI = \frac{\sigma_{HH}^0 - \sigma_{HV}^0}{\sigma_{HH}^0 + \sigma_{HV}^0}$	Relação projetada para avaliar a força do mecanismo de duplo salto, o que é útil para diferenciar a vegetação	Saatchi et al. (2010); Mitchard et al. (2012)

Nota: Fc = fator de calibração (-83); A= fator de conversão (32). Os valores obtidos em escala logarítmica foram convertidos em escala linear para os cálculos dos atributos.

Fonte: O autor (2017).

As decomposições coerentes consistem em um espalhamento determinístico, sendo utilizando normalmente com a matriz de Sinclair [S], o qual caracteriza os estados de polarização (SINCLAIR, 1950; ULABY e ELACHI, 1990 ; CLOUDE, 2009; LEE e POTTIER, 2009). Conforme apresentado na Tabela 4.3. as equações se relacionam com a onda espalhada ou recebida à onda incidente ou transmitida (ULABY e ELACHI, 1990).

Tabela 4.3. Atributos Coerentes

Coerente	Equação	Descrição	Autores
Diferença de Fase	$\Delta\varphi_{HH} - VV = \arg(S_{HH}S_{VV}^*)$ $\Delta\varphi_{HV} - HV = \arg(S_{HH}S_{HV}^*)$ $\Delta\varphi_{HV} - VV = \arg(S_{HV}S_{VV}^*)$	Indicação da estrutura e quantidade de biomassa	Henderson e Lewis (1998)
Entropia	$H = - \sum_{i=1}^3 p_i \log_3(p_i); p_i = \frac{\lambda_i}{\sum_{j=1}^3 \lambda_j}$	Relacionado à complexidade da estrutura da floresta. Floresta mais complexa e diversificada apresenta alta H, baixa A e perto de 45°	Cloude e Pottier (1997)
	$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3}$		
	$\alpha = \sum_{i=1}^3 p_i \alpha_i$		
Contribuição de dispersão de volume para potência total	$P_V = \frac{8 f_V}{3}$	Proporção da retroespalhamento volumétrica associada à estrutura da floresta e ao conteúdo de biomassa	Freeman e Durden (1998)
Contribuição de dispersão de Double-bounce para a potência total	$P_d = f_d(1 + \beta ^2)$	Indicação da abertura do dossel, densidade e número de árvores (troncos)	Freeman e Durden (1998)
Contribuição de dispersão de Surface potência total	$P_s = f_s(1 + \beta ^2)$	Relacionado à abertura do dossel	Freeman e Durden (1998)

Nota: Fc = fator de calibração (-83); A= fator de conversão (32). Os valores obtidos em escala logarítmica foram convertidos em escala linear para os cálculos dos atributos.

Fonte: O autor (2017).

Como apresentado na Tabela 4.4. os atributos de texturas trabalham com a intensidade do sinal, assim distinguindo-se em diferentes momentos. De acordo com Wijaya et al. (2010), Gallardo-Cruz et al. (2012) e Laurin et al. (2013), as informações sobre a textura melhoram o resultado das classificações em ambientes tropicais complexos.

Tabela 4.4. Atributos Texturas

Textura	Equação	Descrição	Autores
VI	$VI_{xy} = \frac{\sum_{i,j} (I_{ij})^2 / m}{\left(\sum_{i,j} I_{ij} / m \right)^2} - 1 \cdot (i, j)$	Uma proporção da média da intensidade quadrada para a intensidade média ao quadrado.	Oliver and Quegan, 1998
VA	$VA_{xy} = \frac{\sum_{i,j} (I_{ij})^2 / m}{\left(\sum_{i,j} A_{ij} / m \right)^2} - 1 \cdot (i, j)$	Uma proporção de intensidade média para a amplitude média quadrática.	Oliver and Quegan, 1998
VL	$VL_{xy} = \left[\sum_{i,j} \ln(I_{ij})^2 / m \right] \cdot \left[\sum_{i,j} \ln(I_{ij}) / m \right]^2$	A diferença entre o valor médio do logaritmo de intensidade quadrado e o quadrado do logaritmo de intensidade média.	Oliver and Quegan, 1998
U	$U_{xy} = \left[\sum_{i,j} \ln(I_{ij}) / m \right] \cdot \ln \left[\sum_{i,j} (I_{ij}) / m \right]$	Uma medida de log normalizada de textura. Calcula a distribuição de K.	Oliver and Quegan, 1998

Note: XY represents either HH or HV.

4.2.4. Descrição das classes de uso e cobertura da terra

Em aplicativo GIS foi realizada a coleta das amostras das diferentes classes de uso e cobertura da terra. Este procedimento foi realizado de forma visual. Para a coleta, foi utilizada a cena do Landsat-5 e para a verdade de campo, foram utilizadas imagens do Google Earth e as imagens ortorretificadas do levantamento aéreo, disponibilizadas pelo Estado de Santa Catarina Governo do Estado de Santa Catarina (2017). As amostras foram coletadas em formato *shapefile* e posteriormente transformada em Região de Interesse (ROI). Estas transformações foram realizadas no aplicativo ENVI. As seis classes de uso e cobertura da terra podem ser observadas na Tabela 4.5.

Tabela 4.5. Descrição das Classes de Uso e cobertura da terra.

Floresta Nativa	Áreas cobertas por vegetação de porte florestal nativa.
Reflorestamento	Área cobertas por vegetações de porte florestal plantado (reflorestamentos de eucalipto e pinus).
Solo Exposto	Área em que houve corte raso da vegetação natural e/ou colheita florestal e/ou agronômica.
Água	Áreas como rios, lagos, represas e entre outras.
Área Urbanizada	Áreas com feições e características de ocupações urbanas observadas nas imagens.
Campo Nativo	Área cobertas por vegetações rasteira nativa e/ou plantada (manejada).

Fonte: O autor (2017).

4.2.5. Classificação dos dados remotamente situados

As classificações foram realizadas no *software EnMAPToolBox* (VAN DER LINDEN et al..2015). Posteriormente foram realizadas as avaliações de acurácia, tanto para os classificadores como para os tamanhos de amostras e cenários (Números de variáveis).

Para a determinação do tamanho de amostras de treinamento e validação, foram geradas amostras aleatórias simples, sendo eles 10% para treinamento e 90% para validação (10/20). Em outras palavras, de 1.000 pixels de cada classe, 100 pixels foram selecionados para treinamento e os 900 pixels remanescentes para validação. Portanto, foram geradas amostras para 10/90, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20 e 90/10, respectivamente. Já para a avaliação do tempo de execução de n variáveis, foram propostos diferentes cenários, integrando cenas do Landsat-5/TM com as decomposições do ALOS/PALSAR. Os cenários estão explanados na Tabela 4.6 que expõe os diferentes cenários, variáveis e número total de variáveis para cada cenário.

Tabela 4.6. Combinação dos cenários para classificação de uso e ocupação do solo.

Cenário	Descritores	Variáveis	Nº variáveis
S1	Coerentes	ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} , H, A, α , Pv, Pd, Os, Pedestal, α , RVI e Autovalores (11, 12, 13)	14
S2	Coerentes + Texturas	ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} , H, A, α , Pv, Pd, Os, Pedestal, α , RVI, Autovalores (11, 12, 13), $VI_{(HH, HV, VH \text{ e } VV)}$, $VA_{(HH, HV, VH \text{ e } VV)}$, $VL_{(HH, HV, VH \text{ e } VV)}$ e $U_{(HH, HV, VH \text{ e } VV)}$	30
S3	Incoerentes + coerentes	σ_{HH} , σ_{HV} , σ_{VH} , σ_{VV} , R_{co} , P_T , BMI, CSI, VSI, RFDI, ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} , H, A, α , Pv, Pd, Os, Pedestal, α , RVI e Autovalores (11, 12, 13)	24
S4	Incoerentes	σ_{HH} , σ_{HV} , σ_{VH} , σ_{VV} , R_{co} , P_T , BMI, CSI, VSI e RFDI.	10
S5	Incoerentes + Texturas	Incoerente + Textura	26
S6	SAR total	Σ Coerentes + Textura + incoerentes	40
S7	SAR+ Óptica	Σ Coerentes + Textura + incoerentes + Ópticos	46
S8	Óptica+ Incoerentes	Σ Óptica + Incoerente	16
S9	Óptica + Incoerentes + Texturas	Σ Óptica + Incoerente + Textura	32
S10	Óptica + Incoerentes + Coerentes	Σ Óptica + Incoerente+ Coerente	30
S11	Óptica + Coerentes + Texturas	Σ Óptica + Coerente + Textura	36
S12	Óptica + Coerentes	Σ Óptica+ Coerente	20
S13	Dados ópticos	$\rho 485, \rho 560, \rho 660, \rho 830, \rho 1625, \rho 2215$	6

Nota: ϕ_{HH-VV} , ϕ_{HH-HV} , ϕ_{HV-VV} = Diferença de fase; H= Entropia A= anisotropia; α =Alfa médio; Pv= Contribuição de dispersão de volume; Pd= Contribuição de dispersão de Double-bounce; Ps= Contribuição de dispersão de Surface; σ = Retroespalhamento; R_{co} = Relação de Co-Polarização; R_{cross} = Relação de Polarização Cruzada; P_T = Potência Total; BMI= índice de Biomassa; CSI= índice de estrutura de Copo; VSI= índice de espalhamento de Volume; RFDI= índice de degradação de florestas.

Fonte: O autor (2017).

As avaliações da qualidade da classificação foram geradas através das matrizes de confusão e do índice Kappa, sugerido por Cohen (1960) e Congalton (1991). As significâncias dos classificadores e dos tamanhos de amostras e cenários, foram realizadas pelo Teste-Z, cujos valores críticos para 95% e 99% são de 1,96 e 2,57, respectivamente. Para a avaliação do tempo de execução dos classificadores em relação ao tamanho de amostra de entrada e o *n* variáveis, foram computados o tempo em que o próprio aplicativo ofereceu ao final de cada etapa (i.e. Parametrização e Classificação).

As classificações dos dados remotamente situados, foram realizadas empregando os seguintes algoritmos: *Support Vector Machine* (SVM) e *Random Forest* (RF). O SVM foi introduzido por meio da teoria estatística de aprendizagem assumindo como apoio a superfície de decisão, conhecida como hiperplano ideal, o qual elevar ao máximo os limites de separação de classes (VAPNIK, 1999; NASCIMENTO et al., 2009). Já o RF é um algoritmo proposto por Breiman (2001) é fundamentado em árvores de decisões, “Conjunto” (ensemble), que de acordo com Kuncheva (2004) são métodos que geram muitas classificações e combinam seus resultados, priorizando os melhores resultados.

Por meio das avaliações de qualidade e do tempo de execução, foram realizadas duas análises de correlação de Spearman, ou seja, análise de correlação não-paramétricas, pois os dados não se apresentaram com distribuição normal. A análise “a” trata-se da comparação entre o tamanho de amostras, enquanto a análise “b” está relacionada aos números de variáveis.

a. Tamanho de amostras (N), índice Kappa de RF (KNRF) e SVM (KNSVM), Tempo de parametrização de RF (TPNRF) e SVM (TPNSVM) e tempo de classificação de RF (TCNRF) e SVM (TCNSVM); e

b. N° de variáveis (S), índice Kappa de RF (KSRF) e SVM (KSSVM), Tempo de parametrização de RF (TPSRF) e SVM (TPSSVM) e tempo de classificação de RF (TCSR) e SVM (TCSSVM).

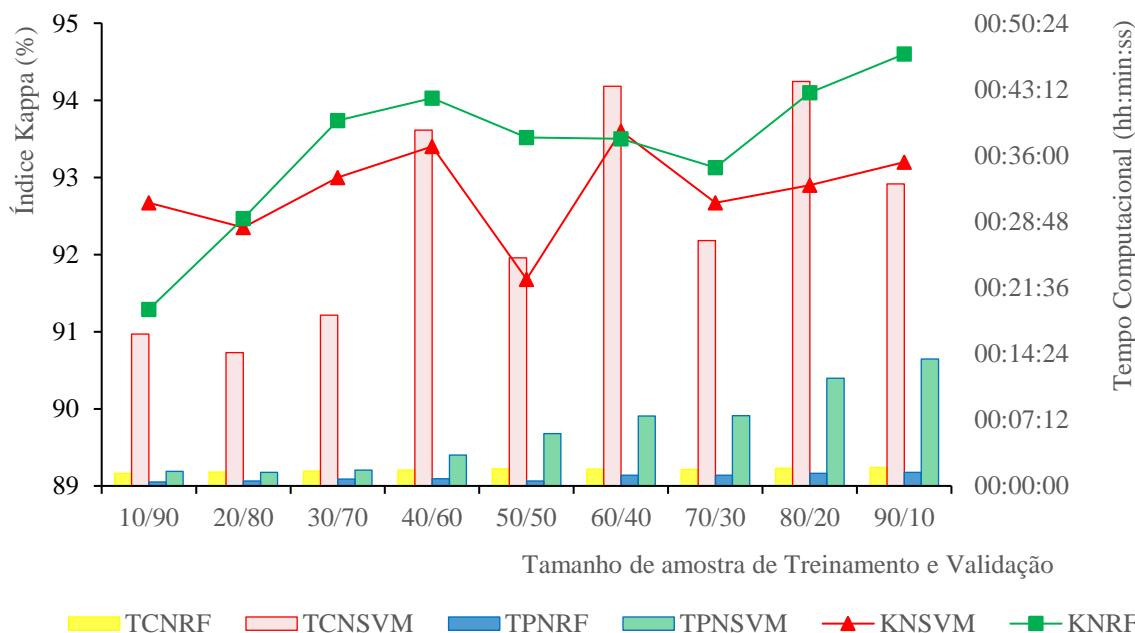
A fim de comparar o tempo de execução computacional dos dois classificadores pelo mesmo *software*, foi definida a utilização de um computador de uso comum, apresentando as seguintes configurações: processador Intel® Core™ i3- 3110M CPU @ 2.40GHz, com memória instalada (RAM) de 4,00 GB e um sistema operacional de 64 Bits. Esta configuração é adotada pela Secretaria de Ensino do Estado de São Paulo.

4.3. RESULTADOS E DISCUSSÃO

4.3.1. Tempo computacional para diferentes tamanhos de amostra de treinamento e validação

Os tempos de execuções dos classificadores variaram conforme o tamanho da amostra de entrada. A Figura 4.3., mostra que o classificador SVM, demandou maior tempo computacional para a classificação e parametrização. Por outro lado, o classificador RF, obteve os melhores índices Kappa, e apresentou ainda menor esforço computacional.

Figura 4.3. Tamanho de amostra de treinamento e validação em relação ao tempo computacional e índice Kappa.



Nota: Tempo de parametrização de *Random Forest* (TPNRF) e *Support Vector Machine* (TPNSVM). Tempo de classificação de RF (TCNRF) e SVM (TCNSVM). Índice Kappa de RF (KNSRF) e SVM (KSSVMM).

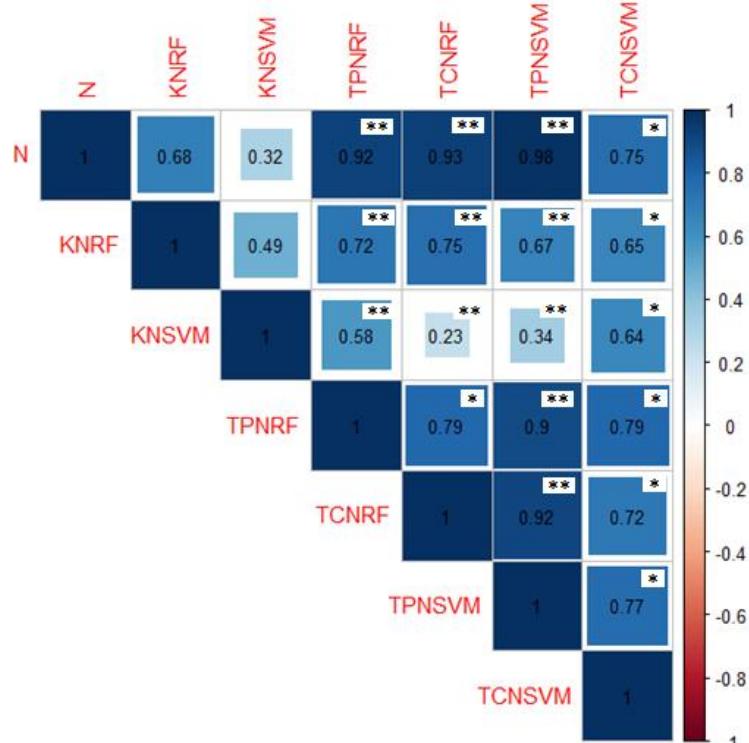
Fonte: O autor (2017).

A correlação de Spearman, apresentada na Figura 4.4., mostrou que o tamanho amostral possui uma correlação direta com o esforço computacional, em ambos classificadores. Assim, quanto maior a amostra de treinamento, maior o tempo computacional, corroborando com a teoria de Medina e Ferting (2006). Estes autores afirmaram que existe uma relação direta entre o tempo de execução e o tamanho da amostra de entrada.

Foi notado também que os índices Kappa apresentaram correlações positivas e significativas com o tempo de execução dos algoritmos. Entretanto, o RF obteve correlações positivas maiores que o SVM, como apresentado na Figura 4.4. Conforme apresentado na Figura 4.3, o SVM exerce um tempo de execução computacional maior que o RF.

Ao relacionar o tempo de parametrização com o tempo de classificação, estes apresentaram correlação positiva sendo significativo a 5% de probabilidade de erro. Portanto pode ser findado que o tamanho das amostras, ou seja, quantidade de pixels amostrados, demanda um maior tempo de parametrização e consequentemente o tempo de classificação de ambos classificadores.

Figura 4.4. Correlação de Spearman para Níveis de Tratamento



Obs.: * Teste significativo a 5%; ** Teste significativo a 1%; ___Teste não significativo.

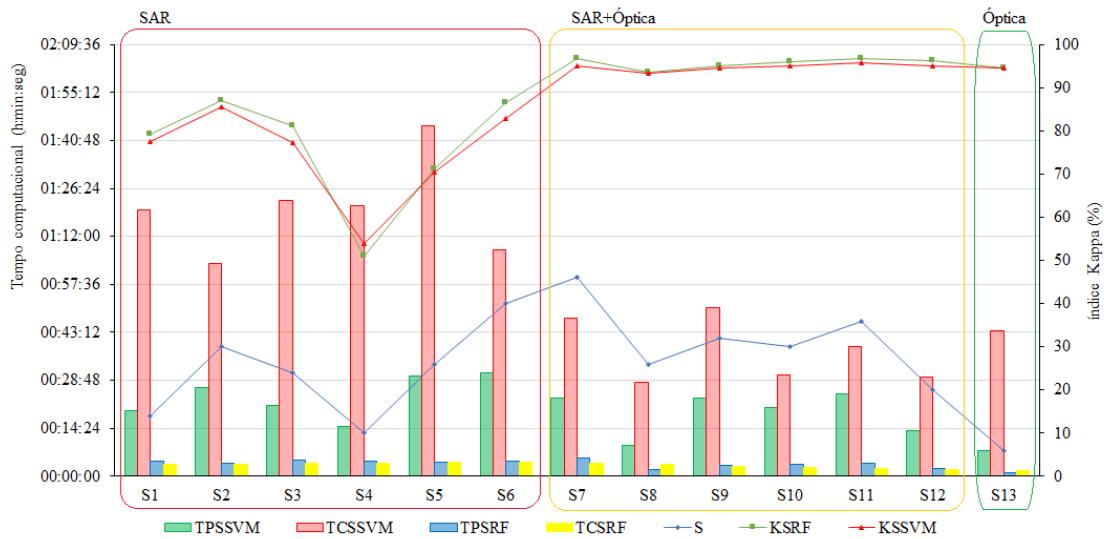
Nota: Tempo de parametrização de *Random Forest* (TPNRF) e *Support Vector Machine* (TPNSVM). Tempo de classificação de RF (TCNRF) e SVM (TCNSVM). Índice Kappa de RF (KSRF) e SVM (KSSVM).

Fonte: O autor (2017).

4.3.2. Tempo Computacional para diferentes cenários

Na Figura 4.5. pode ser notado que para ambos classificadores, o tempo computacional é maior para os cenários envolvendo produtos SAR. Todavia, ao adicionar parâmetros ópticos aos cenários SAR, o tempo computacional tende a diminuir. A correlação de Spearman, de acordo com a Figura 4.6., mostra que os cenários não interferem no tempo computacional. Todavia, apenas para o tempo de parametrização do SVM, o número de variáveis obteve significância positiva, ou seja, conforme o número de bandas aumenta, eleva também o tempo computacional.

Figura 4.5 .Cenários em relação ao tempo computacional e índice Kappa.



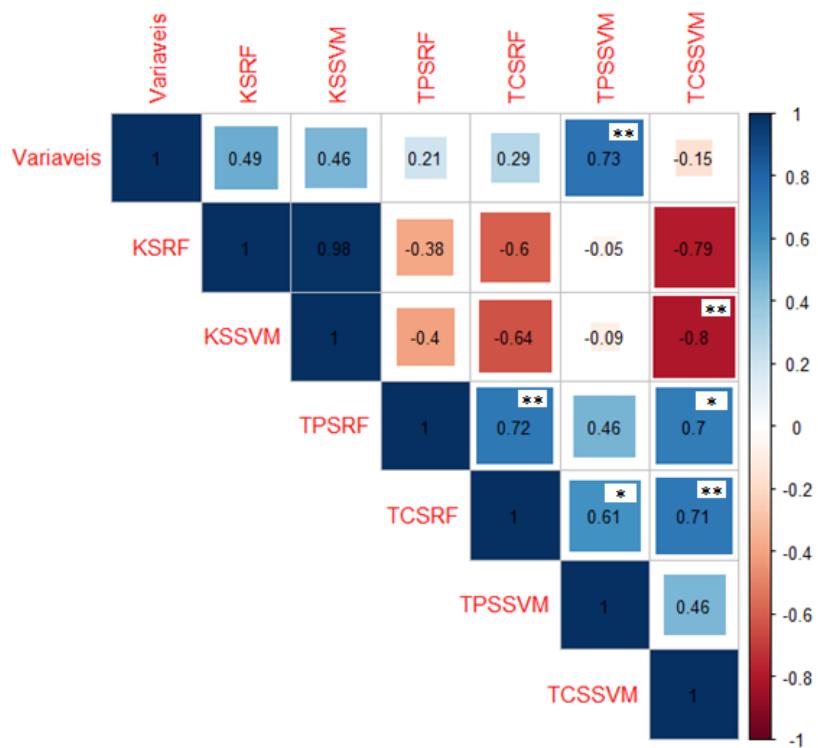
Nota: Tempo de parametrização de *Random Forest* (TSPRF) e *Support Vector Machine* (TPNSVM). Tempo de classificação de RF (TCNRF) e SVM (TCNSVM). Índice Kappa de RF (RF) e SVM (KSSVM).

Fonte: O autor (2017).

Pode ser observado que o índice Kappa do SVM proporcionou correlação negativa com o tempo de parametrização. Em outras palavras, a melhoria do índice Kappa não está correlacionado com o tempo de execução do SVM. O algoritmo RF para o tempo de execução não apresentou correlações significativas com os números de variáveis. Entretanto, foi possível notar que houve uma correlação positiva entre o tempo de parametrização com o de classificação.

Ao serem correlacionados os tempos de execução dos diferentes classificadores, pode ser percebida a obtenção de uma correlação positiva, tanto para a parametrização quanto para a classificação. Este fato proporciona a afirmação que os dois apresentam os mesmos mecanismos de tempo computacional, entretanto o SVM, como apresentado na Figura 4.5, apresentou um tempo de execução muito elevado ao ser comparado com RF.

Figura 4.6. Correlação de Spearman em relação aos cenários de combinações.



Obs.: * Teste significativo a 5%; ** Teste significativo a 1%; ___Teste não significativo.

Fonte: O autor (2017).

4.4 CONCLUSÕES

Diante dos resultados expostos, pode ser findado que o classificador RF exige menor tempo de execução computacional e apresentou melhores índices Kappa, ao ser comparado com SVM. Pode ser concluído ainda, que além da escolha ideal para a melhor classificação, ou seja, melhor índice Kappa, o tamanho de amostra de treinamento e validação interfere no tempo computacional para ambos classificadores (SVM e RF). Por outro lado, o número de variáveis, ou seja, os diferentes cenários, não interferiu no tempo de execução computacional de ambos classificadores.

Assim, pode ser recomendada novas pesquisas que visem a estabelecer classificadores paramétricos e não paramétricos, diferenciando aplicativos de processamento de imagem digital e apresentando resultados que indiquem qual o melhor algoritmo de classificação e os melhores aplicativos para determinadas áreas e sensores.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAF. Anuário Estatístico-Associação Brasileira de Produtores de Florestas Plantadas. **Anuaário Estatístico ABRAF**, p. 146, 2013.
- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; DE MORAES GONÇALVES, J. L.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.
- BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.
- BOERNER, W. M. et al. On the Basic Principles of Radar Polarimetry: the Target Characteristic Polarization State Theory of Kennaugh, Huynen's Polarization Fork Concept, and its Extension to the Partially Polarized Case. **IEEE Transactions on Geoscience and Remote Sensing**, v. 79, n. October, p. 1538–1550, 1991.
- CLOUDE, S. R.; POTTIER, E. An entropy based classification scheme for land applications of polarimetric SAR. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 1, p. 68–78, 1997.
- COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37–46, 1960.
- CONGALTON, R. G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. **Remote Sensing Environment**, v. 46, n. October 1990, p. 35–46, 1991.
- CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: Principles and practices**. Lewis Publ ed. New York: Lewis Publishers, 1999.
- COSTA, T.; SOUZA, M.; BRITES, R. Delimitação e caracterização de áreas de preservação permanente por meio de um sistema de informações geográficas (SIG). **Revista Árvore**, v. 20, n. 1, p. 129-135, 1996.
- FELDE, G. W.; ANDERSON, G. P.; COOLEY, T. W.; et al. **Analysis of Hyperion Data with the FLAASH Atmospheric Correction Algorithm**. In: IEEE International Geoscience and Remote Sensing Symposium. New York, **Anais...** . p.90–92, 2003.
- FIGUEIREDO, D. **Conceitos Básicos de Sensoriamento Remoto**. Brasília - DF, 2005.
- FREEMAN, A. e DURDEN, S. A three-component scattering model for polarimetric SAR data. **IEEE Transactions on Geoscience and Remote Sensing**. v.36, p. 963-973, 1998.
- GALLARDO-CRUZ, J. A.; MEAVE, J. A.; GONZÁLEZ, E. J.; et al. Predicting tropical dry forest successional attributes from space: Is the key hidden in image texture? **PLoS ONE**, v. 7, n. 2, p. 38–45, 2012.
- GOVERNO DO ESTADO DE SANTA CATARINA, O. Sistema de Informações Geográficas de Santa Catarina. Disponível em: <<http://sigsc.sds.sc.gov.br/>>..
- KUNCHEVA, L. I. **Combining pattern classifiers: methods and algorithms**. New Jersey:

John Wiley & Sons, 2004.

LAURIN, G. V.; LIESENBERG, V.; CHEN, Q.; et al. Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa. **International Journal of Applied Earth Observation and Geoinformation**, v. 21, n. 1, p. 7–16, 2013. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.jag.2012.08.002>>..

LEE, J. SEN; GRUNES, M. R.; POTTIER, E. Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 11, p. 2343–2351, 2001.

LIM, H. S.; MATJAFRI, M. Z.; ABDULLAH, K.; SALEH, N. M. High Spatial Resolution Land Cover Mapping Using ALOS Data Over Kedah, Malaysia. First Joint PI Symposium of ALOS DataNodes for ALOS Science Program Kyoto. **Anais...** . p.1–4, 2007.

MEDINA, M.; FERTING, C. Algoritmos e Programação: Teoria e Prática. **Novatec ED**, São Paulo, p. 386, 2006.

MENESES, P. R.; ALMEIDA, T. DE. Introdução ao processamento de imagem de sensoriamento remoto. **Cnpq/Unb**, p. 266, 2012.

MITCHARD, E. T. A. et al. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: Overcoming problems of high biomass and persistent cloud. **Biogeosciences**, v. 9, n. 1, p. 179–191, 2012.

NOVO, E. M. L. DE M. **Sensoriamento Remoto: Princípios e Aplicações**. São Paulo: Edgard Blucher, 2010.

OLIVER,C.J e QUEGAN. Understanding Synthetic Aperture Radar Images. **Artech House**. p.480,1998.

POPE, K. O.; REY-BENAYAS, J. M.; PARIS, J. F. Radar remote sensing of forest and wetland ecosystems in the Central American Tropics. **Remote Sensing of Environment**, v. 48, n. 2, p. 205–219, 1994.

SAATCHI, S. S. et al. **Estimation of Forest Biomass Change from Fusion of Radar and Lidar Measurements.2010**. Disponível em: <<https://pt.slideshare.net/grssiee/estimation-of-forest-biomass>>. Acesso em: 1 set. 2017.

SINCLAIR, G. The Transmission and Reception of Elliptically Polarized Waves. **Proceedings of the IRE**, v. 38, n. 2, p. 148–151, 1950.

ULABY, F.; ELACHI, C. **Radar Polarimetry for Geoscience Applications**. Artech Hou ed. 1990.

VAN DER LINDEN, S.; RABE, A.; HELD, M.; et al. The EnMAP-box-A toolbox and application programming interface for EnMAP data processing. **Remote Sensing**, v. 7, n. 9, p. 11249–11266, 2015.

WIJAYA, A.; LIESENBERG, V.; GLOAGUEN, R. Retrieval of forest attributes in complex

successional forests of Central Indonesia: Modeling and estimation of bitemporal data. **Forest Ecology and Management**, v. 259, n. 12, p. 2315–2326, 2010.

CONSIDERAÇÕES FINAIS

A alta demanda por produtos coletados por sensores remotos é bastante difundida, principalmente quando se trata de sensores remotos ópticos. Entretanto, em algumas regiões da superfície terrestre há muita presença de nuvens, e os sensores ópticos são sensíveis a este efeito. Logo, os sensores micro-ondas surgem como alternativas, pois ultrapassam nuvens (dependendo da densidade da nuvem e do comprimento de onda do sensor) com facilidade. Porém, as pesquisas e trabalhos com sensores micro-ondas, ainda é incipiente.

Portanto, o presente estudo realizou quatro objetivos, sendo o primeiro a realização de uma breve revisão bibliográfica sobre o imageamento de sensores ópticos e de micro-ondas. O segundo, foi avaliar a performance das classificações supervisionadas oriundas dos dados ópticos em diferentes tamanhos amostrais. O terceiro, avaliou a performance de dados óptico e dados SAR, de forma isolada e combinada, para o mapeamento, e o quarto avaliou o tempo computacional de dois classificadores não-paramétricos *Support Vector Machine* (SVM), e o *Random Forest* (RF) para fins de classificação, tanto para diferentes tamanhos de amostras, quanto para diferentes números de variáveis de entrada.

Assim, é plausível concluir que é recomendado a utilização de algoritmos não-paramétricos para sensores orbitais ópticos e que o conjunto amostral 60% de treinamento e 40% para validação, apresentou melhores índices Kappa. Posteriormente do tamanho ideal de amostra definido, foi possível notar que a integração dos dados micro-ondas com os ópticos, demonstraram melhorias significativas na acurácia e que é recomendado a utilização individual do sensor micro-ondas na banda L para a classificação de uso e cobertura da terra. Em relação aos algoritmos estudados, é legítimo concluir que o RF apresentou menores tempos computacionais e maiores índices Kappa, ao ser comparado com o SVM, recomendando a utilização deste algoritmo de classificação.