



Acurácia temática para classificação de imagens utilizando abordagens por pixel e por objetos

Thematic Accuracy for image classification using per pixel and per objects approaches

*André Fabiano Meller Mastella*¹
*Carlos Antonio Oliveira Vieira*²

Recebido em abril de 2018.
Aprovado em novembro de 2018.

RESUMO

O objetivo do artigo consiste em realizar a validação de uma classificação supervisionada, em uma imagem de satélite Landsat 8, sensor OLI, do município de Nova Veneza-SC, usando 03 (três) diferentes classificadores: Mínima Distância Euclidiana (MDE), Máxima Verossimilhança (MAXVER) e Classificação SEGCLASS. Os dois primeiros, baseados numa classificação por pixel, e o último, numa abordagem por objeto/região, como unidade de classificação. A partir das classificações, foram geradas matrizes de erro (ME), usadas posteriormente para derivar índices, e validar e comparar os níveis de acurácia entre os métodos de classificação. A avaliação da acurácia foi realizada através dos coeficientes Kappa, Tau e teste Z estatístico, e de suas variâncias, a partir da matriz de significância. Os resultados dos experimentos indicaram que o desempenho da classificação supervisionada por objeto foi superior às classificações supervisionadas seguindo abordagens por pixel. Através do teste Z estatístico, verificou-se que os valores obtidos para as classificações individuais são estatisticamente melhores do que uma classificação aleatória, e entre elas são estatisticamente diferentes. Comparando-se os índices Kappa e Exatidão Global, o melhor desempenho foi obtido pelo classificador SEGCLASS, seguido do MAXVER e, por último, o MDE, apesar dos classificadores apresentarem valores muito próximos. Por fim, o melhor mapa temático foi gerado pelo classificador supervisionado por objeto/região, SEGCLASS.

PALAVRAS-CHAVE: Classificação de imagens. Classificação por pixel e por objeto. Acurácia temática.

¹Programa de Pós-graduação em Engenharia de Transportes e Gestão Territorial, Universidade Federal de Santa Catarina, Brasil. E-mail: a.f.m.mastella@posgrad.ufsc.br

²Departamento de Geociências. Programa de Pós-graduação em Engenharia de Transportes e Gestão Territorial. Universidade Federal de Santa Catarina, Brasil. E-mail: carlos.vieira@ufsc.br

ABSTRACT

The objective of this article is to perform an accuracy assessment of a supervised classification on a Landsat 8 satellite image, OLI sensor, from the municipality of Nova Veneza-SC, using 03 (three) different classifiers: Minimum Distance Euclidean (MDE), Maximum Likelihood (MAXVER) and SEGCLASS Classification. The first two, by a pixel-based classification, and the last one, on an object-based classification. Based on the classification, error matrices (ME) were produced separately for each map and used to derive indexes, validate and compare the accuracy levels among the classification methods. Accuracy assessment was performed using the Kappa, Tau and statistical Z test, and their variances, from the significance matrix. The results of the coefficients indicated that the performance of the classification by object was better than the classifications by pixel. Through the statistical Z test, it was found that the values obtained for the individual classifications are statistically better than a random classification, and between them are statistically different. Comparing the Kappa and Global Accuracy indexes, the best performance was obtained by the SEGCLASS classifier, followed by the MAXVER and, finally, by the MDE, although the classifiers presented very close values. Finally, the best thematic map was generated by the SEGCLASS object/region supervised classifier.

KEYWORDS: Image classification. Pixel and object based classification. Thematic accuracy.

* * *

Introdução

Através de técnicas e produtos do sensoriamento orbital, a partir de classificadores, é possível classificar imagens de satélite, a fim de estudar a cobertura da Terra, além de gerar mapas temáticos.

No entanto, todo processo de classificação de imagens e geração de mapas temáticos deve haver uma posterior validação. O uso de matrizes de erro (ME) e índices estatísticos derivados, neste caso, permitem estimar/validar a acurácia destes produtos.

Existe uma tendência de mudança de estratégia, após o surgimento de imagens de alta resolução, seguindo uma abordagem de classificação por pixel, para uma abordagem de classificação por objeto/região, embora não haja muitas comparações entre a acurácia destes produtos (DARWISH, LEUKERT e REINHARDT, 2003; WHITESIDE e AHMAD, 2005).

O objetivo deste artigo é avaliar a acurácia de mapas temáticos da cobertura da Terra obtidos a partir de classificadores supervisionados da Mínima Distância Euclidiana (MDE), Máxima Verossimilhança (MAXVER), ambos seguindo abordagens por pixel, e do classificador por objeto/região SEGCLASS, a partir dos índices derivados de ME e teste Z estatístico, sumarizados numa matriz de significância.

1 Classificação de Imagens

A classificação de imagens multiespectrais é responsável pela criação de mapas temáticos, onde a imagem é subdividida em rótulos/classes (SCHOWENGERDT, 2007).

O primeiro passo à classificação de imagens é o reconhecimento de categorias dos objetos do mundo real (corpos d'água, pastagem, etc.). O segundo passo, baseia-se na rotulagem numérica das entidades (pixel/objeto) a serem classificadas (ADAMI, PINHEIRO e MOREIRA, 2002). No entanto, deve-se determinar, *a priori*, o número e natureza das categorias, onde a cobertura da Terra é descrita (floresta, pastagem, dentre outras).

As duas abordagens para treinar e rotular o pixel são conhecidas na literatura como classificação supervisionada e não-supervisionada (MATHER, 2004, p. 203).

Na classificação supervisionada, o analista define suas amostras de treinamento previamente. Assim, o valor espectral de cada pixel no ponto de treinamento é usado para definir o limite de decisão para tal classe. O algoritmo, então, classifica todos os pixels, ou objetos, de acordo com alguma regra de decisão. As amostras da área devem ser representativas às respectivas classes (SCHOWENGERDT, 2007).

Já a classificação não-supervisionada é conhecida pelo seu caráter exploratório (GABRIEL, 2013). O computador separa o pixel, ou objeto/região, dentro de uma classe, utilizando algoritmos computacionais, permitindo uma maior uniformidade da composição espectral, não necessitando de

conhecimento prévio das classes (MASTELLA, 2017; SABINS, 1987; SCHOWENGERDT, 2007, p. 399; VIEIRA, 2000).

Congalton e Green (1999) afirmam que a implementação de um robusto sistema de classificação é crucial para o sucesso de mapeamento e acurácia temática, e que a acurácia de um mapa depende de muitos fatores, incluindo nível de detalhes (esquemas de classificação) e a variabilidade das categorias a ser mapeada.

1.1 Classificadores ou regras de decisão

Classificadores são utilizados para reconhecimento de classes/padrões, utilizados para rotular a que classe determinado pixel, ou objeto, pertence. O que difere um classificador do outro é a forma como são criadas as fronteiras de decisão a partir das amostras de treinamento (SANTOS et al., 2003).

O classificador supervisionado da regra da Mínima Distância Euclidiana (MDE), utilizado nesta pesquisa, envolve o cálculo da distância euclidiana entre um vetor de atributos a ser classificado, e o centroide das classes previamente definidas por amostras.

Assim, o classificador examina as distâncias entre um pixel e as médias das classes (centroides), atribuindo o pixel à classe que apresentar a menor distância euclidiana (FREITAS e PANCHER, 2011; VIEIRA, 2000).

Já o classificador por pixel da Máxima Verossimilhança (MAXVER), utilizado nesta pesquisa, é baseado na hipótese de funções de probabilidade de densidade. Esta função é paramétrica, pois usa parâmetros estatísticos derivados dos dados de treinamento (MATHER, 2004, p. 214).

Esse classificador estima a variância e covariância dos padrões espectrais das amostras de treinamento, pressupondo que os pixels que formam o conjunto de treinamento apresentam uma distribuição normal. Desta maneira, a média entre os centroides da distribuição das classes controlará a localização da distribuição, e a variância, sua extensão (LIU, 2006, p. 766; MENESES e ALMEIDA, 2012, p. 202).

Por fim, o classificador supervisionado por objeto SEGCLASS, também utilizado nesta pesquisa, classifica imagens segmentadas a partir do agrupamento de pixels adjacentes, que facilitam a classificação (GABRIEL, 2013, p. 19; MENESES e ALMEIDA, 2012).

O algoritmo SEGCLASS intersecciona os objetos, a partir da imagem segmentada criada, atribuindo os objetos à uma classe, a partir de sua assinatura espectral e informação espacial em relação aos seus vizinhos, como seus aspectos geométricos e temáticos (localização e forma) (ANTUNES, 2003; SOARES, 2013; VIEIRA, 2017).

Para maiores detalhes a respeito dos classificadores, ver Jensen (2015), Liu (2006), Mather (2004), Meneses e Almeida (2012), Schowengerdt (2007) e Vieira (2000).

1.2 Segmentação de imagens

Na segmentação de imagem, grupos de pixels adjacentes são segmentados de acordo com suas características, incluindo o tamanho, a homogeneidade espectral, a homogeneidade espacial e a forma do objeto (LUZ, ANTUNES e JUNIOR, 2009; MENESES e ALMEIDA, 2012; WHITESIDE e AHMAD, 2005).

Os parâmetros de manipulação para segmentar uma imagem podem ser o número de bandas e o tamanho da janela de varredura, que desloca-se por toda a imagem avaliando a similaridade espectral e verificando a variância existente.

Também são utilizados a proporção entre a média e desvio padrão, e a similaridade permitida/tolerada entre segmentos (quanto maior o valor desta similaridade, menor é o número de segmentos) (EASTMAN, 2009; GABRIEL, 2013, p. 20; JENSEN, 2015, p. 416-17). Desta forma, a imagem é particionada e influenciada por parâmetros de escala, tonalidade e forma.

O parâmetro de escala influencia a heterogeneidade dos pixels, definindo o tamanho dos objetos. A tonalidade equilibra a homogeneidade das

cores dos segmentos com a homogeneidade da figura. E o parâmetro forma é o equilíbrio entre a suavidade da borda de um segmento e sua compacidade (LUZ, ANTUNES e JUNIOR, 2009; WHITESIDE e AHMAD, 2005).

Para operação de segmentação, é necessário especificar o número de bandas da imagem [de satélite]. Por conseguinte, são dadas às bandas pesos iguais por padrão. A janela de varredura indica a variância da imagem. Eastman (2009) sugere que uma janela de 3 x 3 garante bons resultados.

Já os pesos para o fator médio da variância são usados para avaliar a similaridade entre os segmentos vizinhos. Por fim, a similaridade de tolerância é usada para controlar o nível de generalização. Os valores de similaridade consistem em uma sequência de números que começam em 0 (zero), indicando o início da partição dos limites das regiões.

Whiteside e Ahmad (2005) alcançaram bons resultados utilizando uma inspeção visual para determinar a visão geral dos valores dos pesos, com parâmetro de escala “10” (dez), fator de contorno/forma “0.4”, densidade “0.7”, e suavidade, “0.3”. Após experimentações, Jensen (2015, p. 420) admite que uma escala de segmentação de 20 (vinte) fora melhor para criação de objetos/regiões, e que o fator “cor” foi mais importante que os contornos, no processo de segmentação.

Neves (2013) testou vários níveis de tolerância de similaridade de 1 (um) a 10 (dez), trabalhando com o nível 4 (quatro), embora o valor tenha sido escolhido aleatoriamente. É importante ressaltar que a segmentação é essencialmente empírica (tentativa e erro) e, em geral, ajustável aos diferentes tipos de imagem, com limiares definidos face à complexidade dos alvos investigados (ANTUNES, 2003, p. 27; JENSEN, 2015, p. 420).

1.3 Esquema de amostragem e número de amostras

Segundo Congalton e Green (2009, p. 79-80), os dois esquemas de amostragem mais comumente usados são a amostragem aleatória simples e a amostragem aleatória estratificada.

Numa amostragem aleatória simples, cada unidade de amostra na área de estudo tem uma chance igual e independente de ser selecionada. A maioria dos índices de validação baseiam-se neste esquema de amostragem.

Já a amostragem aleatória estratificada, é semelhante à amostragem aleatória simples, no entanto, é necessário algum conhecimento prévio sobre a área de estudo para dividir a área em grupos/estratos, sendo amostrados, em seguida. Sua principal vantagem é que todos os estratos (classes), por menores que sejam, serão incluídos na amostra (fator importante para coleta de amostras representativas de cada classe).

A seleção de um esquema de amostragem adequado é fundamental para gerar uma matriz de erro que seja representativa para todo o mapa (CONGALTON e GREEN, 2009, p. 81).

Congalton e Green (2009, p. 75) assumem um mínimo de 50 amostras para cada classe de mapa, para mapas de menos de 1 milhão de acres em tamanho, e menos de 12 classes.

Segundo Mather (2004, p. 214), para casos multivariados (mais de uma banda espectral), o tamanho da amostra deve ser de pelo menos 30 vezes o número de variáveis discriminantes por classe. Para esta pesquisa, empregou-se a metodologia de Mather (2004), utilizando-se 180 amostras, para 06 (seis) bandas espectrais ($30 \times 6 = 180$ amostras/classe).

1.4 Validação da classificação e coeficientes estatísticos

Stehman e Czaplewski (1998) sugerem que para uma boa avaliação da acurácia temática dos produtos do sensoriamento orbital, são necessários três passos: 1) modelo ou esquema de amostragem; 2) modelo de resposta (uso da melhor regra de decisão), e 3) estimativa e análise da acurácia temática (índices derivados da ME).

A acurácia temática é uma verificação para constatar se o mapa está correto no tocante ao rótulo das classes, além de aumentar a qualidade da informação do mapa, identificando e corrigindo as fontes de erros. Assim, a

avaliação da acurácia determina a qualidade do mapa criado a partir de imagens e produtos do sensoriamento remoto, qualitativa ou quantitativamente, indicando e mensurando os erros do mapa (CONGALTON e GREEN, 2009, p. 2). Para avaliação da acurácia quantitativa, comumente são utilizadas matrizes de erro (ME) (CONGALTON e GREEN, 1999, Cap. 5, p. 3)

Uma ME, descrita na Figura 1, é um arranjo quadrado disposto em linhas e colunas, e expressa o número de unidades amostrais, inicialmente atribuídas, confrontando os dados de referência (coluna) com os dados classificados (linhas), representando a acurácia do mapa (CONGALTON e GREEN, 1999; CONGALTON e GREEN, 2009).

Figura 1 – Exemplo de ME mostrando os “Dados de Referência” (colunas), “Dados Classificados” (linhas), diagonal principal com as amostras classificadas corretamente (negrito), exatidão global dos dados, acurácia dos dados do produtor e dados do usuário

		Dados de Referência			
		A	B	C	Total (linhas)
Dados Classificados	A	10	20	30	60
	B	40	50	60	150
	C	70	80	90	240
Total (colunas)		120	150	180	450
Acurácia “Dados do Produtor”		Acurácia “Dados do Usuário”		“Exatidão Global”	
A	$10/120*100=8,3\%$	A	$10/60*100= 16\%$	$10+50+90/450$	
B	$50/150*100=3,3\%$	B	$50/150*100=3,3\%$	$150/450*100=33\%$	
C	$90/180*100=5\%$	C	$90/240*100=3,8\%$		

Fonte: elaborada pelos autores.

Na ME (Figura 1), as colunas representam os dados de referência, e as linhas, a resposta do classificador. A diagonal principal da ME representa os valores de concordância entre os dados de referência e dados classificados, classificados corretamente.

Para gerar a exatidão global, basta somar os valores da diagonal principal e dividir pelo valor total das amostras $((10+50+90) / 450)*100=33\%$.

Para produzir a acurácia dos dados do produtor, basta dividir o valor da diagonal de cada classe, com o total da classe, disposta no total da coluna ($A = (10/120) * 100 = 8,3\%$). Para produzir a acurácia dos dados do usuário, usa-se o valor da diagonal de cada classe, dividido pelo total da classe disposta no total de linhas ($A = (10/60) * 100 = 16\%$).

Congalton e Green (1999), afirmam que é importante, para outros casos, outras medidas de acurácia. Não obstante, através da ME, é possível derivar índices estatísticos como Kappa, Tau, teste Z estatístico, dentre outros.

O coeficiente Kappa é uma medida de quão bem a classificação concorda com os dados de referência. É expresso através de medidas de concordância, indicada pelos elementos da diagonal da ME, subtraídos da discordância, indicada pelos produtos das somas marginais das linhas e colunas da ME (CONGALTON e GREEN, 1999; VIEIRA, 2000, p. 82).

Já o índice Tau fornece uma medida quantitativa relativamente precisa e intuitiva sobre a acurácia da classificação, sendo superficialmente similar ao Kappa (FIGUEIREDO e VIEIRA, 2007).

Quanto ao teste Z estatístico, é usado para analisar o quanto uma medida se afasta da média, em termos de desvio padrão, analisando 01 (uma) ou 02 (duas) ME normalmente distribuídas, identificando, individualmente, um índice Kappa, ou para duas classificações, onde para um teste individual, Z pode ser testado, assumindo a hipótese nula ($Z=0$), que é rejeitada se $Z \neq 0$ (CONGALTON E GREEN, 1999).

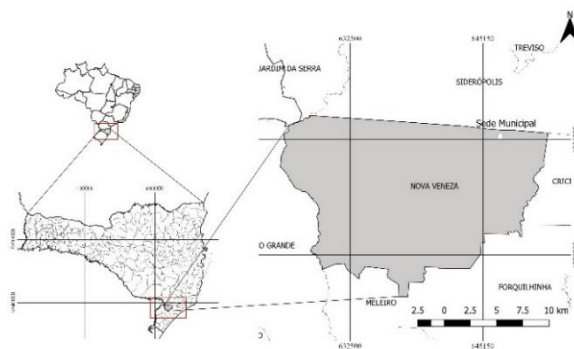
O teste Z estatístico pode utilizar a matriz de significância (MS), que serve como abordagem alternativa para comparação de ME independentes, com os resultados sumarizados na forma de matriz.

Para consultar as equações dos coeficientes Kappa, Tau, teste Z estatístico e matriz de significância, ver Vieira (2000), Figueiredo e Vieira (2007), Congalton e Green (2009), Mastella (2017), Mastella e Vieira (2017).

2 Área de Estudo

O município de Nova Veneza, apresentado na Figura 2, fica localizado no sul do estado de Santa Catarina. Limita-se ao norte com os municípios de Siderópolis, ao sul com Meleiro, a leste com Criciúma e Forquilha, a sudoeste e oeste com Morro Grande, e a noroeste com Bom Jardim da Serra e São José dos Ausentes (RS).

Figura 2 – Mapa de localização de Nova Veneza



Fonte: Elaborada pelos autores.

Nova Veneza conta com uma população de 13.309 habitantes, tendo a sua cobertura da Terra voltada à rizicultura (7.800 ha), mineração de argila e seixos (29 ha), banana (15 ha), pastagens (228 ha), pecuária (87 ha), milho (750 ha), fumo (90 ha), feijão (195 ha), aviários e indústria (463 ha). Tem ocorrido conflitos no tocante ao “abastecimento de água” x “agricultura”, entre produtores. Ainda, há evidências da poluição fluvial do rio Mão Luzia, um dos principais do município (MASTELLA e VIEIRA, 2017).

Possui um clima mesotérmico, tipo úmido (Cfb), com chuvas para o trimestre outubro, novembro e dezembro de 2016, para este estudo de caso, entre 500-600mm (100 a 200mm por mês), e temperaturas entre 20-22°C, e 18-20°C, na parte noroeste do município (MASTELLA, 2017).

3 Material e métodos

3.1 Material

Para esta pesquisa foram utilizados 01 (uma) imagem do satélite Landsat 8, órbita ponto 220/080, sensor OLI, do dia 03 de novembro de 2016, com uma resolução espacial de 30m, 4.70% de nuvens, qualidade 09, disponibilizada pelo Instituto Nacional de Pesquisas Espaciais (INPE) <http://www.dgi.inpe.br/siteDgi/index_pt.php>, em 15 de novembro de 2016.

O sensor OLI está acoplado ao satélite Landsat 8, e provê dados multiespectrais a cada 16 dias, orbitando aproximadamente a 705 km de distância da Terra, órbita síncrona ao sol, completando-a em 98.9 minutos, adquirindo 650 a 725 cenas diárias, e produtos em GeoTIFF.

Possui 11 bandas espectrais, das quais foram utilizados, para esta pesquisa, as bandas 2 (azul), 3 (verde), 4 (vermelho), 5 (infravermelho próximo), 6 (infravermelho médio) e 7 (infravermelho médio) (USGS, 2016; USGS, 2016a, p. 05, 07-9).

3.2 Métodos

A imagem de satélite adquirida foi reamostrada, de 16 bits para 8 bits, para reduzir a variância interna, e aumentar a distância das classes, no ambiente SIG livre QGIS *Wien*. Ela ainda foi recortada à área de estudo, através do sistema *OSGeo4w Shell* (GDAL), e georreferenciada a um sistema cartográfico projetivo conhecido (SIRGAS2000/UTM fuso 22S)

No software IDRISI 16.03, *The Taiga Edition*, da Clarck University, a imagem de satélite foi vetorizada com base nas amostras para criação do mapa de referência, além do uso dos classificadores por pixel MDE e MAXVER, e do classificador por objeto/região, SEGCLASS, para criação dos mapas temáticos e formulação das matrizes de erro (ME).

A partir do software *Google Earth Pro* (2016), em seguida à aquisição da imagem de satélite, foi feita a coleta amostral para definição das classes e confecção do mapa de referência, para posterior seleção de amostras para treinamento e validação da classificação.

A coleta amostral, a partir do software *Google Earth Pro*, foi definida por 04 (quatro) classes informacionais: floresta, culturas agrícolas, solos expostos e pastagem. Não foi possível fazer uma análise multitemporal, a fim de melhorar a interpretação dos alvos, melhorando a coleta amostral.

As amostras foram coletadas aleatoriamente, via *Google Earth Pro* (*Image* © 2016 Digital Globe e *Image* © 2016 CNES / Airbus), conforme a seção 1.3.

A categoria bordas/classes abandonadas são as áreas urbanas e hidrografia, e representam valor 0 (zero), sendo mascarada para não atrapalhar o processo de classificação. A ME do classificador por objeto SEGCLASS computou 3.991 amostras da classe abandonada, e sua linha, na ME, foi excluída.

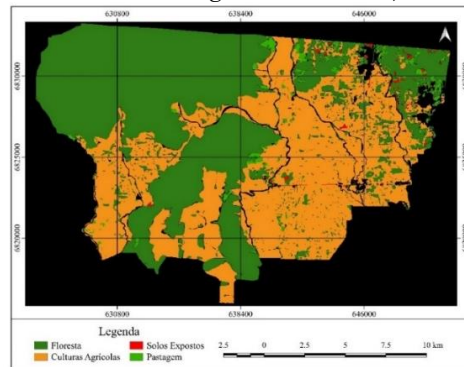
A classe floresta constituiu-se de toda vegetação ombrófila densa, sub-montana, floresta de Eucalipto (*Eucalyptus*) e talhões em estágio fenológico médio e avançado.

A classe culturas agrícolas baseou-se em culturas de terraços (arroz) e culturas de sequeiro (milho, feijão, cana-de-açúcar, fumo, banana, dentre outras) e áreas agrícolas em processo de crescimento. A classe solo exposto abarcou solos que sofreram ação antrópica, como aterros, bem como topos de morros. A classe pastagem refere-se aos campos entre gramínea à capoeirinha, abarcando pequenas propriedades, cuja área florestal foi desmatada, dando lugar a pequenos campos naturais. Ainda no que tange à capoeirinha, baseia-se, também, nos estádios iniciais de regeneração que ocorrem na floresta ombrófila densa, montana e sub-montana, e nas matas secundárias, além de poteiros e lavouras abandonadas (campos sujos).

O tamanho mínimo da seleção de amostras por classe foi de 180/classe (MATHER, 2004, p. 214), e foram utilizadas as bandas 2 (dois) à 7 (sete) da imagem de satélite.

No software IDRISI, após o processo de vetorização, os polígonos das amostras de treinamento foram transformados e salvos em formato matriz (*raster*), no software IDRISI. A partir deste arquivo, mais a imagem Landsat 8, bandas espectrais 2 a 7, e o mínimo do tamanho de amostras por classe (180), criou-se mapa de referência (Figura 3).

Figura 3 – Mapa de referência confeccionado a partir das amostras definidas via Google Earth, coletadas da imagem Landsat 8, bandas 2 a 7.



Fonte: Elaborada pelos autores.

Para classificação, todos os classificadores utilizaram as mesmas bandas de satélite (2 à 7), além do mapa de referência (Figura 3). Para a classificação do MAXVER, foi assumida uma definição de igual probabilidade *a priori*.

Já para a classificação SEGCLASS, foram utilizados uma janela de varredura 3 x 3 (centralidade do pixel), fator de peso médio “0.5”, similaridade de tolerância “10” (dez), e fator de variação de peso “0.5”. Através do mapa temático gerado pelo algoritmo MAXVER, criou-se o mapa temático pelo método SEGCLASS (Figura 6).

Após sobreposição dos mapas temáticos com o mapa de referência (Figura 3), no software IDRISI (*Assesment Accuracy*), foi gerado a ME de cada classificador. Para confecção das ME, todos os mapas gerados pelos classificadores utilizaram o mapa de referência, mais seus respectivos mapas.

Nas ME, foi possível observar a acurácia do produtor e do usuário, a exatidão global, e com o uso do software Kappa Analysis, desenvolvido em C padrão, disponibilizado por Vieira (2000), foram realizadas as análises da acurácia, a partir de índices estatísticos e suas respectivas variâncias, refinando, assim, a acurácia temática. Os resultados obtidos pelo Kappa Analysis foram sumarizados numa matriz de significância.

4 Apresentação e análise dos resultados

4.1 Apresentação dos resultados

A Tabela 1 apresenta a ME do classificador MDE. A diagonal principal apresenta as amostras corretas das classes floresta, culturas agrícolas, solos expostos e pastagem.

Tabela 1 – ME do mapa de referência x dados do mapa temático gerado pelo classificador Mínima Distância Euclidiana.

	F	C	S	P	Total	DR*
F	142.340	7.312	2.670	10.921	163.243	0.1280
C	17.321	72.214	22.857	28.870	141.262	0.4888
S	36	289	435	279	1.339	0.4511
P	504	598	586	4.018	5.706	0.2958
Total	160.201	80.413	26.848	44.088	311.550	
DC*	0.1115	0.1020	0.9726	0.9089		0.2961

*DR = Dados de Referência; *DC = Dados Classificados; F= floresta; C= culturas a.; S= solos e; P= pastagem;

Fonte: elaborada pelos autores

A Tabela 2 apresenta a ME do classificador MAXVER. A diagonal principal apresenta as amostras classificadas corretamente das classes floresta, culturas agrícolas, solos expostos e pastagem.

Tabela 2 – ME do mapa de referência x dados do mapa temático gerado pelo classificador Máxima Verossimilhança.

	F	C	S	P	Total	DR*
F	145.934	2905	1.798	12.606	163.243	0.1060
C	10.266	78.131	14.986	37.879	141.262	0.4469
S	69	145	667	458	1.339	0.5019
P	621	444	325	4.316	5.706	0.2436
Total	156.890	81.625	17.776	55.259	311.550	
DC*	0.0698	0.0428	0.9625	0.9219		0.2648

***DR** = Dados de Referência; ***DC** = Dados Classificados; **F**= floresta; **C**= culturas a. **S**= solos e; **P**= pastagem;

Fonte: elaborada pelos autores

A Tabela 3 apresenta a ME do classificador SEGCLASS. Em sua diagonal principal têm-se as amostras classificadas corretamente das classes floresta, culturas agrícolas, solos expostos e pastagem.

Tabela 3 – ME do mapa de referência x dados do mapa temático gerado pelo classificador por objeto SEGCLASS.

	F	C	S	P	Total	DR*
F	147.540	3.262	1.095	9.670	163.243	0.0962
C	10.700	77.347	12.895	38.409	141.262	0.4525
S	155	67	435	516	1.339	0.6751
P	1.011	334	93	4.030	5.706	0.2937
Total	159.406	81.010	14.518	52.625	311.550	
DC*	0.0744	0.0452	0.9700	0.9234		0.2638

***DR** = Dados de Referência; ***DC** = Dados Classificados; **F**= floresta; **C**= culturas a.; **S**= solos e; **P**= pastagem;

Fonte: elaborada pelos autores

Os três classificadores trabalharam com intervalos de confiança de 90% (+/-0.0013) 95% (+/-0.0015) e 99% (+/- 0.0021). Em relação à matriz de significância (Tabela 4), é possível analisar os 03 (três) classificadores (MDE, MAXVER e SEGCLASS) e suas comparações.

Tabela 4 – Comparação entre os classificadores mostrando os coeficientes Kappa pelo teste Z estatístico (individual na diagonal) e entre dois Kappa (fora da diagonal).

CLASSIFICADORES	MDE	MAXVER	SEGCLASS
EXATIDÃO GLOBAL	70.4	73.5	74.6
KAPPA	0.5180	0.5686	0.5798
TAU	0.605230	0.646918	0.660956
VAR	0.000001	0.000001	0.000001
MDE	518.087		
MAXVER	35,723	568.607	
SEGCLASS	43,647	7,924	579.813

Fonte: elaborada pelos autores

O classificador por pixel MDE obteve um Kappa de 0,518087, e índice Tau de 0,605230. O classificador por pixel MAXVER registrou um Kappa de 0,568607, e índice Tau de 0,646918. O classificador por objeto/região SEGCLASS obteve um Kappa de 0,579813, e índice Tau de 0,660956. Todos os classificadores possuíam a mesma variância (0,0000010).

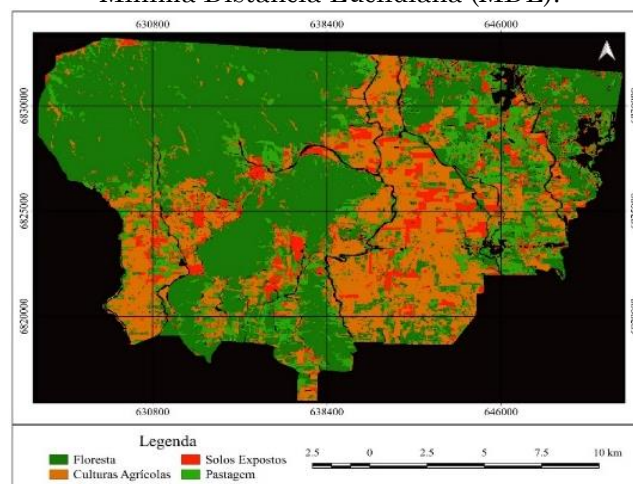
Por fim, com um nível de significância de 0,05, as classificações são melhores que uma classificação aleatória, e com 95% das amostras de treinamento, a longo prazo, a acurácia verdadeira dos classificador por pixel MDE será de 70.2% ou maior, contra 73.4% do classificador por pixel MAXVER, e 74.4%, do classificador por objeto/região SEGCLASS (VIEIRA, 2000).

Em relação à matriz de significância, os valores na diagonal representam o teste Z estatístico para classificações individuais, a partir das formulações apresentadas por Congalton e Green (1999), resultando em valores de Z estimados acima do Z tabelado (1,96), sendo todas as classificações significativamente melhores do que uma classificação aleatória.

Para o teste Z estatístico para 02 (dois) coeficientes Kappa fora da diagonal principal, conforme a Tabela 4, todos os valores estimados de Z acima do valor tabelado (1.96) tiveram valores estatisticamente diferentes, significando que a hipótese nula ($Z=0$) foi rejeitada ($Z_{cal} > Z_{tab}$), indicando que todos os classificadores tiveram performances estatisticamente diferentes.

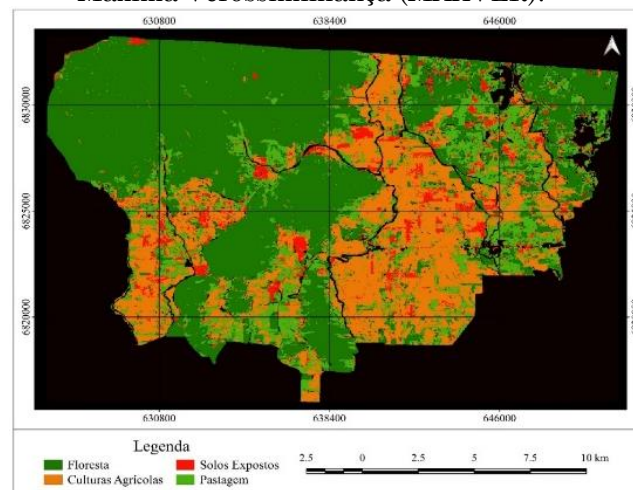
A seguir, são apresentados 03 (três) mapas temáticos (Figuras 4, 5 e 6), frutos de uma classificação supervisionada, usando 02 (dois) classificadores supervisionados por pixel (MDE e MAXVER), e 01 (um) por classificação supervisionada segmentada por objeto/região (SEGCLASS).

Figura 4 – Mapa temático gerado pelo classificador Mínima Distância Euclidiana (MDE).



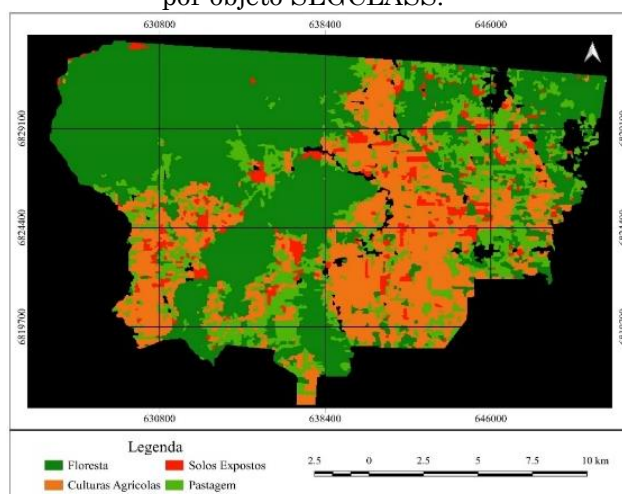
Fonte: Elaborada pelos autores.

Figura 5 – Mapa temático gerado pelo classificador Máxima Verossimilhança (MAXVER).



Fonte: Elaborada pelos autores.

Figura 6 – Mapa temático gerado pelo classificador por objeto SEGCLASS.



Fonte: Elaborada pelos autores.

Na Figura 4 é possível analisar 4 (quatro) categorias classificadas (floresta, culturas agrícolas, solos expostos e pastagem), pelo classificador MDE.

A parte noroeste do município foi classificada majoritariamente como floresta, bem como o centro-sudoeste e nordeste do município. A parte central do município foi classificada como culturas agrícolas e solos expostos. Por último, a região central de Nova Veneza (nordeste), sudeste, sul, e bordas da categoria floresta, localizadas à noroeste, foi classificada como pastagem, além da grande heterogeneidade de todas as classes ao redor do município.

A Figura 5 é fruto de uma classificação supervisionada por pixel, pelo classificador MAXVER, apresentando 4 (quatro) categorias (floresta, culturas agrícolas, solos expostos e pastagem).

A região noroeste, centro-sul e sudoeste, e parte da região nordeste do município, foi classificada como floresta. Já a região centro-sul, sudoeste, e uma pequena parte da região leste, foi classificada como culturas agrícolas. A região centro-leste e oeste, teve a cobertura da Terra classificada como solos expostos. Por último, a região central de Nova Veneza (nordeste), sudeste, sudoeste, teve classificação pastagem. Há distribuição de todas as categorias, em menor número, ao redor do município.

A Figura 6, fruto de uma classificação por objeto/região, teve as bordas das categorias floresta, culturas agrícolas, solos expostos e pastagem bem discernidas. Numa análise visual, baseado no mapa de referência (Figura 3), constata-se que algumas categorias de solos expostos se sobrepuseram à classe culturas agrícolas, mas nenhuma discrepância concernente à classes sobrepondo a categoria floresta, quando confrontada ao mapa de referência (Figura 3). A generalização de uma classificação supervisionada por objeto é visível, com áreas homogêneas grandes.

4.2 Análise dos resultados

A ME que teve melhor resultado em relação à exatidão global (Tabela 4) foi a do classificador por objeto SEGCLASS, obtendo um valor de 74.6%, seguido dos classificadores supervisionados por pixel MAXVER (73.5%), e MDE (70.4%). A longo prazo, 95% das amostras de treinamento que melhor terão acurácia verdadeira serão do classificador SEGCLASS, com 74.4%, seguido dos classificadores MAXVER, 73.4%, e do MDE, com 70.4%.

Quanto ao índice Tau, o melhor resultado ficou com o classificador por objeto SEGCLASS (0.660956), contra 0.646918 (MAXVER) e 0.605230 (MDE), dos classificadores por pixel (Tabela 4).

Quanto ao teste Z estatístico, considerando seu valor tabelado ($Z=1,96$), demonstrado na matriz de significância, todos os classificadores rejeitaram ($Z \neq 0$) a hipótese nula ($Z=0$), mostrando-se individualmente diferentes. O teste Z para dois Kappa teve o resultado mais expressivo com os classificadores MAXVER x SEGCLASS, de similaridade 7.924, contra o teste do MDE x MAXVER, 35.723, e 43.647, do MDE x SEGCLASS (estatisticamente denotando uma leve diferença). Todos possuíam variâncias muito próximas (0.000001), não interferindo significativamente nos resultados.

Utilizando a calculadora de campo do SIG, e baseado no mapa de referência (Figura 3), o classificador MDE (Figura 4) sobrepôs algumas áreas da categoria floresta às classes culturas agrícolas (1.704,51 ha). A classe solo

exposto sobrepôs-se à classe culturas agrícolas (526,60 ha). A classe solo exposto do classificador MDE apresentou menor classificação em relação à classificação do MAXVER (MDE, 435 amostras; MAXVER, 667 amostras), com os limites das classes do mapa temático apresentando-se dissolvidos.

Quanto ao classificador supervisionado por pixel MAXVER (Figura 5), teve uma boa classificação da classe floresta, porém, sobreposta por pastagem (11,34 ha), devido a possíveis respostas espectrais semelhantes, além de sobreposição da classe solos expostos à classe culturas agrícolas (460,83). Em contrapartida, apresentou uma boa classificação no tocante à solos expostos (667 amostras), em relação ao mapa de referência (Figura 3, 1339 amostras) e aos demais classificadores (MDE, 435; SEGCLASS, 435). A classe pastagem adentrou a classe floresta (284,70 ha) e culturas agrícolas (148,90 ha), mas obteve melhores resultados que o classificador MDE (Tabela 1 e Figura 4). Quanto ao limite das classes, apresentaram-se um pouco apagados, complexos.

Por fim, o classificador por objeto SEGCLASS (Figura 6) teve um bom desempenho, com pouca sobreposição da classe solos expostos à classe culturas agrícolas (555,24 ha). Teve, assim como o classificador MDE (Tabela 1), baixa performance para a classe solos expostos (435 amostras), em relação ao classificador MAXVER (Tabela 2, 667 amostras). Porém, houve uma boa classificação da classe pastagem (4.991,90 ha), com o limite da classe apresentando-se mais nítido, entretanto, pouco realista.

Comparando os resultados com estudos similares (ABRÃO et al., 2015), utilizando imagens do sensor OLI, 30 m, e classificação supervisionada pelo classificador por pixel MAXVER, esta pesquisa obteve menor resultado comparando-se o índice Kappa ($0.568,607 < 0.60$), e melhor Exatidão Global ($73.5\% > 64\%$).

Comparando os resultados desta pesquisa ao trabalho de Cohenca e Carvalho (2015), trabalhando com classificadores por pixel e objeto/região MAXVER, MDE e OBIA, observa-se um Kappa menor para o classificador MAXVER ($0.56 < 0.57$), maior em relação ao MDE ($0.518,087 > 0.45$), e

SEGCLASS maior em relação ao OBIA (0.579,813>0.53). Em relação à Exatidão Global, esta pesquisa obteve melhores resultados em relação aos classificadores por pixel MDE (70.4%>0.45), MAXVER (73.5%>0.61) e por objeto/região (SEGCLASS, 74.6%>0.57%, OBIA).

Jensen (2015, p. 568), em estudos de cobertura da Terra, alcançou resultados excelentes no tocante à Exatidão Global (93.86%) e coeficiente Kappa (92,1%), se comparados aos resultados dos três classificadores desta pesquisa (MDE, 70.4%; MAXVER, 73.5%, e; SEGCLASS, 74.6%).

Por fim, comparando-se os resultados desta pesquisa com Tilahun e Teferie (2015), referente à avaliação da acurácia temática de imagens Landsat 8, sensor OLI, esta pesquisa obteve uma Exatidão Global dos classificadores por pixel e por objeto inferior a 82%, bem como Kappa menores de 77.02%, segundo o resultado dos autores.

5 Conclusões

Foi possível alcançar o objetivo geral desta pesquisa, onde através das ferramentas e produtos do sensoriamento remoto orbital, pôde-se avaliar o nível de acurácia dos mapas temáticos do município de Nova Veneza-SC, por classificações supervisionadas a partir de classificadores por pixel e por objeto/região, a partir de coeficientes derivados das ME.

Conclui-se que, no geral, o uso de classificadores supervisionados por pixel (MDE e MAXVER) e por objeto/região (SEGCLASS) apresentaram resultados satisfatórios.

Quanto à avaliação da acurácia, constatou-se que ela é uma etapa de validação da classificação fundamental à obtenção de melhores resultados para os mapas temáticos gerados através dos classificadores, a partir de imagens de satélite.

Quanto aos resultados alcançados na pesquisa, para o mês de novembro de 2016, têm-se resultados satisfatórios, frente à complexidade da cobertura

da Terra, das culturas temporárias em início de plantio, e das similaridades físicas e assinaturas espectrais do terreno.

A classe floresta foi bem classificada pelos 03 (três) classificadores devido Nova Veneza possuir mais de 3.510 ha de matas e/ou florestas, APPs e Silvicultura. Já a classe culturas agrícolas também foi bem classificada devido o município possuir aproximadamente 8.550 ha de terras para plantio, majoritariamente arroz e milho.

Por último, os classificadores tiveram baixo desempenho quanto à classe solos expostos. O problema em relação à classificação pode estar no classificador, classificando terraplanagem, topos de morros, pedreiras, da classe solos expostos, às culturas agrícolas em fase de plantio, devido a possível similaridade espectral e maior cobertura da Terra no tocante à culturas agrícolas.

Em relação aos classificadores, o classificador supervisionado por pixel MDE, e por objeto SEGCLASS, apresentaram o pior resultado no tocante à classificação solos expostos. O classificador MAXVER obteve um bom resultado em face à classificação das amostras corretas na diagonal principal da ME da classe floresta e culturas agrícolas. Já o classificador supervisionado por objeto/região SEGCLASS obteve melhor resultado entre os demais classificadores por pixel, comprovado a partir da exatidão global, coeficiente Kappa, Tau e teste Z estatístico.

Quanto aos mapas temáticos, o melhor mapa gerado foi do classificador SEGCLASS, com o limite das classes mais nítido em relação aos classificadores por pixel MDE e MAXVER, embora pouco realista em relação às classes.

A matriz de significância e as ME apresentaram resultados diferentes para cada classificador, elencando o melhor resultado do classificador supervisionado por objeto/região SEGCLASS, seguido da classificação supervisionada por pixel MAXVER e MDE, embora com valores aproximados.

Os coeficientes Kappa, Tau e teste Z estatístico apresentaram performances estatisticamente diferentes, com melhor resultado para a

classificação supervisionada por objeto/região SEGCLASS, seguido da classificação supervisionada por pixel MAXVER e MDE.

O teste Z estatístico (um e dois Kappa) pôde indicar que todos os valores obtidos para as classificações individuais são maiores que o valor de Z tabelado, de 1.96, rejeitando-se ($Z \neq 0$) a hipótese nula ($Z=0$), e mostrando que todas as classificações são estatisticamente diferentes e melhores que uma classificação aleatória.

Destaca-se, também, a alta resolução radiométrica (níveis de cinza) entre as classes (solos expostos similares à culturas agrícolas, pastagem similares a culturas agrícolas em processo de crescimento), que faz com que a variação das classes aumente, e a variação interclasses diminua, dificultando a discriminação entre as classes, por classificadores tradicionais, prejudicando a classificação.

Recomenda-se fazer a coleta amostral o quanto mais perto à obtenção da imagem de satélite possível, a fim de haver compatibilidade da imagem do satélite com a imagem do mapa de referência, e a utilização de outros classificadores, como o classificador supervisionado Redes Neurais Artificiais (RNA), *Support Vector Machine*, que possibilitam a manipulação eficiente de grandes quantidades de dados, não assumindo nenhuma distribuição normal.

Por fim, recomenda-se o uso da classificação supervisionada por objeto/região, já que ela obteve os melhores resultados, generalizou e definiu melhor os limites e descontinuidades das classes.

Referências

- ABRÃO, C. M. R. et al. Avaliação de classificadores supervisionados e não supervisionados para mapeamento de uso e cobertura da Terra a partir de dados Landsat-8/OLI. **Anais XVII Simpósio de Sensoriamento Remoto – SBSR**, João Pessoa-PB, Brasil, INPE, 2015.
- ADAMI, M.; PINHEIRO, E. S.; MOREIRA, M. A. **Aplicação de diferentes algoritmos para a classificação de imagens ETM+/LANDSAT-7 no**

- mapeamento agrícola.** Ministério da Ciência e Tecnologia. Instituto Nacional de Pesquisas Espaciais – INPE. São José dos Campos, 2002.
- ANTUNES, A. F. B. **Classificação de ambiente ciliar baseada em orientação a objeto em imagens de alta resolução espacial.** Tese de doutoramento. Pós-Graduação em Ciências Geodésicas. Universidade Federal do Paraná - UFPR. Paraná, 2003.
- COHENCA, D.; CARVALHO, R. Comparação de métodos de classificação OBIA, Máxima Verossimilhança e Distância Mínima Euclidiana em imagem OLI/Landsat-8 em área de alta diversidade de uso do solo. **Anais XVII Simpósio de Sensoriamento Remoto – SBSR**, João Pessoa-PB, Brasil, INPE, 2015.
- CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices.** Lewis Publishers, Boca Raton, FL, 1999.
- CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices** / Russell G. Congalton and Kass Green. -- 2nd ed., Boca Raton, FL, 2009. ISBN 978-1-4200-5512-2
- DARWISH, A.; LEUKERT, K.; REINHARDT, W. Urban land-cover classification: an object based perspective. **2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas.** Germany, 2003.
- EASTMAN, J. R. **IDRISI Taiga guide to GIS and image processing.** Clark Labs, Clark University, Worcester, MA, 2009.
- FIGUEIREDO, G. C.; VIEIRA, C. A. O. Estudo do Comportamento dos índices de Exatidão Global, Kappa e Tau, comumente usados para avaliar a classificação de imagens do sensoriamento remoto. **Anais do XIII Simpósio Brasileiro de Sensoriamento Remoto.** Florianópolis, 2007. pp. 5755-5762.
- FREITAS, M. I. C.; PANCHER, A. M., 2011. **Classificação de Imagens.** Universidade Estadual Paulista (UNESP). Rio Claro, 2003. 33p. Site <http://www.rc.unesp.br/igce/planejamento/download/isabel/sist_inf_geografica/Aula%2011/classificacao_imagens.pdf>, acessado em junho de 2016.
- GABRIEL, C. G. F., 2013. **Análise comparada de segmentação e classificação orientada por objectos de uma imagem worldview-2.** Dissertação de Mestrado. Universidade Nova de Lisboa. Faculdade de Ciências Sociais e Humanas. Lisboa, 2013.

- GOOGLEEARTHPRO. Site <<https://www.google.com/earth/>>, acessado em novembro de 2016.
- JENSEN, J. R. **Introductory digital image processing: a remote sensing perspective** / John R. Jensen, university of South Carolina, 2015. ISBN 978-0-13-405816-0
- LIU, W. T. H. **Aplicações de sensoriamento remoto** / William Tse Horng Liu. - Campo Grande : Ed. UNIDERP, 2006. 908 p : il. color. ISBN 85-7704-040-0
- LUZ, N. B.; ANTUNES, A. F. B.; JUNIOR, J. B. T. Segmentação multirresolução e classificação orientada a objetos aplicados a imagens Spot-5 para o mapeamento do uso da terra. **FLORESTA**, v. 40, n. 2, 2009, pp. 429-446.
- MASTELLA, A. F. M. **Avaliação da acurácia temática para classificação de imagens de satélite: estudo de caso no município de Nova Veneza/SC**. Trabalho de Conclusão de Curso de Graduação. Universidade Federal de Santa Catarina – UFSC. Centro de Filosofia e Ciências Humanas – CFH. Florianópolis, 2017. 92p.
- MASTELLA, A. F. M.; VIEIRA, C. A. O. Avaliação da acurácia temática para classificação de imagens utilizando abordagens por pixel e por objetos. **Anais do XXVII Congresso Brasileiro de Cartografia e XXVI Expositiva**, SBC, Rio de Janeiro-RJ, 2017. pp.442-446.
- MATHER, P. M. **Computer processing of remotely-sensed images: An introduction**. 3rd ed. Chichester: John Wiley & Sons, 2004. 324p.
- MENESES, P. R. et al. **Introdução ao processamento de imagens de sensoriamento remoto**. Brasília: UNB/CNPq, 2012.
- NEVES, A. I. P. B. **Classificação da ocupação do solo através da segmentação de uma imagem de satélite de alta resolução**. Projeto de Mestrado. Universidade Nova de Lisboa. Faculdade de Ciências Sociais e Humanas. Lisboa, 2013.
- SABINS, F. F. **Remote Sensing: principles and interpretation**. W. H. Freeman and Company, 2nd ed. New York, 1987. 449p.
- SANTOS, R. L. et al. A teoria da incerteza aplicada a classificação de imagens de satélite. **Congresso Brasileiro de Cartografia**, vol. 21, 2003.
- SCHOWENGERDT, R. A. **Remote Sensing: Models and Methods for Image Processing**. 3ª ed., - Department of Electrical and Computer Engineering

(College of Optical Sciences and Office of Arid Lands Studies), University of Arizona, Tucson. Arizona, 2007.

SOARES, F. A. A. M. **Segmentação de imagens: Processamento Digital de Imagens, Aula 10.** Site <<http://www.inf.ufg.br/~fabrizzio/mestrado/pdi/aulas/aula10.pdf>>, acessado em dezembro de 2016.

STEHMAN, S. V.; CZAPLEWSKI, R. L. Design and analysis for thematic map accuracy assessment fundamental principles. **Remote Sensing of Environment**, 64. New York, 1998. pp. 331-344.

TILAHUN, A.; TEFERIE, B. Accuracy assessment of land use land cover classification using Google Earth. **American Journal of Environment Protection**. Vol. 4, No. 4, 2015, pp. 193-198.

USGS, United States Geological Survey (**Landsat Mission**). Site <<http://landsat.usgs.gov/>>, acessado em abril de 2016.

USGS, United States Geological Survey (**Landsat 8: Data Users Handbook (Version 2.0)**). Site <<https://landsat.usgs.gov/documents/Landsat8DataUsersHandbook.pdf>>, acessado em abril de 2016.

VIEIRA, C. A. O. **Accuracy of Remotely Sensing Classification of Agricultural Crops: A Comparative Study**. Tese de doutoramento. University of Nottingham, United Kingdom, 2000. 321p.

VIEIRA, C. A. O. Visualizando a precisão temática no processo de classificação de produtos do sensoriamento remoto: abordagem por pixel e por objeto. **Anais do XXVII Congresso Brasileiro de Cartografia e XXVI Expositiva**, SBC, Rio de Janeiro - RJ, 2017. pp. 437-441.

WHITESIDE, T.; AHMAD, W. A comparison of object-oriented and pixel-based classification methods for mapping land cover in northern Australia. Proceedings of SSC2005 Spatial Intelligence, Innovation and Praxis: The National Biennial Conference of the **Spatial Sciences Institute**, 2005.