



7, 8 e 9
Março 2018
ÉVORA
Évora Hotel

GESTÃO DOS
RECURSOS HÍDRICOS:
NOVOS
DESAFIOS

AVALIAÇÃO DOS MÉTODOS DE CLASSIFICAÇÃO DE USO E COBERTURA DO SOLO POR IMAGENS ORBITAIS

Uma aplicação na região lagunar de Maceió e demais municípios

Marcio Luciano Ferreira de SÁ FILHO¹; Rosangela Sampaio REIS²

¹ Graduando em Engenharia Ambiental e Sanitária, Universidade Federal de Alagoas, marcio.filho@ctec.ufal.br

² Dra. em Ciências da Engenharia Ambiental, Universidade Federal de Alagoas, rosangelareis_al@hotmail.com

RESUMO

Diferentes métodos de classificações podem ser utilizados para definir por meio de softwares o uso e cobertura do solo através de imagens orbitais. Neste artigo, foram avaliados os métodos de classificação do paralelepípedo (Piped), da mínima distância para médias (Mindist) e da máxima verossimilhança (Maxlike) representantes do grupo supervisionado, e os métodos de agrupamento (Cluster), Kmédias (Kmeans) e agrupamento por vizinho mais próximo (Chaincluster) como representantes do grupo não supervisionado. Sabe-se que para cada método de classificação serão encontrados resultados diferente devido as particularidades do algoritmo de cada um deles. Entretanto, esperasse que tais resultados possuam a mesma ordem de grandeza observada na realidade.

Foi gera de uma “amostra verdade”, definida manualmente por polígonos, com base na identificação visual das características da imagem analisada, onde se conhece os seguintes tipos de uso e cobertura do solo da região: vegetação alta (remanescente de floresta), vegetação baixa (agricultura, pastagem ou grama), água lagunar, água continental (rios), água do oceano, areia, solo exposto e área urbana.

A imagem analisada possui uma resolução de 0,70m x 0,70m para cada pixel, em uma composição multiespectral, R: Banda1, G: Banda2 e B: Banda3. A área total analisada encontra-se entre os municípios de Maceió, Coqueiro Seco e Marechal Deodoro.

Os resultados qualitativos foram obtidos através de comparação entre a amostra verdadeira e o valor resultante em cada um dos métodos de classificação utilizados, por meio do índice de Kappa desta forma, pôde-se obter o método que melhor descreve a região do estudo.

Palavras-chave: Sensoriamento Remoto, Sistema de Informação Geográfica, Komsat-3, Classificação Supervisionada, Classificação não supervisionada.

Tema: qualidade da água e dos ecossistemas.

1. INTRODUÇÃO

As mudanças do uso e cobertura do solo da sua condição original para os diversos usos, agricultura, urbanização, mineração e etc., encontram-se em constante expansão, especialmente em países em desenvolvimento. Por sua vez, estes países podem negligenciar a ocupação de algumas áreas que deveriam ser protegidas. Um exemplo é a cidade de Maceió que se encontra delimitada por duas Áreas de Proteção Ambiental (APA), a APA do Pratagy, a leste do município, e a APA de Santa Rita que se localiza a oeste. Nos últimos anos estas APAs tem sido impactadas com o avanço antrópico sobre elas, principalmente por fatores como o crescimento econômico e a especulação imobiliária, além de já serem afetadas historicamente por áreas de agricultura.

Segundo o censo do Instituto Brasileiro de Geografia e Estatística (IBGE) no ano 2000 Maceió possuía 797.759 habitantes, passando para 932.748 no censo de 2010 e em 2017 a população de Maceió foi estimada em 1.029.129 habitantes. O crescimento populacional muitas vezes não ocorre com o devido planejamento. Neste processo áreas de proteção ambiental podem ser impactadas, de forma direta, com alteração do uso e cobertura do solo da região, ou em decorrência de alterações nas proximidades destas, como por exemplo, a construção de empreendimentos residenciais de grande porte em áreas de proteção ambiental, ou lançamento de efluentes nas águas do Complexo Estuarino Lagunar Mundaú Manguaba (CELMM) pela cidade de Maceió.

O crescimento da área urbana de Maceió foi apontado por Bispo (2011) como um dos fatores de desequilíbrio hídrico na região. O CELMM, por exemplo, tem sido impactado pelo avanço antrópico sobre suas áreas de drenagem e outros efeitos da urbanização. Sendo este estuário alvo de diversos estudos acadêmicos, isso se deve ao fato deste sistema estuarino ser um dos mais importantes do país e sofrer com um processo acelerado de degradação ambiental. Segundo a Agência Nacional de Águas (ANA), o crescimento desordenado da área urbana de Maceió resulta numa situação crítica para cerca de 260 mil habitantes que vivem no ou do CELMM.

Tendo em vista a importante influência do crescimento urbano sobre o CELMM e nas APAs do Rio Pratagy e de Santa Rita, é de suma importância que se avalie os métodos de classificação de uso e cobertura do solo disponíveis para estudos do avanço antrópico na região. Dessa maneira, foram selecionados os métodos supervisionados *Maxlike*, *Mindist* e *Piped*, além dos não supervisionados *Cluster*, *Chaincluster* e *Kmeans*, para que fossem avaliados em relação ao índice de Kappa. O objetivo final foi definir qual dentre estes apresenta melhor resultado para a região em termos da classificação de uso e cobertura do solo.

2. MATERIAL E MÉTODOS

2.1. ÁREA DE ESTUDO

Foi delimitada uma porção ao noroeste da APA de Santa Rita para o estudo, tal região abrange os municípios alagoanos de Maceió, Coqueiro Seco e Marechal Deodoro, possuindo as seguintes coordenadas geográficas: 851392,0278251 - 856112,1277447 e 8927902,2171522 - 8931204,1170960, no fuso 24S do sistema de coordenadas Universal Transversa de Mercator (UTM). A superfície delimitada possui 1.343,232 ha com características de uso e cobertura do solo diversificadas com áreas urbanizadas, de vegetação, espelhos d'água e solos expostos. Área de estudo está delimitada na figura 1.

Área de Estudo, região lagunar Maceió, Coqueiro Seco e Marechal Deodoro.



Figura 1. Delimitação da área de estudo. Fonte: 9°40'17.87"S e 35°46'49.69"W. Google Earth. Maio de 2017. 29 de Janeiro de 2018.

2.2. IMAGEM ORBITAL

A imagem utilizada foi capturada pelo sensor AEISS (Advanced Earth Imaging Sensor System) do satélite Sul Coreano KOMPSAT-3, com altitude orbital de 685,1 Km, inclinação de 98,13°, período de 98,5 minutos e resolução de 0,7 metros, obtida

no período de julho de 2016. As bandas utilizadas para a composição RGB foram, banda 1 (vermelha), banda 2 (verde) e banda 3 (azul), esta composição foi utilizada tanto na obtenção visual uma amostra verdade, quanto para realizar a classificação com os métodos selecionados.

2.3. MÉTODOS DE CLASSIFICAÇÃO

O passo inicial para a classificação foi a escolha do software, sendo este o Idrisi em sua versão Selva. Em seguida realizou-se o corte das bandas da imagem adquirida para região desejada. Após a elaboração composição colorida, R:1, G:2 e B:3, iniciou-se a criação de vetores para as classes e então definição da amostra verdade de uso e cobertura do solo, por diferenciação visual.

Foram necessárias 9 classes para classificar a imagem quanto ao uso e cobertura do solo, sendo três representantes de espelhos d'água (água do oceano, água doce e água lagunar), duas de áreas de vegetação (remanescente e agricultura), e as demais classes solo exposto, areia, urbana e transporte.

Áreas com baixa representatividade foram enquadradas em uma das outras classes como foi o caso de gramados que foram absorvidos pela classe agricultura. Houve, também, a subdivisão de classes, como a água, pois ao utilizar os algoritmos de classificação notou-se uma certa confusão causada pelas águas do oceano, lagunar e continental.

Dos métodos de classificação de imagem:

2.3.1. AMOSTRA VERDADE

A criação da amostra verdade foi fundamental para se realizar a comparação com os algoritmos de classificação, uma vez que para se calcular o índice kappa este seria fundamental. Esta amostra, foi gerada através da criação de vetores com a ferramenta apropriada. Sendo esta realizada de forma visual, por questões temporais, não foi possível o levantamento das áreas por meio de visitas, portanto o levantamento foi realizado conforme o prévio conhecimento das classes presentes na região. Assim, foram demarcadas as áreas de acordo com as categorias.

2.3.2. MÁXIMA VEROSSIMILHANÇA (MAXLIKE)

É um método de classificação supervisionada, necessitando da criação de uma assinatura de classes. A assinatura de classes que foi criada acabou por ser utilizada para os três métodos de classificação supervisionada, *maxlike*,

mindist e *piped*. No *maxlike* a assinatura foi usada com igual probabilidade de ocorrência para cada classe. Com uma proporção de 0% de exclusão dos pixels.

2.3.3. DISTÂNCIA MÍNIMA (MINDIST)

É um outro método supervisionado, que foi processado com a mesma assinatura de classe dos demais. O método é baseado na reflectância média atribuída a cada banda para a dada assinatura, sendo aos pixels, por sua vez, atribuído à classe com a média mais próxima de seu valor. Esta foi configurada para o tipo bruto (raw) de distância e sem limitação da distância de busca do algoritmo.

2.3.4. PARALELEPÍPEDO (PIPED)

O último dos supervisionados avaliado foi o paralelepípedo, este método é baseado na reflectância de limiar superior e inferior. Já se sabe que o algoritmo é o mais rápido entre as rotinas de classificação, entretanto é potencialmente o menos preciso, conforme o manual do Idrisi Selva. Para o processamento da imagem foi necessária a utilização de um z-score igual a 0,6, o que representou a não classificação de cerca de 58% dos pixels, pois para z-scores maiores que 0,6 houve a supressão da maior parte das categorias de uso e cobertura de solo.

2.3.5. AGRUPAMENTO (CLUSTER)

O primeiro dos métodos não supervisionados de classificação foi o *cluster*, este método permite a classificação por avaliação dos picos do histograma, onde a média entre os picos define a divisão de classes. Os níveis de cinza foram configurados para 10, com saturação em 1,0, nível de generalização amplo e regra de agrupamento com redução de classes menos significativas em 1%.

2.3.6. AGRUPAMENTO EM CADEIA (CHAINCLUSTER)

O agrupamento em cadeia consiste em duas passagens, a primeira realiza o agrupamento, enquanto à segunda se aplica uma distância mínima para a classificação média. Para este método se recomenda um ajuste que é obtido

através do processamento das imagens algumas vezes com o intuito de identificar o conjunto ideal de valores. Assim, para o resultado final do método foi selecionado uma distância de raio 190, com um número máximo de clusters de 9 e a frequência de leitura dos pixels das imagens foi da primeira linha para a última e para a direita.

2.3.7. KMÉDIAS (KMEANS)

O método *Kmeans*, último método não supervisionado, usa a distância euclidiana para calcular a distância entre pixels e o centroide do cluster, repetindo o processo até a correção dos centroides. Sabe-se que o desempenho deste método é dependente da inicialização dos centroides. Para os melhores resultados do método foram selecionadas as seguintes configurações, um número máximo de clusters de 9, o centroide inicial alimentado aleatoriamente, uma percentagem de migração dos pixels menores ou iguais a 1%, um número máximo de 100 iterações e a fusão de clusters com proporção menor ou igual a 1%.

2.4. ÍNDICE KAPPA

A análise do índice Kappa consiste em uma técnica discreta e multivariada amplamente usada na avaliação de precisão no meio acadêmico. Esta técnica se utiliza dos elementos de uma matriz de confusão para seu cálculo. O índice Kappa é a medida da concordância real, excluindo-se a concordância por chance. O índice pode ser calculado pela seguinte equação:

$$K = \frac{n \sum_{i=1}^c x_{ii} - \sum_{i=1}^c x_{i+} x_{+i}}{n^2 - \sum_{i=1}^c x_{i+} x_{+i}}$$

Onde,

K é o índice de exatidão Kappa;

r é número de linhas da matriz;

x_{ij} é número de observações na linha i e coluna j;

x_{i+} e x_{+i} são os totais marginais da linha i e coluna j, respectivamente;

n é número total de observações.

O nível de exatidão de uma classificação avaliada conforme o índice Kappa pode ser verificado na tabela a seguir:

Índice Kappa	Qualidade
< 0	Péssimo
$0 < k \leq 0,2$	Ruim
$0,2 < k \leq 0,4$	Razoável
$0,4 < k \leq 0,6$	Boa
$0,6 < k \leq 0,8$	Muito Boa
$0,8 < k \leq 1,0$	Excelente

Tabela 1. Índice Kappa pelo desempenho da classificação. Fonte: Adaptada de Landis e Koch, 1977 em Brasileiro et al, 2016.

A ferramenta *Errmat* disponível no *software Idrisi Selva* foi utilizada para realização deste cálculo.

3. RESULTADOS E DISCUSSÕES

As figuras abaixo, 2 a 8, trazem os mapas de uso e cobertura do solo na região lagunar de Maceió, Coqueiro Seco e pequena porção de Marechal Deodoro, englobando parte da APA de Santa Rita, para o período de julho de 2016.



Figura 1. Amostra Verdade



Figura 3. Classificação Agrupamento em Cadeia (Chaincluster).



Figura 5. Classificação Mínima Distância (Mindist).



Figura 4. Classificação Kmédias (kmeans).



Figura 2. Classificação Máxima Verossimilhança (Maxlike).



Figura 6. Classificação Agrupamento (cluster), não supervisionada.



Figura 7. Classificação Paralelepípedo (piped), supervisionada.

Para os mapas que foram obtidos com os métodos supervisionados criou-se uma assinatura com a seguinte quantificação de células amostradas por categoria.

Categorias	Células amostradas
Água Salgada	157576
Água Doce	10310
Água Lagunar	1318712
Remanescente Floresta	1009039
Agricultura	422319
Solo Exposto	322823
Areia	15686
Urbanizada	117956
Vias de Transporte	14080

Tabela 2. Quantidade de células amostradas por categoria na criação da assinatura.

A tabela 3, traz a o índice kappa para cada uma das categorias estabelecidas, assim como o índice global da classificação e a área de cada uma das categorias geradas na amostra verdade.

		Índice Kappa (Classe x Método)					
	Amostra Verdade (ha)	Maxlike	ChainCluster	Kmeans	Mindist	Cluster	Piped
Água Salgada	25.67	65.65%	36.41%	35.54%	51.08%	16.68%	12.15%
Água Doce	4.63	47.49%	0.00%	0.00%	31.37%	0.00%	12.50%
Água Lagunar	244.53	76.67%	69.64%	54.70%	48.04%	18.36%	10.60%
Remanescente Floresta	696.69	61.88%	70.08%	76.13%	44.91%	61.04%	14.07%
Agricultura	149.48	67.63%	0.00%	13.62%	38.20%	25.18%	20.22%
Solo Exposto	79.18	56.70%	14.97%	0.00%	13.97%	10.26%	11.44%
Areia	4.15	48.26%	73.80%	0.00%	78.01%	61.49%	44.52%
Urbanizada	114.38	56.61%	25.33%	23.92%	16.36%	24.48%	15.53%
Vias de Transporte	25.08	24.86%	0.00%	19.71%	14.99%	0.00%	8.35%
Índice Kappa	-	0.64	0.45	0.44	0.39	0.34	0.14
Qualidade	-	Muito Boa	Boa	Boa	Razoável	Razoável	Ruim

Tabela 3. Índices Kappa Global e por classe e área de cada classe na amostra verdade.

Durante o processamento das imagens algumas classes apresentaram um certo nível de confusão entre si, por isso parte deles foi subdividida, como o caso da água, e outras incorporadas por outras classes, como no caso das diferentes culturas agrícolas, que foi resumida em a uma classe. Estas alterações produziram melhores resultados de forma global para as classificações, não deixando de cobrir as principais categorias de uso e cobertura do solo. Outro ponto a ser avaliado, é a alta resolução da imagem que pode ter sido responsável por confusões geradas na classificação, visto que a riqueza de dados pode atrapalhar na criação das assinaturas com áreas puras de cada categoria, por exemplo.

Durante a classificação, notou-se que alguns métodos não foram capazes de classificar todas as categorias previamente definidas, sendo a água doce a mais omitida (pelos métodos *Chaincluster*, *Kmeans* e *Cluster*), seguida da categoria transporte (omitida pelo método *chaincluster* e *cluster*), pela agricultura omitida pelo método *chaincluster*, por fim, o solo exposto e areia que foram omitidas pelo método *kmeans*. Entre as omissões a mais

expressiva é a da classe agricultura, omitida pela *chaincluster*, mesmo possuindo uma área de aproximadamente 150 ha. Ainda assim, o índice Kappa (K) para o método *chaincluster* foi considerado bom, com $K=0,45$.

Os métodos de classificação supervisionada *mindist* e *piped* tiveram resultados razoável e ruim, respectivamente com $K=0,39$ para o *mindist*, enquanto o *piped* teve um $K=0,14$. Por outro lado, o método supervisionado *maxlike* obteve o melhor resultado entre os algoritmos analisados, sendo classificado como muito bom, com um $K=0,64$. Ao passo que, os métodos não supervisionados tiveram suas classificações consideradas boas com os classificadores *chaincluster* e *kmeans*, $K=0,45$ e $K=0,44$, respectivamente, e o classificador *cluster* foi considerado razoável, com $K=0,34$.

Finalmente, o cálculo da exatidão foi realizado com a ferramenta *ErrMat* disponibilizada no software Idrisi Selva. Através desta ferramenta foram gerados os índices de Kappa para cada um dos métodos de classificação, estes índices ficaram dentro da faixa esperada, verifica na literatura.

4. CONCLUSÃO

Em suma, foi observado que o algoritmo *maxlike* obteve o melhor resultado para a região estudada, com índice Kappa considerado muito bom na avaliação da exatidão. Outros métodos, mais rápidos e de simples aplicação, os métodos não supervisionados *chaincluster* e *kmeans*, obtiveram resultados considerados bons, entretanto esses métodos devem ser melhor avaliados, uma vez que categorias de grande relevância foram omitidas no resultado final da classificação, assim, sua aplicação para área deve ser realizada com cautela. Os demais métodos *mindist*, *cluster* e *piped*, mostraram-se inadequados para a classificação das imagens orbitais da região lagunar de Maceió e demais municípios.

De modo geral nenhum dos métodos de classificação estudado foi capaz de realizar um mapeamento das classes de maneira plenamente confiável, porém, considerando o ganho de velocidade no processamento das imagens torna-se viável a utilização do método *maxlike* para classificação do uso e cobertura do solo. Sugere-se que estudos posteriores sejam realizados com métodos mais complexos e imagens de satélites disponibilizadas de forma gratuita, uma vez que a utilização de imagens oriundas de projetos como Landsat-8, CBERS-4 e Sentinel-2 podem gerar redução nos custos dos estudos e avaliação para a região.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida A. J. P. ; Ferreira Neto J. V. (2015). Análise do processo de urbanização e impermeabilização do solo da bacia endorreica do Tabuleiro dos Martins, Maceió/AL. In: XI Encontro Nacional da ANPEGE, Dourados. Anais do XI ENANPEGE, Presidente Prudente, p. 5711-5722.

ANA - Agência Nacional de Águas. Descrição do Complexo Estuarino Lagunar Mundaú Manguaba - CELMM.

http://www2.ana.gov.br/Paginas/servicos/planejamento/planoderecursos/Celmm_Apresentacao.aspx (acedido a 16 de Janeiro de 2018).

Bispo T. C., Levino N. A. (2011). Impactos ambientais decorrentes do uso e ocupação desordenada do solo: um estudo da região da periferia de Maceió/AL. In Anais XXXI Encontro Nacional de Engenharia de Produção - Enegep, Belo Horizonte.

Brasileiro, F. G.; Oliveira, C. M. M.; Rodrigues, R. A.; Delgado, R. C. (2016). Classificação de imagem orbital pelo método máxima verossimilhança em Quixeramobim, Ceará, Brasil. Revista de Geografia Acadêmica v.10, 81 - 92.

Demarchi J. C., Sartori A. A. C., Zimback C. R. L. (2011). Métodos de classificação de imagens orbitais para o mapeamento do uso do solo: estudo de caso na Sub-Bacia do Córrego das Três Barras. In Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Curitiba. pp.2644.

Eastmans J. R. (2012). IDRISI Selva Manual. 17ª ed. Clark University. Worcester, EUA.

Figueiredo G. C., Vieira C. A. O. (2007). Estudo do comportamento dos índices de Exatidão Global, Kappa e Tau, comumente usados para avaliar a classificação de imagens do sensoriamento remoto. In Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, INPE, p. 5755-5762.

IBGE - Instituto Brasileiro de Geografia e Estatística. (2010). Cidades, Alagoas, Maceió, População.

<https://www2.ibge.gov.br/cidadesat/painel/populacao.php?codmun=270430&search=alagoas%7Cmaceio%7Cinphographics:-demographic-evolution-and-age-pyramid&lang=en> (acedido a 28 de Janeiro de 2018).

IBGE - Instituto Brasileiro de Geografia e Estatística. (2017). Brasil em Síntese, Alagoas, Maceió, Panorama. <https://cidades.ibge.gov.br/brasil/al/maceio/panorama> (acedido a 28 de Janeiro de 2018).

Moreira M. A. (2001). Fundamentos de Sensoriamento Remoto e Metodologias de Aplicação. 1ª ed. Instituto Nacional de Pesquisas Espaciais - INPE. São José dos Campos, Brasil.

Nascimento A. N. (2007). Estudo da capacidade de adensamento urbano do litoral norte de Maceió, nos bairros de Jacarecica e Guaxuma. Dissertação de Mestrado em Desenvolvimento e Meio Ambiente. Universidade Federal de Alagoas, Maceió. pp. 22-30.