

---

# ANÁLISE COMPARATIVA DE CLASSIFICADORES EM IMAGENS LANDSAT 8 (OLI) COM E SEM CORREÇÃO ATMOSFÉRICA NO ENTORNO DE PETROLÂNDIA-PE

LUIZ FILIPE CAMPOS DO CANTO<sup>1</sup>

JOÃO RODRIGUES TAVARES JUNIOR<sup>2</sup>

ANA LÚCIA BEZERRA CANDEIAS<sup>3</sup>

Universidade Federal de Pernambuco -  
UFPE Centro de Tecnologia e Geociências - CTG  
Departamento de Engenharia Cartográfica, Recife, PE  
luizfilipec@gmail.com<sup>1</sup>, joaoufpe@gmail.com<sup>2</sup>, analucia@ufpe.br<sup>3</sup>

---

**RESUMO**-Uma tarefa comum em análise dos dados nas imagens obtidos por sensores remotos é a geração de mapas temáticos a partir da classificação da cobertura terrestre. Normalmente, procura-se desenvolver diferentes algoritmos de classificação e depois aquele que apresenta o melhor desempenho, ou seja, maior acurácia é escolhido. Este tipo de metodologia pode acarretar perdas de importantes informações contidas nos classificadores descartados. O principal objetivo deste trabalho é avaliar e comparar, visual e estatisticamente, os classificadores pixel a pixel através da classificação supervisionada pelos métodos da Máxima Verossimilhança Gaussiana e Distância Mínima e a classificação não supervisionada com o método *K-médias*, com o uso de doze classes temáticas, através do uso de técnicas de Processamento de Imagem em composições, fusão e índices em imagens do satélite Landsat 8 – OLI, com valores de Número Digital e Reflectância de Topo da Atmosfera, sobre um região de desenvolvimento Sertão do Itaparica, no estado de Pernambuco.

**ABSTRACT**-A common task in data analysis on the images obtained by remote sensing is the generation of thematic maps from the land cover classification. Usually, we try to develop different ranking algorithms and then the one that performs best, ie, greater accuracy is chosen. This type of methodology can lead to loss of important information contained in discarded classifiers. The main objective of this study is to evaluate and compare, visually and statistically, the pixel classifiers pixel by supervised classification by the methods of Maximum Likelihood Gaussian and Minimum Distance and unsupervised classification with the K-means method, using twelve classes themes, through the use of image processing techniques in compositions, fusion and index images of the satellite Landsat 8 - OLI, with Digital number of values and Reflectance Top of atmosphere, on the Hinterland development region of Itaparica in the state of Pernambuco.

---

## 1. INTRODUÇÃO

Os classificadores tradicionais, tais como pixel a pixel e por regiões, vêm sendo alvo de diversos estudos comparativos, sendo basicamente dividido em duas categorias: a Classificação Supervisionada e a Não Supervisionada.

A Classificação Supervisionada depende de amostras de treino que sejam representativas das classes presentes na imagem. Assim, o algoritmo classifica os pixels para cada classe. Na Classificação Não Supervisionada os padrões de treino não se encontram classificados, onde os algoritmos têm que encontrar uma estrutura nos dados que permita segmentar grupos.

Os métodos de classificação ao nível do pixel, a que constituem objeto de estudo deste trabalho, são os métodos da Distância Mínima e Método da Máxima Verossimilhança Gaussiana, para a Classificação Supervisionada, e a Classificação Não-Supervisionada com algoritmo *K-médias*.

Para a obtenção de um bom resultado nos testes dos classificadores, foi necessária escolher um número razoável de pixels para cada amostra de treinamento da classe, com base em doze classes temáticas, para obter parâmetros comparativos entre os métodos.

As imagens produzidas por sensores orbitais sofrem influência da atmosfera. Jensen (2009) informa que diferentes fluxos de energia radiante oriunda de várias fontes podem penetrar no campo de visada introduzindo ruído no processo de imageamento por sensoriamento remoto.

A correção destas distorções atmosféricas pode ou não influir nos resultados alcançados. Deste modo, foram

---

geradas imagens com valores de reflectância dos alvos presentes, para estudar o comportamento espectral dos alvos, com o objetivo de comparar as classificações e avaliar as possíveis diferenças entre as imagens com valores dos números digitais.

Nesse contexto, nas páginas seguintes, serão expostos os procedimentos e avaliação, processos que se aplicam ao pré-processamento (para ajuste de inconsistências na imagem) e o processamento de imagem, através de composições, fusão e índices, que tem por finalidade reconhecer objetos categóricos associados a temas de interesse, proposta dos algoritmos de classificação utilizando imagens de satélite Landsat 8 - OLI da região de desenvolvimento Sertão do Itaparica, no estado de Pernambuco, a fim analisar o desempenho e efetuar um comparativo entre os métodos de classificação e consequente mapa temático.

## 2. METODOLOGIA DO TRABALHO

### 2.1 Área de estudo

O trabalho foi realizado sobre uma área teste em uma região de desenvolvimento Sertão do Itaparica, no estado de Pernambuco, incluindo o município de Petrolândia – PE e Baía de Comandantes que reunia diferentes padrões de ocupação de cobertura do solo. Localizado às margens do Rio São Francisco e uma distância de aproximadamente 499,8 km da capital pernambucana, Recife. Localiza-se a uma latitude 08°58'45" Sul e a uma longitude 38°13'10" Oeste, limitando-se ao Norte com Floresta, ao sul com Jatobá, a leste com Tacaratu e a Oeste com o Estado da Bahia. Petrolândia foi um dos locais inundados com a construção da barragem e a criação do Lago de Itaparica, há duas décadas (IFPE, s.d.).

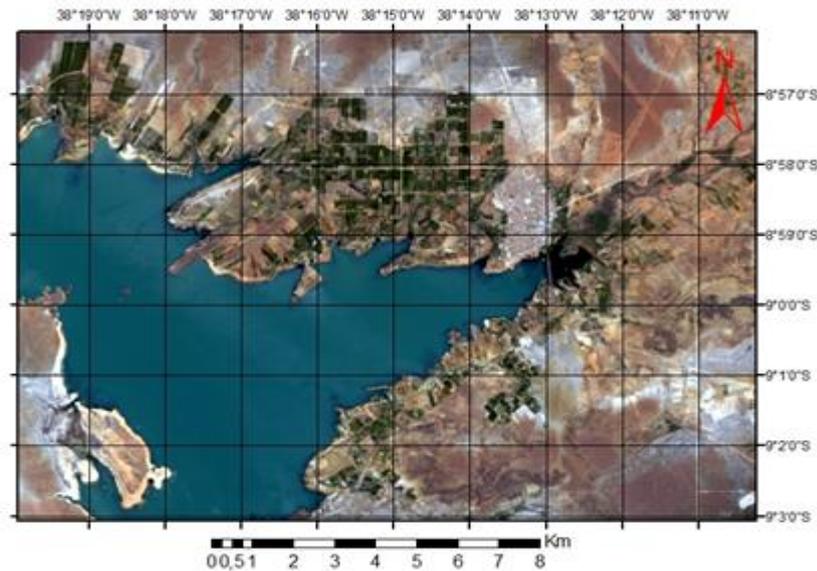


Figura 1 – Área de Estudo.

Esta área de estudo faz parte do subprojeto SP5-MR3 Land-use change inserido no Projeto INNOVATED (INterplay between the multiple use of water reservoirs via inNOVative coupling of substance cycles in Aquatic and Terrestrial Ecosystems) aprovado pelo Ministério de Ciência e Tecnologia (MCT) e pelo Ministério Alemão de Educação e Pesquisa - BMBF.

### 2.2 Imagens

Neste trabalho foram utilizadas imagens e dados do sensor OLI (*Operational Land Imager*), o qual está a bordo do satélite Landsat - 8 com resolução espacial de 30 metros e 15 metros, para a banda pancromática, com tamanho aproximado da cena de 170 km ao norte-sul por 183 km a leste-oeste, na órbita/ponto 216/66 datada do dia 02/09/2015 com horário de imageamento aproximadamente às 16h, com diferença de 15 minutos para mais ou para menos e resolução temporal de 16 dias, período de seca onde há pouca presença de nuvens.

### 2.3 Softwares

Os programas computacionais utilizados para o pré-processamento e processamento das imagens foram Quantum GIS 2.8.1 Wien, SPRING 5.3 e Google Earth.

### 2.4 Métodos

O presente trabalho apresenta um fluxograma (Figura 2) divulgando os passos desenvolvidos. No pré-processamento foram criadas novas imagens com Reflectância de Topo da Atmosfera e correção atmosférica, efetuadas a reprojeção e definidas a área de estudo, para as imagens com Número Digital e as com Reflectância de Topo da Atmosfera.

Foram efetuados os treinamentos, com o objetivo de reconhecer as assinaturas espectrais e identificar como representantes das classes em estudo. Em seguida foram aplicados nos treinamentos as Classificações Supervisionadas com algoritmos do Método da Distância Mínima e Método da Máxima Verossimilhança e a Classificação Não-Supervisionada com algoritmo *K-médias*, na composição (4R-3G-2B) das imagens com reflectância aparente e de Número Digital.

Os resultados foram divididos em: visual e numérico, sendo necessária a análise dos índices Kappa e avaliadores estatísticos para as classificações das composições. Por fim, uma análise interpretativa dos resultados, conclusões e recomendações deste trabalho.

### 2.5 Pré-processamento

Foram efetuados recortes das imagens com o intuito de permitir um melhor enfoque, nas áreas listadas no primeiro parágrafo. O enquadramento destas foi definido pelas seguintes coordenadas, superior esquerdo latitude N 9010772 m e longitude E 572961 m; canto inferior direito latitude N 9000000 m e longitude E 592000 m. Essas coordenadas serão referenciadas ao elipsóide de revolução *World Geodetic System* (WGS-84) e também utilizadas em sistemas de coordenadas Universal Transversa de Mercator (UTM).

Em seguida foi feita a reprojeção das imagens no software Quantum GIS 2.8.1, para Datum WGS 1984 UTM Zone 24S, visto que as imagens do satélite Landsat-8 são orientadas ao norte verdadeiro.

As imagens Landsat 8 possuem a correção geométrica realizada pela USGS/NASA, de modo que as imagens já são disponibilizadas ortoretificadas, não sendo assim, necessário realizar processos de georreferenciamento antes da utilização desses dados, diferentemente do que acontece com as imagens geradas pelos satélites mais antigos da série, onde a acurácia das correções foram realizadas e comparadas por Duarte et al.(2015) .

### 2.6 Correção atmosférica e transformação em reflectância

Foram criadas imagens com Reflectância de Topo da Atmosfera e classificadas com os mesmos métodos de Classificação Supervisionada e Não-Supervisionada utilizadas com as imagens com Número Digital normalmente quando são baixadas, com o objetivo de comparar os resultados obtidos das classificações com estes dois tipos de imagens.

Técnicas e programas que visam corrigir os efeitos da atmosfera vêm sendo desenvolvidos e aprimorados. Foi utilizado o complemento do programa Quantum GIS 2.8.1 chamado Semi-Automatic Classification Plugin ou SCP. Este complemento apresenta diversas funções pré-processamento e pós-processamento de imagens para vários satélites, sendo usado o método de correção atmosférica denominada DOS1 (Dark Object Subtraction) presente no complemento SCP para imagens do LANDSAT 8. Este método transforma números digitais em reflectância no topo da atmosfera.

### 2.7 Composição colorida

Foram realizadas três composições coloridas: Uma composição “Infravermelho” (Banda 5-R; Banda 4-G e Banda 3-B), “Falsa Cor 1” (Banda 6-R; Banda 5-G e Banda 4-B), “Falsa Cor 2” (Banda 7-R; Banda 5-G e Banda 3-B) e uma com “Cor Natural” (Banda 4-R; Banda 3-G e Banda 2-B). Na Figura 3 são apresentadas as composições coloridas obtidas.

Essas combinações permitem visualizar alvosem terras emersas com cores reconhecidas pela visão humana, possibilitando uma melhor visualização das diferentes feições terrestres existentes na área e ainda identificar diferentes tipos de uso, como discriminação entre área urbana, solo exposto, área de cultivo e água.

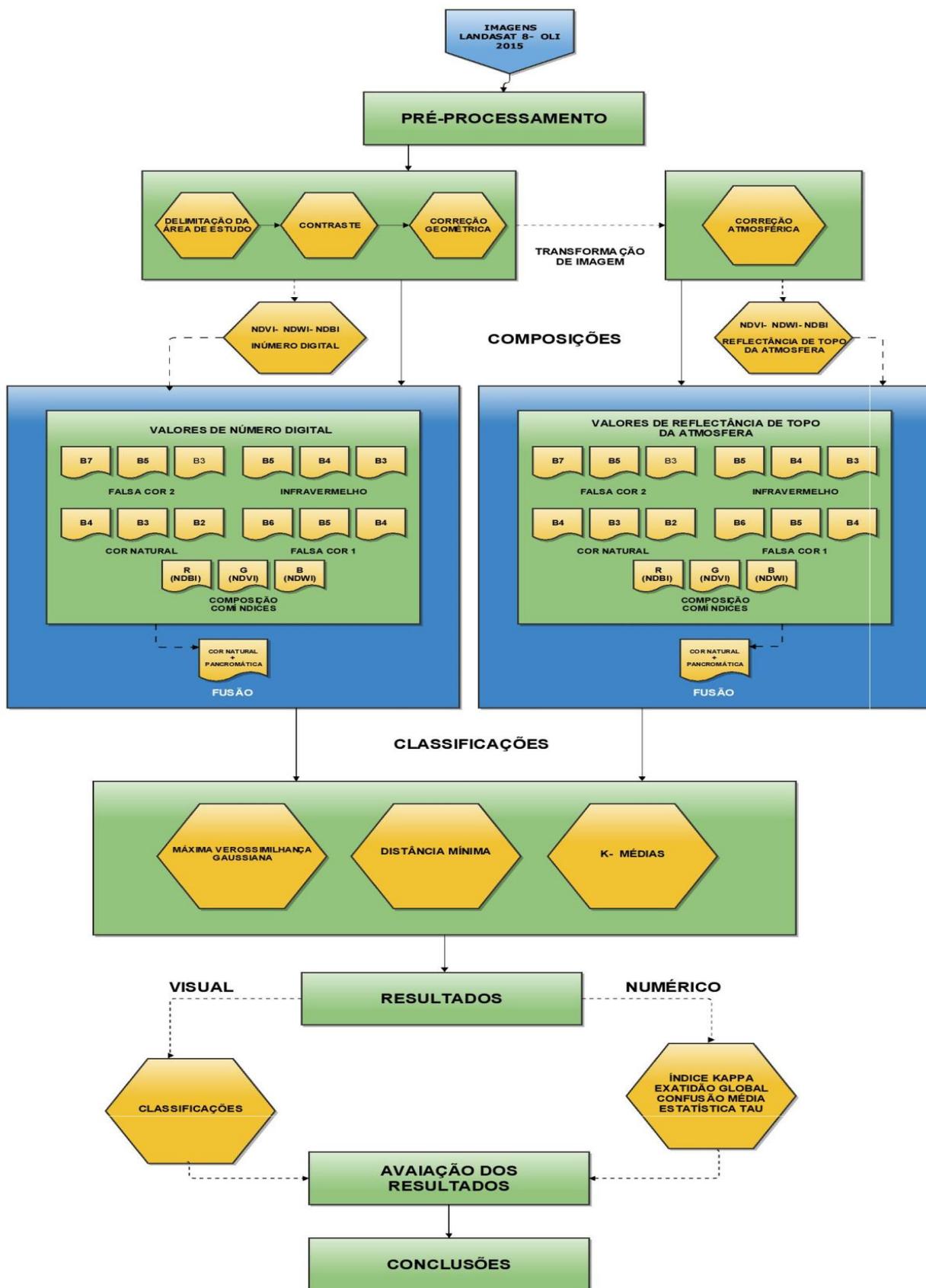


Figura 2–Fluxograma da pesquisa.

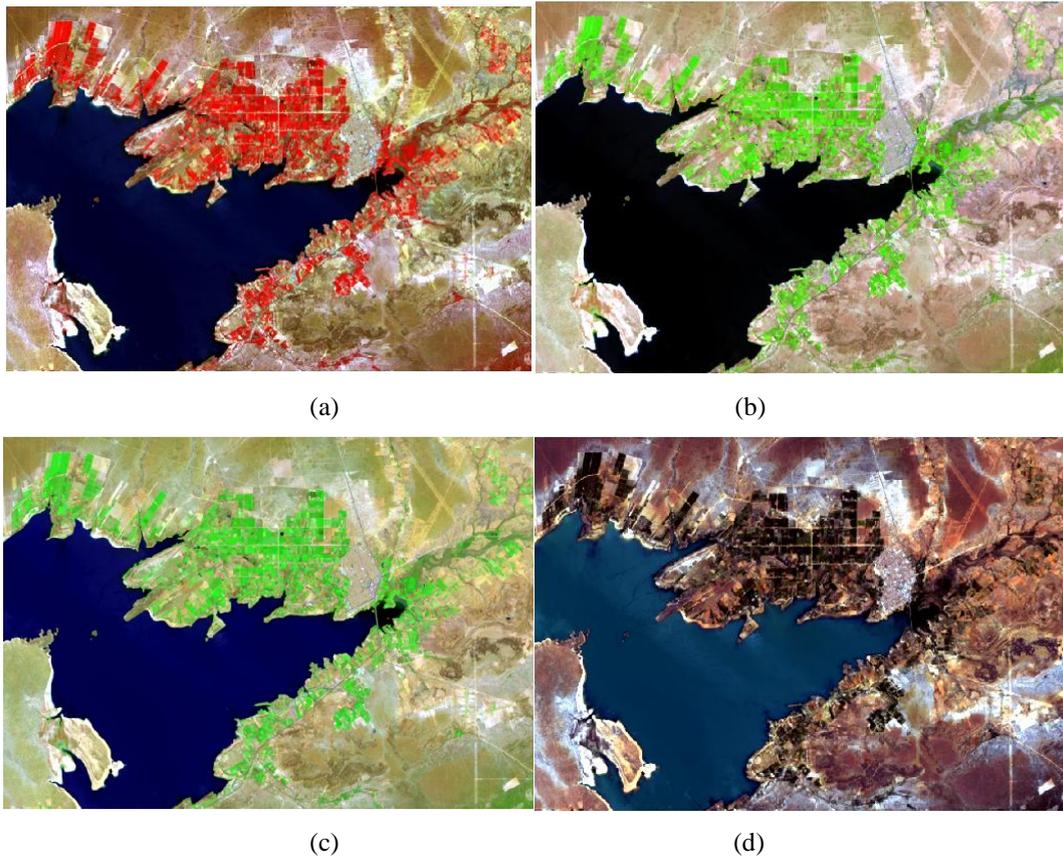


Figura 3 - Composições coloridas. (a) “Infravermelho” (5R-4G-3B). (b) “Falsa Cor 1” (6R-5G-4B). (c) “Falsa Cor 2” (7R-5G-3B). (d) “Cor Natural” (4R-3G-2B).

## 2.8 Cálculo dos índices

Para obter o índice NDVI foram usadas a banda 4 (Refletância do Vermelho - V) e banda 5 (Refletância no Infravermelho Próximo - IVP) na equação 1. Na equação 2 foram usadas a banda 3 (Refletância do Verde - Vv) e banda 5 (Refletância no Infravermelho Próximo - IVP) obtendo o índice NDWI. O índice NDBI foi obtido através do uso da banda 7 (Refletância no Infravermelho Médio - IVM) e banda 5 (Refletância no Infravermelho Próximo - IVP) na equação 3.

$$\text{NDVI} = (\text{IVP} - \text{V}) / (\text{IVP} + \text{V}) \quad (1)$$

$$\text{NDWI} = (\text{Vv} - \text{IVP}) / (\text{Vv} + \text{IVP}) \quad (2)$$

$$\text{NDBI} = (\text{IVM} - \text{IVP}) / (\text{IVM} + \text{IVP}) \quad (3)$$

Com os cálculos dos índices radiométricos obtidos, foi gerada a composição RGB, onde R = NDBI, G = NDVI e B = NDWI, que serviu como ferramenta na obtenção de contrastes relevantes na identificação de classes como uso áreas construídas, solo exposto, área de cultivo e água.

Devido à proposta dos índices físicos em realçar características de alvos de forma automatizada, a geração das imagens a partir destes cálculos (Figura 4) foram usadas para análise e identificação dos tipos de cobertura da superfície terrestre imageada, chamadas de temas. Para obter uma melhor visualização dos diferentes tipos de classes, foram realizadas manipulações de contraste nas imagens obtidas a partir do cálculo dos índices físicos.

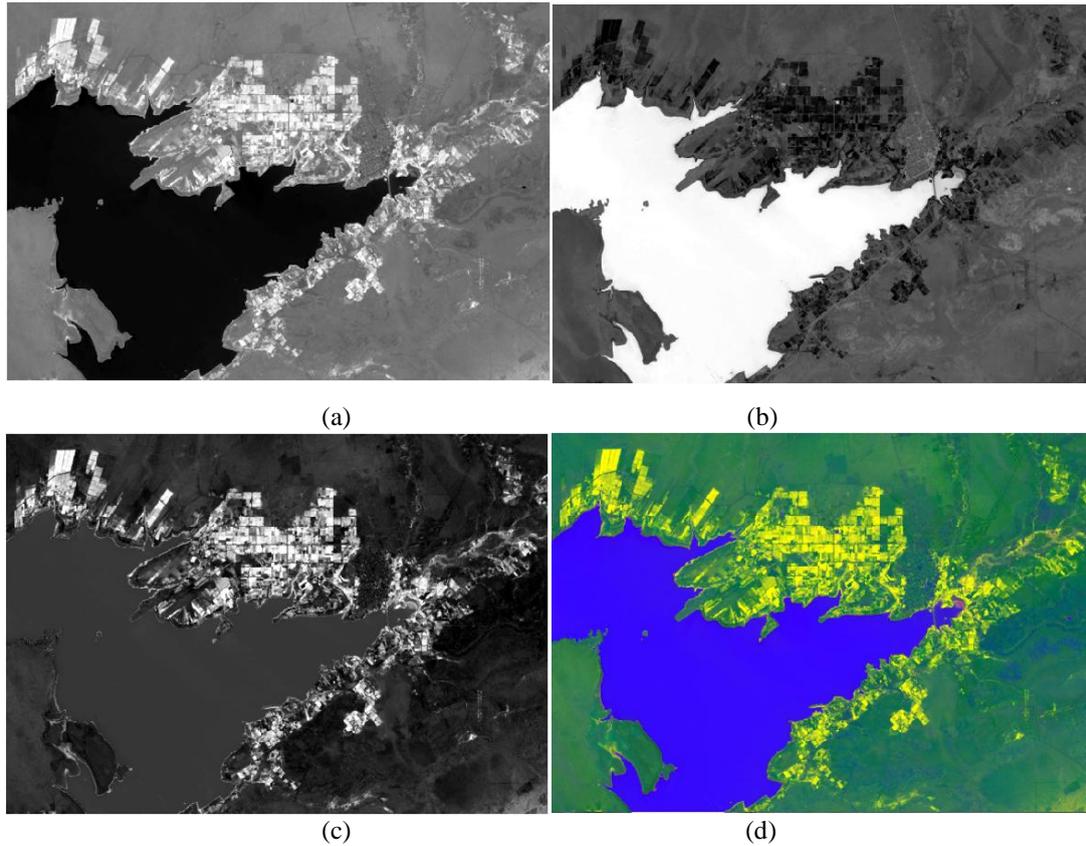


Figura 4 – Índices Físicos. (a) NDVI. (b) NDWI. (c) NDBI. (d) R (NDBI) G (NDVI) B (NDWI).

### 2.9 Definição das classes

Para avaliar o desempenho dos classificadores realizou-se a seleção das amostras de treinamento com base em doze classes informacionais, a saber: Classe de Solo I, Classe de Solo II, Classe de Solo III, Classe de Solo IV, Classe de Solo V, Classe de Solo VI, Áreas Edificadas, Tipo de Cultivo I, Tipo de Cultivo II, Classe de Água I, Classe de Água II e Classe de Água III. O número de amostras e o total de pixels coletados em cada classe são mostrados na figura 5.

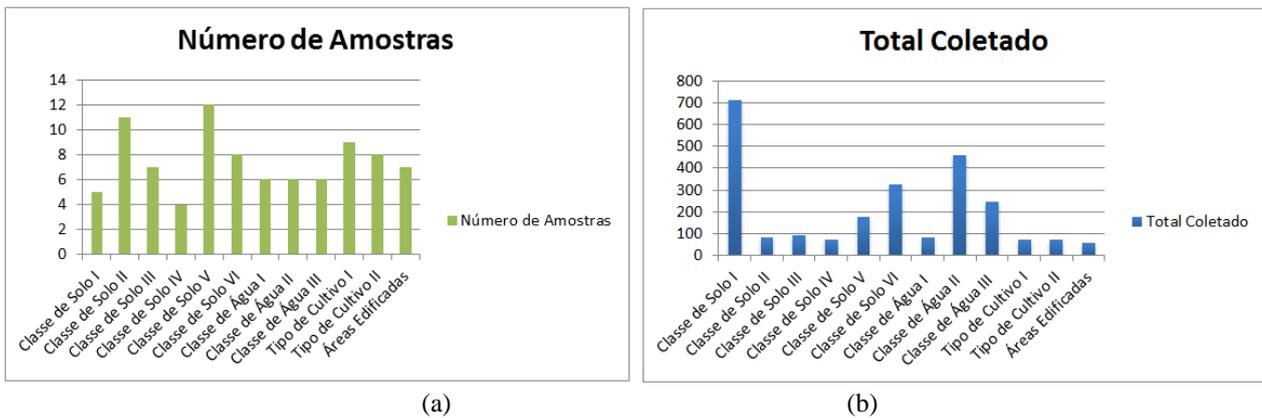


Figura 5– Número de Amostras(a) e Total de Pixels Coletados(b).

Para efetuar o comparativo dos resultados, das Classificações Supervisionadas e Não Supervisionadas, foram utilizadas as mesmas amostras selecionadas sobre as imagens de Número Digital e Reflectância de Topo da Atmosfera, sendo avaliada em cada imagem a correspondência das amostras com suas respectivas classes.

### 3. RESULTADOS E DISCUSSÕES

#### 3.1 Resultado visual

Foram utilizados os métodos de Classificação Supervisionada (Distância Mínima e Máxima Verossimilhança Gaussiana (Figura 5)) e Não Supervisionada (K-Médias (Figura 6)), para o processo de extração de informações temáticas sobre a composição Cor Natural (4R-3G-2B), com a finalidade de reconhecimento de padrões e objetos pré-adquiridos, cujos pontos das Regiões de Interesse (ROIs) estão referidos a coordenadas de imagem e cartográfica, para serem utilizados em todos os métodos de Classificação Supervisionada propostos nas composições de Cor Natural geradas por imagens com Número Digital e imagens com Reflectância de Topo da Atmosfera, tendo como objetivo final o comparativo de resultados entre os classificadores. As classificações foram realizadas no programa SPRING 5.3 com um limiar de aceitação de 100%.

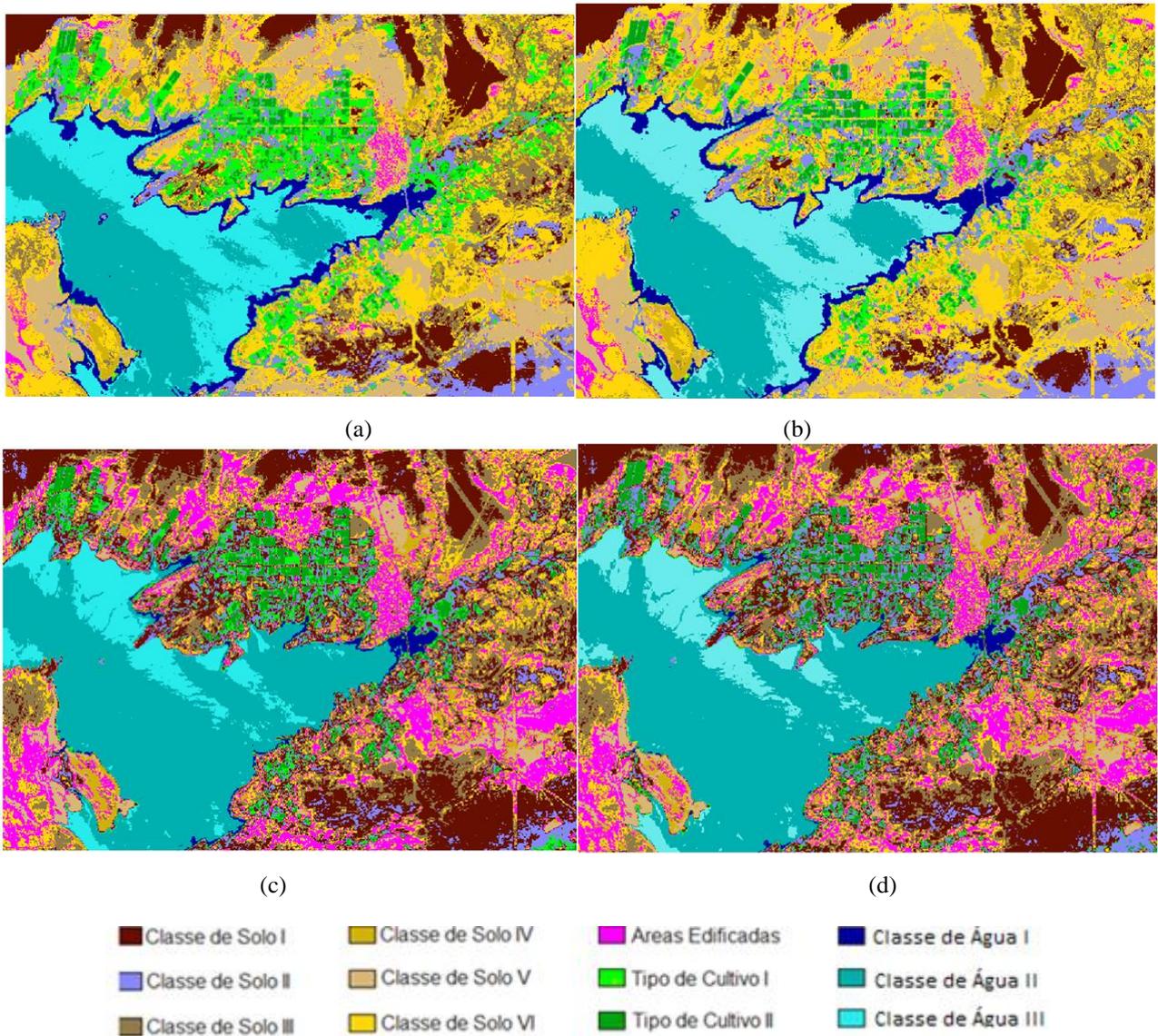


Figura 5– Classificação pelo método Distância Mínima, com o uso de doze classes. (a) Realizada na imagem com valores de Número Digital. (b) Realizada na imagem com Reflectância de Topo da Atmosfera.

O método Máxima Verossimilhança Gaussiana possui aceitável resultado na identificação de alvos cujo padrão de assinatura espectral apresenta pouca diferença, principalmente nas Classes de Solos, e excelentes resultados para as regiões com assinatura espectral com ampla diferença, apontando para regiões classes de natureza distinta, tais como corpos d’água e solo exposto. O mesmo não pode ser afirmado na classe de “Áreas Edificadas”, onde foi constatada certa quantidade de pixels classificados em uma pequena área e distribuições distintas pelo mapa que, na realidade, são

caracterizadas como “Classe de Solo V”.

A diferença entre as imagens com Número Digital e Reflectância de Topo da Atmosfera, classificadas com o método Máxima Verossimilhança Gaussiana foram poucas, tendo destaque para uma melhor identificação e distribuição dos Tipos de Cultivo e melhor identificação da realidade nas imagens com valores de Reflectância de Topo da Atmosfera.

O método da Distância Mínima, no geral, apresentou bons resultados na identificação das classes. A análise permitiu verificar a confusão nas classes de assinatura espectral com pouca diferença, nas imagens com valor de Número Digital e Reflectância de Topo da Atmosfera, destacando a “Classe de Solo III” e “Classe de Solo VI”, entre as classes “Tipo de Cultivo I” e “Tipo de Cultivo II” e identificar, também, a confusão entre as classes “Classe de Solo II” e “Tipo de Cultivo I”. Na classe “Áreas Edificadas” verifica-se, visualmente, uma intensa confusão generalizada por todo o mapa temático com a classe “Classe de Solo V”.

Uma terceira classificação foi proposta por meio de classificação Não Supervisionada, utilizando o algoritmo de k-médias, método que não requer amostras de treinamento.

O classificador necessita que o usuário determine o número de classes que os espectros serão divididos e a quantidade de iterações (quantas vezes será calculado um novo valor médio). A cada nova iteração, é gerado um novo valor médio e ao fim das iterações, cada espectro é atribuído a uma classe na qual o espectro médio está mais próximo. Para a realização da classificação foram usados como parâmetros de entrada do classificador 12 temas com 20 iterações.

No resultado da classificação com o uso de imagem com valor de Nível de Cinza e Reflectância de Topo da Atmosfera, não houve diferenças significativas. A Figura 6 mostra a classificação k-médias com o uso de doze classes realizada sobre a composição Cor Natural (4R-3G-2B).

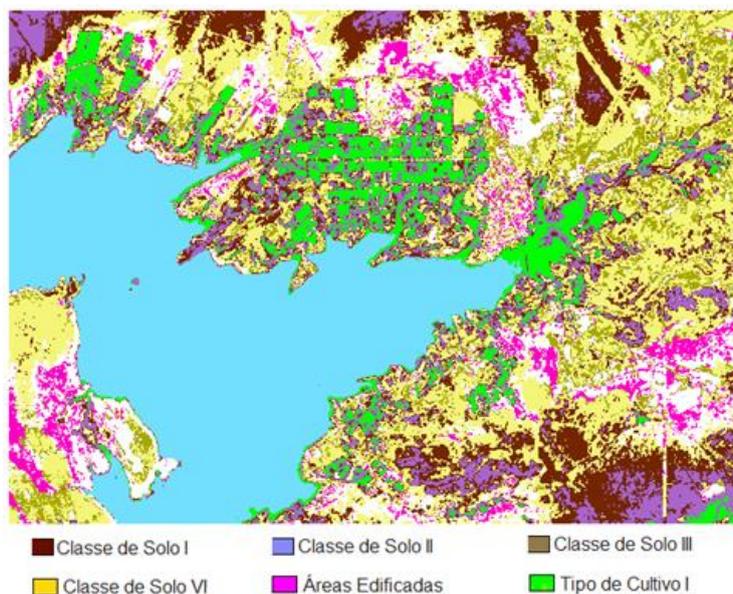


Figura 6 - Classificação Não Supervisionada pelo método k-médias com o uso de doze classes.

O algoritmo apresentou grande confusão nas separações das classes, visto que este tipo de classificação, do k-médias, inicialmente divide o conjunto de pixels em grupos, faz a média de cada grupo, e em seguida calcula a distância Euclidiana entre cada pixel do grupo e a média calculada.

No método do k-médias, para diferentes partições iniciais dos pixels têm-se diferentes agrupamentos finais. Sendo assim o algoritmo buscou automaticamente alinhar em classes as informações do recorte, mas os padrões amostrados não representaram as mesmas, criando assim classes que não corresponderam as buscadas, sendo por fim eliminadas, já que o destino é comparar as mesmas classes entre os classificadores.

Para o método com doze classes, apenas seis (Classe de Solo I, Classe de Solo II, Classe de Solo III, Classe de Solo VI, Áreas Edificadas, Tipo de Cultivo I) fizeram parte das classes escolhidas. Houve uma confusão ponderada entre as classes, tendo como destaque os conflitos entre as classes “Classe de Solo I” e “Classe de Solo II” e “Classe de Solo III” com “Classe de Solo VI”. As classes “Áreas Edificadas” e “Tipo de Cultivo I” tiveram boas classificações, mostrando maior semelhança com as feições reais obtidas pelo classificador.

### 3.2 Resultado numérico

A avaliação dos resultados foi obtida através dos índices de acurácia. Este processo consistiu no cruzamento

das informações de todas as classes resultantes e as classes observadas nas imagens de referência. Logo, obtém-se um valor global de precisão para cada resultado. Assim, torna-se possível realizar uma análise comparativa entre tais produtos.

Os dados obtidos foram tabulados em matrizes de confusão. Através destas matrizes foi possível realizar a avaliação de desempenho global através do índice Kappa, Exatidão global, Confusão Média e Estatística TAU obtidos pelos classificadores e demonstrados na Figura 7.

Não foram avaliados os resultados numéricos do método K-médias devido à incompatibilidade das classes obtidas e a não disponibilidade destas informações no programa SPRING 5.3 para este método.

São apresentados a seguir os valores dos resultados obtidos nos classificadores através dos diferentes métodos e imagens:

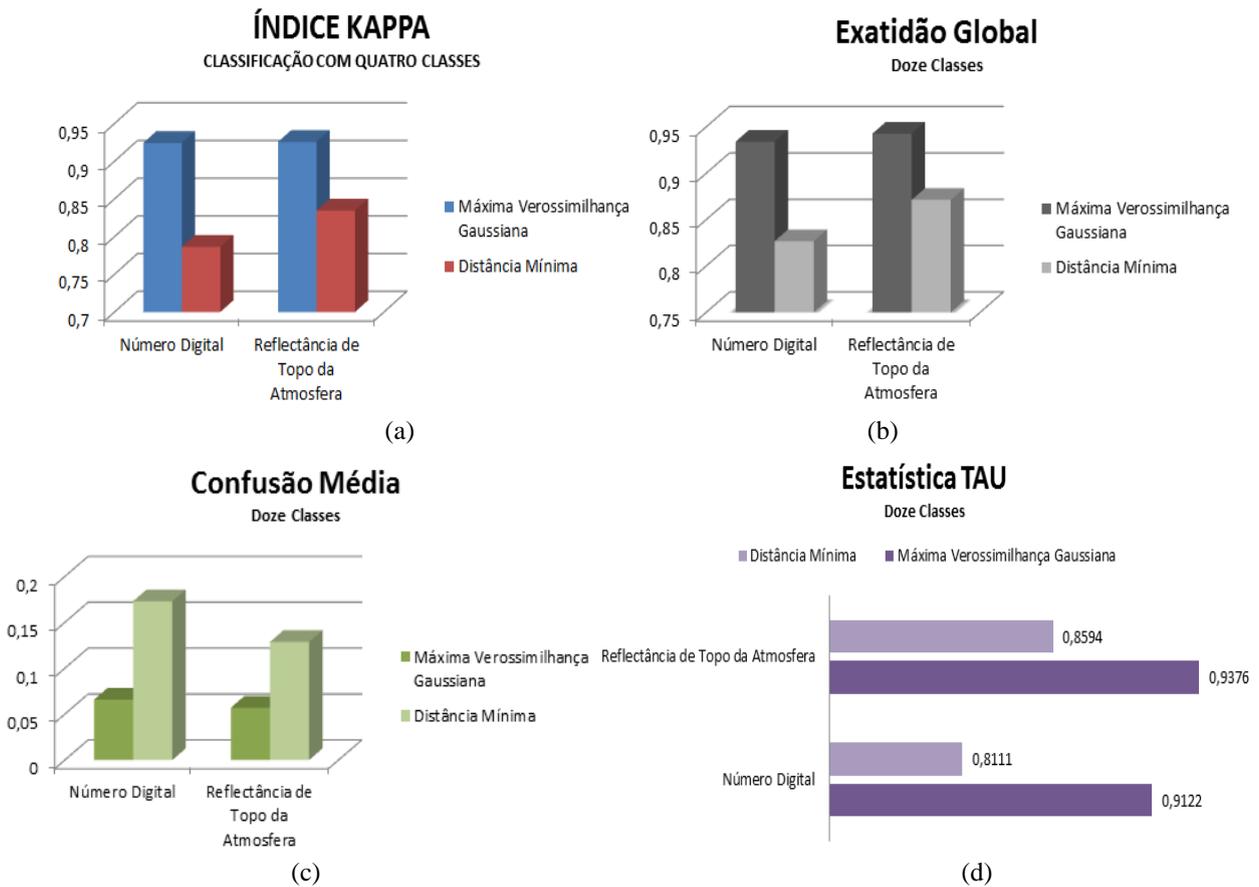


Figura 7–Valores na classificação com doze classes: Kappa (a), Exatidão Global (b), Confusão Média (c) e Estatística TAU (d).

Pode-se observar na análise dos índices Kappa, um desempenho considerado excelente na classificação com doze classes, para os métodos de Máxima Verossimilhança Gaussiana e Distância Mínima obtendo médias de 92,51% e 81,05% respectivamente, mantendo-se um desempenho excelente para o método de Máxima Verossimilhança Gaussiana com 81,82%.

Para a Exatidão Global houve valores acima de 93% no método de Máxima Verossimilhança Gaussiana e na Distância Mínima com média de 81,05%.

Nos índices TAU, foi obtido média de 91,1% para método de Máxima Verossimilhança Gaussiana e na Distância Mínima com média de 81,1%.

Os valores de coeficiente Kappa obtidos foram menor do que a Exatidão global, pelo fato do índice Kappa considerar todas as células da matriz de erros; o mesmo ocorrendo com o índice Tau.

Os resultados referentes à classificação supervisionada quando aplicado e o Método da Distância Mínima, apontam para o melhor desempenho do método de classificação Máxima Verossimilhança Gaussiana. O uso das classificações na imagem com dados de Reflectância de Topo da Atmosfera alcançou os melhores resultados em todas as etapas e métodos utilizados, comprovando a importância da correção devido à possibilidade dos alvos invariáveis, temporalmente, sejam interpretados.

#### 4. CONCLUSÕES

A análise dos resultados permite concluir que entre os mapas temáticos gerados pela classificação Supervisionada e Não Supervisionada quando aplicado o método da Distância Mínima, Máxima Verossimilhança Gaussiana e K-Médias nas imagens OLI do satélite Landsat - 8, e, considerando os resultados referentes aos índices Kappa, Exatidão global, Confusão Média e Estatística TAU, chegou-se a conclusão que, para a área de estudo, a classificação pelo método Máxima Verossimilhança Gaussiana obteve resultados excelentes, onde os mapas temáticos gerados podem ser considerados satisfatórios.

O pior desempenho dentre os classificadores foi verificado pelo classificador K-médias, devido as diferentes classificações finais e uma qualidade questionável nas classes correspondentes.

A imagem com dados de Reflectância de Topo da Atmosfera alcançaram os melhores resultados em todas as etapas e métodos utilizados, comprovando a importância da correção devido à possibilidade dos alvos invariáveis, temporalmente, sejam interpretados.

A coleta das amostras de treinamento é importante para o resultado final, contudo, a escolha adequada do método de classificação é fundamental para se obter um resultado

Houve dificuldade na separabilidade entre a classe de Áreas Edificadas nas imagens utilizadas neste estudo em todos os métodos. A tentativa de distinção destas classes, pelo método de Distância Mínima e K-Médias, aumentou a confusão e diminuiu a qualidade da classificação, devido a baixa resolução que dificulta discriminar estas áreas com algumas classes de solos.

Estes métodos de classificação estudados que utilizam apenas as características espectrais são incapazes de diferenciar de forma precisa classes que possuem semelhança espectral tais como solo e edificações, sendo necessária a implementação de novos métodos que visem classificar de forma mais acurada.

#### Agradecimentos

PIBIC / CNPq pela bolsa de apoio para o primeiro autor e ao projeto INNOVATE (*Innovate Interplay between the multiple use of water reservoirs via innovative coupling of substance cycles in aquatic and terrestrial ecosystems*).

#### Referências Bibliográficas

Duarte, G. S.; Costa G. E.; Oliveira A. P.; Barros, R. S.; Cruz C. B. M. **O uso da geometria do Landsat 8 como base para georreferenciamento semiautomático visando estudos espaço-temporais.** In: Simpósio Brasileiro de Sensoriamento Remoto, João Pessoa. Anais. São José dos Campos, SP: INPE, p. 5216-5223, 2015.

IFPE. **Petrolândia**, s.d. Disponível em: <[http://www.ifsertao-pe.edu.br/reitoria/pro-reitorias/prodi/observatorio/micronegiao\\_itaparica/petrolandia.pdf](http://www.ifsertao-pe.edu.br/reitoria/pro-reitorias/prodi/observatorio/micronegiao_itaparica/petrolandia.pdf)> Acesso em Dezembro de 2015.

Jensen, J. R. **Sensoriamento Remoto do Ambiente: uma perspectiva em recursos terrestres.** 2ed. São José dos Campos: Parêntese, 2009.