



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE TECNOLOGIA E RECURSOS NATURAIS  
UNIDADE ACADÊMICA DE ENGENHARIA CIVIL  
CAMPUS CAMPINA GRANDE - PB

AVALIAÇÃO DA CLASSIFICAÇÃO DA COBERTURA DO SOLO NO BIOMA  
CAATINGA POR ANÁLISE MULTITEMPORAL E UNITEMPORAL

**João Maria de Andrade**

**Orientadores: Profa. Dra. Iana Alexandra Alves Rufino**

**Prof. Dr. John E. de Brito Leite Cunha**

CAMPINA GRANDE – PB

2018

JOÃO MARIA DE ANDRADE

AVALIAÇÃO DA CLASSIFICAÇÃO DA COBERTURA DO SOLO NO BIOMA  
CAATINGA POR ANÁLISE MULTITEMPORAL E UNITEMPORAL

Trabalho de Conclusão de Curso apresentado ao Curso  
de Engenharia Civil da Universidade Federal de  
Campina Grande como requisito para obtenção do  
grau de bacharel em Engenharia Civil.

Área de habilitação: Recursos Hídricos

**Orientadores: Profa. Dra. Iana Alexandra Alves Rufino**

**Prof. Dr. John E. de Brito Leite Cunha**

CAMPINA GRANDE - PB

2018

JOÃO MARIA DE ANDRADE

AVALIAÇÃO DA CLASSIFICAÇÃO DA COBERTURA DO SOLO NO BIOMA  
CAATINGA POR ANÁLISE MULTITEMPORAL E UNITEMPORAL

**BANCA EXAMINADORA**

Profa. Dra. Iana Alexandra Alves Rufino  
(Orientadora – Universidade Federal de Campina Grande)

Prof. Dr. John Elton de Brito Leite Cunha  
(Coorientador – Universidade Federal de Campina Grande)

Prof. Dr. Carlos de Oliveira Galvão  
(Examinador interno -Universidade Federal de Campina Grande)

Profa. Dra. Rochele Sheila Vasconcelos  
(Examinadora externa - Pós-doc – Capes/ANA - PPGECA  
Universidade Federal de Campina Grande)

CAMPINA GRANDE

2018

*Dedico esta, aos meus pais, Manoel Maria Neto e  
Benedita de Andrade Maria, minhas maiores  
inspirações.*

## AGRADECIMENTOS

Quero agradecer primeiramente a Deus, que me guiou em todos os momentos e que agraciou minha vida com Seu amor infinito, pela oportunidade de cursar e concluir este curso.

Aos meus pais Manoel Maria Neto e Benedita de Andrade Maria que através de seus esforços me auxiliaram ao longo da minha caminhada, sempre demonstrando a importância da dedicação e resiliência em todos os aspectos da vida, e às minhas irmãs Moanna Maria de Andrade e Polianna Maria de Andrade que sempre estiveram ao meu lado, me apoiando e orientando quando necessário.

Quero agradecer também a companhia dos amigos e colegas que conheci durante estes anos, e aos docentes que me ensinaram este ofício tão nobre.

E também agradeço à minha companheira de curso e de vida, Carol, por todo o apoio e carinho que me proporcionou durante todo este tempo.

Agradeço ao Prof. John Cunha que me acolheu e incentivou a pesquisa no laboratório de Hidráulica 2, me mostrando toda a grandeza da ciência em minha jornada na universidade, e também por todas as orientações e conselhos. E à minha orientadora Professora Iana Alexandra, minha gratidão pelos ensinamentos, confiança e disponibilidade.

Agradeço também ao Laboratório de Hidráulica – II “BU” e todos seus integrantes, por proporcionar a realização da minha pesquisa e meu crescimento pessoal.

## LISTA DE FIGURAS

Figura 1: Obtenção de imagens por Sensoramento Remoto .....	16
Figura 2: Cronologia histórica do Programa Landsat.....	17
Figura 3: Árvore de Desciões: Método <i>Random Forest</i> .....	24
Figura 4: Matriz de Confusão.....	26
Figura 5 : Localização geográfica da bacia hidrográfica do Rio Sucuru.....	30
Figura 6 : Fluxograma simplificado das etapas metodológicas da classificação do uso do solo .....	31
Figura 7: Amostragem dos pontos das classes do uso do solo da bacia do Rio Sucuro .	32
Figura 8: Classificação da vegetação quanto ao porte.....	33
Figura 9: Classificação das plantas quanto à classe. ....	34
Figura 10: Classificação de comunidades vegetais quanto ao grau de recobrimento da superfície do solo.....	34
Figura 11: Reflectância e índices de vegetação EVI, NDVI e SAVI os tipos de cobertura do solo para uma imagem no período chuvoso (Março de 2016). ....	36
Figura 12: Reflectância e índices de vegetação EVI, NDVI e SAVI os tipos de cobertura do solo para uma imagem no período de seca (setembro de 2016). ....	37
Figura 13: Valor de NDVI para as classes de cobertura do solo.....	38
Figura 14: Valor de Swir-2 para as classes de cobertura do solo .....	38
Figura 15: Classificação do uso do Solo - Série temporal 24 meses EVI .....	39
Figura 16: Classificação do uso do Solo - Série temporal 24 meses NDVI.....	40
Figura 17: Classificação do uso do Solo - Série temporal 24 meses SAVI.....	41
Figura 18: Classificação do uso do Solo - Série temporal 24 meses STI.....	42
Figura 19: Classificação do uso do Solo - Série temporal 24 meses Swir-II .....	43
Figura 20: Classificação Unitemporal do uso do Solo em Setembro de 2016 .....	44
Figura 21: Resumo de acurácia geral e índice kappa .....	45
Figura 22: Classificação do uso do Solo: A- Série temporal 24 meses NDVI, B - Unitemporal com as bandas espectrais (09/2016) .....	46

## LISTA DE TABELAS

Tabela 1 : Bandas (canais) do Satélite LandSat 8 .....	19
Tabela 2 : Pontos base para processamento dos dados para obtenção da classificação da vegetação da Bacia do Sucuri.....	35
Tabela 3: Acurácia do Índice de Vegetação EVI .....	39
Tabela 4 : Acurácia do Índice de Vegetação NDVI .....	40
Tabela 5: Acurácia do Índice de Vegetação SAVI.....	41
Tabela 6: Acurácia do Índice de Vegetação STI .....	42
Tabela 7: Acurácia do Índice de Vegetação Swir-II.....	43
Tabela 8: Acurácia da Classificação Unitemporal – Bandas espectrais .....	44
Tabela 9: Áreas das Classes de Uso do Solo Análise Multitemporla x Análise Unitemporal .....	47

## LISTA DE EQUAÇÕES

Equação 1: EVI .....	20
Equação 2: NDVI .....	21
Equação 3: SAVI .....	18
Equação 4: STI .....	19
Equação 5: Acurácia Global .....	23
Equação 6: Acurácia do Produtor .....	24
Equação 5: Acurácia do Usuário .....	25
Equação 6: Índice Kappa .....	25



## RESUMO

A caatinga é um importante bioma Brasileiro, o qual, abrange 11 % de seu território. É caracterizada por possuir uma grande diversidade de espécies vegetais, que apresentam grande perda de folhagens durante a estação de seca, alterando sua morfologia em um curto período de tempo. A classificação da cobertura do solo desse bioma é de grande importância, pois pode auxiliar no estudo de bacias hidrográficas, análises ambientais e observação da alteração na cobertura do solo, já que esse ambiente vem sofrendo alterações severas ao longo dos anos. Alguns mecanismos que podem ser utilizados para realizar seu monitoramento são as técnicas de sensoriamento remoto e geoprocessamento, que amparados com métodos computacionais (*machine learning*) e dados supervisionados, tornam possível a obtenção de uma classificação da cobertura do uso do solo para esse bioma. Esse estudo foi aplicado na área da Bacia do rio Sucuru, no Cariri paraibano, em que a obtenção da classificação da cobertura do solo dessa região foi feita por meio da criação de um modelo computacional baseado no método *Random Forest*, utilizando o software R. Foram utilizadas imagens dos satélites Landsat 8 e Landsat 7 fornecidas pela USGS, e os índices de vegetação EVI, NDVI, SAVI, STI. A série temporal de 24 meses de NDVI demonstrou ser mais eficiente do que as demais séries para os outros índices para classificação da cobertura do solo. Além disso, foi mais eficiente que a tradicional classificação unitemporal (realizando com as bandas espectrais), possibilitando, assim, a identificação de um padrão para cada tipo de cobertura do solo. A diferença percentual entre a utilização da série temporal de NDVI e a classificação unitemporal foi de +241,03% para classe ASMD - Arbórea Subarbórea Muito Densa, já a classe que apresentou maior decréscimo foi a ATSA - Arbustiva subarbórea Aberta com - 44,32%. Essas variações poderão afetar o desempenho de modelos hidrológicos e climatológicos que utilizam dados de cobertura do solo. Mapas com melhor representação da cobertura do solo serão instrumentos importantes para avaliar as mudanças na cobertura do solo ao longo do processo de ocupação das áreas de Caatinga.

**Palavras-Chave:** *Random forest*, Landsat, Séries Temporais; índices de vegetação.

## ABSTRACT

The caatinga is an important Brazilian biome, which covers 11% of its territory. It is characterized by a great diversity of plant species, which present great loss of foliage during the dry season, changing its morphology in a short period of time. The classification of the land cover of this biome is of great importance, since it can help in the study of hydrographic basins, environmental analyzes and observation of the alteration in the soil cover, since this environment has undergone severe alterations throughout the years. Some mechanisms that can be used to carry out its monitoring are the remote sensing and geoprocessing techniques, which, based on machine learning methods and supervised data, make it possible to obtain a classification of the land use cover for this biome. This study was applied in the area of the Sucuru River Basin, in Cariri, Brazil, where the classification of the land cover of this region was made by means of the creation of a computational model based on the Random Forest method, using the R. software. Images of the Landsat 8 and Landsat 7 satellites provided by the USGS, and the vegetation indices EVI, NDVI, SAVI, STI. The 24-month NDVI time series proved to be more efficient than the other series for the other indices for land cover classification. In addition, it was more efficient than the traditional instantaneous classification (performing with the bands), thus allowing the identification of a standard for each type of soil cover. The percentage between the temporal series NDVI and classification instantaneous was + 241,03% for the ASMD - *Arbórea Subarbórea Muito Densa* class, which was already classified as an ATSA - *Arbustiva subarbórea Aberta* - 44.32%. These deficits affect the performance of hydrological and climatological models that use land cover data. The maps covered with soil cover are important to evaluate the changes in the soil area throughout the process of occupancy of the Caatinga areas.

**Keywords:** Random forest, Landsat, Time Series; indices of vegetation.

## SUMÁRIO

1. INTRODUÇÃO .....	12
1.1 Justificativa .....	13
1.2 Hipótese da Pesquisa .....	14
1.3 Objetivos .....	14
1.3.1 Objetivo Geral .....	14
1.3.2 Objetivos Específicos .....	14
2. FUNDAMENTAÇÃO TEÓRICA .....	15
2.1 Caatinga .....	15
2.2 Geotecnologias e Sensoriamento Remoto .....	15
2.3 Satélites da série Landsat .....	16
2.3.1 Landsat 8 .....	18
2.4 Índices de Vegetação .....	19
2.4.1 EVI .....	20
2.4.2 NDVI .....	21
2.4.3 SAVI .....	21
2.3.4 STI .....	22
2.5. Técnicas de Classificação Supervisionada .....	22
2.5.1 Método <i>Random Forest</i> .....	23
2.5.2. Classificador de floresta aleatório e dados multi-fonte .....	25
2.6 Acurácia .....	25
2.6.1 Matriz de Confusão .....	26
2.6.2 Acurácia do Produtor .....	27
2.6.3 Acurácia do Usuário .....	27
2.6.4 Índice Kappa .....	27
3. MATERIAIS E MÉTODOS .....	29
3.1 Materiais .....	29
3.3 Aquisição de dados .....	30
3.3 Processamento das Imagens de Satélite Landsat .....	35
3.5 Análise da Acurácia .....	35
4. RESULTADOS E DISCUSSÕES .....	36
5. CONCLUSÃO .....	48
6. RECOMENDAÇÕES .....	49
REFERÊNCIAS .....	50
APÊNDICES .....	54

## 1.INTRODUÇÃO

A Caatinga, com aproximadamente os mesmos limites da área de clima semiárido do Nordeste brasileiro, é um dos maiores e mais desconhecidos biomas do mundo. Sua variada cobertura vegetal está, em grande parte, determinada pelo clima, relevo e embasamento geológico que, em suas múltiplas interrelações, resultam em ambientes ecológicos bastante variados (RODAL, et al., 2008).

Apesar da sua importância, o bioma tem sido desmatado de forma acelerada, principalmente nos últimos anos, devido principalmente ao consumo de lenha nativa, explorada de forma ilegal e insustentável, para fins domésticos e industriais, ao sobrepastoreio e a conversão para pastagens e agricultura. Frente ao avançado desmatamento que chega a 46% da área do bioma, segundo dados do Ministério do Meio Ambiente (MMA), o governo busca concretizar uma agenda de criação de mais unidades de conservação federais e estaduais no bioma, além de promover alternativas para o uso sustentável da sua biodiversidade (Ministério do Meio Ambiente, 2017).

Assim, as técnicas de geoprocessamento e sensoriamento remoto podem auxiliar no planejamento e no controle ambiental, assumindo um papel importante na detecção e monitoramento de vários fenômenos meteorológicos, proporcionando uma melhor compreensão do comportamento associado às suas respostas (BORGES, et al., 2016).

Já que as decisões estratégicas exigem uma grande quantidade de informações que podem ser facilmente analisadas com o uso destas ferramentas, tais como: o uso do solo e a influência do relevo, do clima e da hidrografia nas características físicas e ocupacionais de uma dada região. A utilização das técnicas de geoprocessamento vem se tornando uma alternativa viável e confiável nas metodologias de aquisição de dados e classificação do solo, reduzindo consideravelmente as deficiências relativas ao cumprimento das leis (NASCIMENTO, et al., 2005).

O objetivo desse estudo é realizar a classificação do uso solo dessa região utilizando técnicas classificadores supervisionados que são amplamente utilizados na literatura, uma vez que são mais robustos do que abordagens baseadas em modelos (Niemeyer et al., 2014). Esses classificadores podem aprender as características do alvo, classes de amostras de treinamento e identificar essas características aprendidas nos dados não classificados. Um classificador supervisionado eficiente precisa enfrentar os desafios (Millard e Richardson, 2015) envolvidos em (i) mitigar o *fenômeno de Hughes* (ou seja,

a "maldição" de dimensionalidade"), que ocorre quando o número de variáveis é muito maior do que o número de amostras de treino (ii) lidar com não-linearidade das variáveis, (iii) lidar com amostras de treinamento desequilibrado e ruído em ambas as amostras de treino e dados não marcados, e (iv) redução do tempo de computação (BELGIU & DRAGURT, 2016).

Deste modo surge como alternativa para identificação do uso do solo no semi-árido, a aplicação de técnicas automatizadas utilizando o método de “*random forest*” pois estes modelos computacionais, baseados em métodos de aprendizagem supervisionados, criam um padrão de classificação a partir de dados de treinamento (VALBUENA, et al., 2016). Assim pode-se obter uma classificação para o uso do solo da região através de coletas de dados em campo e pela criação de um modelo supervisionado.

### **1.1 Justificativa**

O levantamento do uso do solo é de grande importância, na medida em que os efeitos do mau uso causam deterioração no ambiente. Os processos de erosão intensos, as inundações, os assoreamentos desenfreados de reservatórios e cursos d'água são consequências do mau uso deste solo. Dessa maneira, o mapeamento do uso da terra e cobertura vegetal, mediante a utilização de técnicas de geoprocessamento, representam importante instrumento para o planejamento e administração da ocupação do meio físico, possibilitando a avaliação e o monitoramento do mesmo, a fim de garantir a conservação de seus recursos naturais (ASSIS, 2014).

Do ponto de vista da hidrologia o fator do uso do solo (cobertura vegetal) e às chuvas de alta intensidade (curta duração), em microbacias, possuem grande sensibilidade para esses efeitos, sendo considerados fatores dominantes na geração do deflúvio, sendo assim, as alterações na quantidade e qualidade da água do deflúvio, em função de chuvas intensas e ou em função de mudanças no solo, são detectadas com mais sensibilidade nas microbacias do que nas grandes bacias. (TEODORO, et al., 2007).

Assim, a caracterização da cobertura do solo na Caatinga pode auxiliar na gestão dos recursos hídricos. pois gera importantes indicadores ambientais como uso e ocupação de solo que poderiam funcionar como instrumentos de diagnóstico para auxiliar os tomadores de decisão. (SILVA *et al.*, 2017).

## 1.2 Hipótese da Pesquisa

A análise multitemporal dos índices de vegetação utilizando técnicas de “*machine learning*” (algoritmo “*random forest*”) - pode ser mais eficiente que as técnicas tradicionais de classificação supervisionada que utilizam recorte unitemporal para obter uma classificação de cobertura do solo da Caatinga e, portanto, pode auxiliar na caracterização das bacias hidrográficas.

## 1.3 Objetivos

### 1.3.1 Objetivo Geral

Este estudo tem como objetivo principal determinar uma classificação da cobertura do Solo na Bacia do Sucuru (PB), utilizando técnicas de “*machine learning*” – “*random forest*” (algoritmo random forest).

### 1.3.2 Objetivos Específicos

- Aperfeiçoar técnicas para classificação do uso do solo;
- Verificar a acurácia e índice Kappa dos índices de vegetação para classificação do solo e banda Swir-II;
- Verificar a acurácia e índice Kappa da classificação do uso do solo obtida por recorte unitemporal e multitemporal;
- Verificar a cobertura do solo na bacia do Sucuru;
- Validar uma classificação para cobertura do solo da Caatinga, baseado nesta bacia representativa.

## **2. FUNDAMENTAÇÃO TEÓRICA**

### **2.1 Caatinga**

A Caatinga ocupa uma área de cerca de 844.453 *Km*<sup>2</sup>, o equivalente a 11% do território nacional. Engloba os estados Alagoas, Bahia, Ceará, Maranhão, Pernambuco, Paraíba, Rio Grande do Norte, Piauí, Sergipe e o norte de Minas Gerais. Rico em biodiversidade, o bioma abriga 178 espécies de mamíferos, 591 de aves, 177 de répteis, 79 espécies de anfíbios, 241 de peixes e 221 abelhas (Ministério do Meio Ambiente, 2017).

Esse bioma apresenta um grande potencial para a conservação de serviços ambientais, uso sustentável e bioprospecção que, se bem explorado, será decisivo para o desenvolvimento da região e do país. A biodiversidade da Caatinga ampara diversas atividades econômicas voltadas para fins agrosilvopastoris e industriais, especialmente nos ramos farmacêutico, de cosméticos, químico e de alimentos. Apesar da sua importância, o bioma tem sido desmatado de forma acelerada, principalmente nos últimos anos, devido principalmente ao consumo de lenha nativa, explorada de forma ilegal e insustentável, para fins domésticos e industriais, ao sobrepastoreio e a conversão para pastagens e agricultura (Ministério do Meio Ambiente, 2017).

“A Caatinga, com aproximadamente os mesmos limites da área de clima semi-árido do nordeste brasileiro, é uma das maiores e mais desconhecidas. Sua variada cobertura vegetal está, em grande parte, determinada pelo clima, relevo e embasamento geológico que, em suas múltiplas interrelações, resultam em ambientes ecológicos bastante variados.” (RODAL, et al., 2008, p. 210).

### **2.2 Geotecnologias e Sensoriamento Remoto**

O estudo do espaço geográfico e dos aspectos ambientais nele inseridos pressupõe uma série de conhecimentos e informações que podem ser trabalhadas de maneira mais ágil, fácil e rápida com as novas tecnologias. Assim as geotecnologias irão ocupar esse lugar de destaque devido a sua funcionalidade (FITZ, 2008).

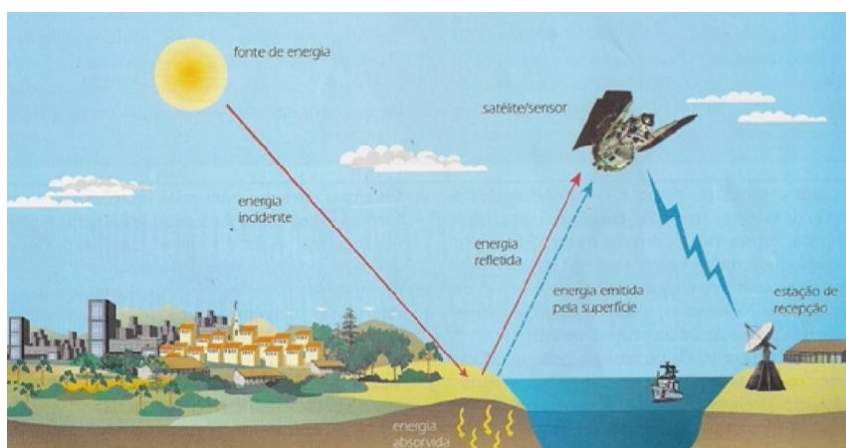
As geotecnologias têm se tornado cada vez mais abrangentes e alcançado um maior número de pessoas, pois, na contemporaneidade, uma grande parte da população está em contato com as diversas tecnologias, como por exemplo, o sensoriamento remoto,

o GPS (Global Positioning System) de campo e/ou automotivo, Google Earth, dentre outros (SILVA, 2017).

Dessa forma, as geotecnologias podem ser entendidas como sendo “as novas tecnologias ligadas às geociências, as quais trazem avanços significativos no desenvolvimento de pesquisas, em ações de planejamento, em processos de gestão, manejo e em tantos outros aspectos relacionados à estrutura do espaço geográfico” (FITZ, 2008).

Sensoriamento remoto é uma técnica de obtenção de imagens dos objetos da superfície terrestre sem que haja um contato físico de qualquer espécie entre o sensor e o objeto (SALLES, et al., 1998), como ilustrado na Figura 1.

Figura 1: Obtenção de imagens por Sensoriamento Remoto



Fonte: (FLORENZANO, 2007)

A condição principal imposta por essa definição clássica, que é o sensor estar a uma distância remota do objeto, estabelece a base para definir o sensoriamento remoto numa concepção um pouco mais científica, que é regida segundo os seguintes preceitos: i) exigência: ausência de matéria no espaço entre o objeto e o sensor; ii) consequência: a informação do objeto é possível de ser transportada pelo espaço vazio; iii) processo: o elo de comunicação entre o objeto e o sensor é a radiação eletromagnética, a única forma de energia capaz de se transportar pelo espaço (MENESES & ALMEIDA, 2012).

### 2.3 Satélites da série Landsat

O programa da série LANDSAT (Land Remote Sensing Satellite) foi desenvolvido pela NASA (National Aeronautics and Space Administration), iniciou-se



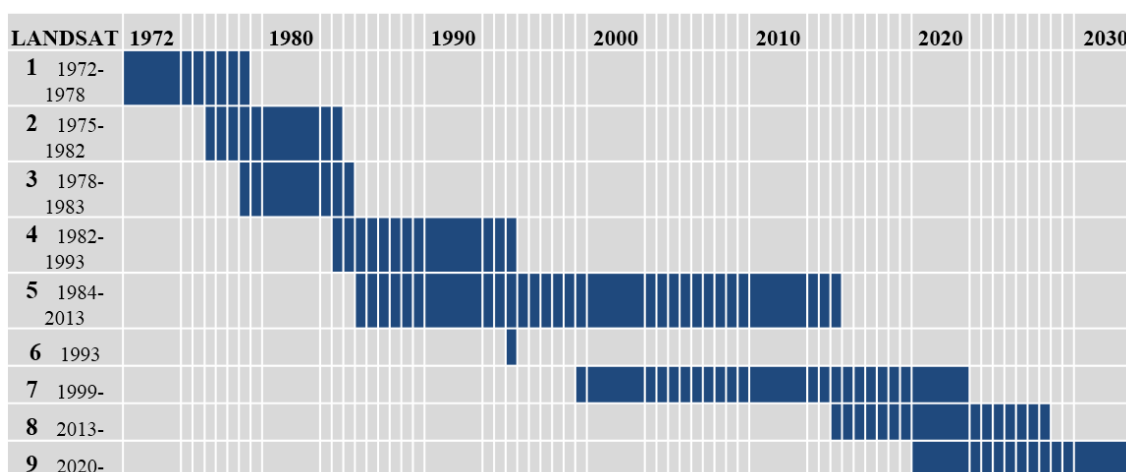
em 1972 com o lançamento do satélite ERTS-1 (Earth Resources Technology Satellite), passando a ser denominado LANDSAT em 1975 (SALLES, et al., 1998).

O Landsat 1 foi lançado em 1972 e foi seguido por uma série de satélites Landsat, (Landsat 2, 3, 4, 5 e 7), que proporcionaram observações reflexivas e de ondas térmicas quase globais com maior fidelidade espectral e espacial. Notavelmente, o registro do Landsat é contínuo, com a maioria dos locais de terra adquiridos pelo menos uma vez por ano desde 1972, capturando um período em que a população humana global mais do que duplicou e a evidência das mudanças climáticas tornou-se discernível (HANSEN, 2012).

O principal objetivo do sistema Landsat foi o mapeamento multispectral em alta resolução da superfície da Terra. Esse foi, e é de longe, o sistema orbital mais utilizado na Embrapa para o monitoramento por satélite no mapeamento da dinâmica espaço-temporal do uso das terras e em todas as aplicações decorrentes. A antena do INPE em Cuiabá recebe de forma contínua imagens de todo o território nacional, desde os anos setenta, e isso constitui para um enorme e único acervo de dados sobre o país (EMBRAPA, 2008).

Em mais de 40 anos, a série de satélites Landsat fornece o registro temporal mais longo de observações de superfície baseadas em espaço Figura 2.

Figura 2: Cronologia histórica do Programa Landsat



Fonte: (OLIVEIRA, 2018)

Notavelmente, o registro do Landsat é contínuo, com a maioria dos locais de terra adquiridos pelo menos uma vez por ano desde do ano de seu lançamento, capturando um período em que a população humana global mais do que duplicou e a evidência das mudanças climáticas tornou-se discernível (HANSEN, 2012).

Os dados Landsat oferecem um registro exclusivo da superfície terrestre e sua modificação ao longo do tempo. A resolução espacial moderada do Landsat é suficiente para permitir a cronologia de mudanças antropogênicas e naturais da escala local à global e, as séries temporais de dados, são calibradas para fornecer um registro consistente e caracterizado que é necessário para permitir a discriminação entre artefatos de dados e mudanças temporais reais da superfície terrestre (ROY, 2014). Eles demonstraram capacidades para o mapeamento e monitoramento da cobertura terrestre e propriedades biofísicas e geofísicas da superfície terrestre (HANSEN, 2012) e potencial de utilidade para a assimilação terrestre e ciclagem biogeoquímica e terra use aplicações de previsão.

As aplicações abordadas com os dados Landsat envolvem descobertas científicas, gerenciamento e monitoramento de recursos para a qualidade econômica e ambiental, saúde pública e bem-estar humano e segurança nacional. As análises dos benefícios econômicos do Landsat variam de US \$ 935 milhões / ano para US \$ 2,19 bilhões / ano (MILLER, 2013) em apoio a aplicações, incluindo análise e gerenciamento de recursos hídricos, análise e gerenciamento de agricultura e floresta, segurança interna, análise de infraestrutura, gerenciamento de desastres, ciência das mudanças climáticas, proteção de zonas húmidas e monitoramento de terra mudança de cobertura.

### **2.3.1 Landsat 8**

O satélite Landsat 8 foi o último lançamento da série Landsat, onde os objetivos gerais da missão são: fornecer continuidade de dados com os Landsats 4, 5 e 7, oferecendo uma resolução temporal de 16 dias da Terra, uma resolução temporal de 8 dias com um deslocamento Landsat 7, construir e atualizar periodicamente um arquivo global de imagens terrestres iluminadas pelo sol, substancialmente sem nuvens. (US Geological Survey, 2018)

O satélite Landsat 8 mede diferentes faixas de frequências ao longo do espectro eletromagnético, cada faixa espectral é chamada de banda, e o Landsat 8 tem 11 bandas (LOYD, 2013).

De suas 11 bandas, apenas aquelas com os comprimentos de onda mais curtos (faixas 1-4 e 8) detectam a luz visível - todas as outras estão em partes do espectro que não podemos ver (Tabela 1).

Tabela 1 : Bandas (canais) do Satélite Landsat 8

<b>Número da Banda</b>	<b><math>\mu\text{m}</math></b>	<b>Resolução Espacial</b>
<b>1</b>	<i>0.433–0.453</i>	<b>30 m</b>
<b>2</b>	<i>0.450–0.515</i>	<b>30 m</b>
<b>3</b>	<i>0.525–0.600</i>	<b>30 m</b>
<b>4</b>	<i>0.630–0.680</i>	<b>30 m</b>
<b>5</b>	<i>0.845–0.885</i>	<b>30 m</b>
<b>6</b>	<i>1.560–1.660</i>	<b>30 m</b>
<b>7</b>	<i>2.100–2.300</i>	<b>30 m</b>
<b>8</b>	<i>0.500–0.680</i>	<b>15 m</b>
<b>9</b>	<i>1.360–1.390</i>	<b>30 m</b>
<b>10</b>	<i>10.6–11.2</i>	<b>100 m</b>
<b>11</b>	<i>11.5–12.5</i>	<b>100 m</b>

Fonte: (LOYD, 2013)

## 2.4 Índices de Vegetação

Os índices de vegetação, indicam a abundância relativa e a atividade da vegetação verde, incluindo índice de área foliar (IAF), porcentagem de cobertura verde, teor de clorofila, biomassa verde e radiação fotossinteticamente ativa absorvida (RFAA, ou APAR, em inglês). Sendo assim, os índices de vegetação são também importantes indicadores da qualidade ambiental, além de serem bastante utilizados nos estudos sobre o meio ambiente (DEMARCHI, et al., 2011).

Os índices de vegetação, possibilitam extrair informações biofísicas da vegetação, os quais têm como objetivo, em princípio, condensar informações espectrais diferenciando a vegetação da superfície não vegetada, bem como avaliar as condições de crescimento das culturas, ocorrências de doenças, pragas, secas, geadas e diversos eventos catastróficos meteorológicos. (SILVA, 2017).

Os cálculos dos índices de vegetação, que são medidas radiométricas adimensionais, foram criados para distinguir as informações espectrais da vegetação em relação às demais superfícies da terra e indicar a quantidade e qualidade da vegetação em uma área imageada (JENSEN, 2011). Portanto, orientam o manejo adequado dos recursos florestais e permitem analisar se a exploração dos recursos está de acordo com a capacidade de recuperação dos sistemas agrícolas, afetando a comercialização e o valor do produto no mercado (LEITE, et al., 2017).

Os índices de vegetação utilizam as faixas do vermelho e do infravermelho próximo por conterem mais de 90% da variação da resposta espectral da vegetação e,

portanto, tais índices realçam o seu comportamento espectral, correlacionando-os com os parâmetros biofísicos da mesma, as faixas espectrais correspondentes ao vermelho (0,65 a 0,72  $\mu m$ ) e ao azul (0,45 a 0,50  $\mu m$ ) são as mais aproveitadas pela folha na fotossíntese, pois os pigmentos da planta responsáveis por ela absorvem a radiação incidente nos comprimentos de onda 0,48  $\mu m$  (carotenóides) e 0,62  $\mu m$  (clorofila). Dessa forma, a reflectância é muito baixa nestas faixas. Já na faixa do infravermelho próximo (0,7 a 1,3  $\mu m$ ), a absorção da radiação eletromagnética é muito baixa (menos de 15 %) e a reflectância pode chegar a 50% dependendo da estrutura anatômica das folhas. A radiação solar sofre difusão e dispersão através da cutícula e da epiderme das células do mesófilo e das cavidades de ar no interior da folha. As múltiplas reflexões e refrações da radiação solar nestas estruturas devem-se à diferença entre o índice de refração do ar e o das paredes de celulose hidratadas, provavelmente em razão da superfície rugosa das microfibrilas das paredes celulares (DEMARCHI, et al., 2011).

#### 2.4.1 EVI

Em inglês, este índice é denominado “Enhanced Vegetation Index” e, por isso, sua sigla original é “EVI” (JENSEN, 2011). Foi desenvolvido para otimizar o sinal de vegetação com sensibilidade melhorada em regiões de alta biomassa e monitoração de vegetação melhorada através de um desacoplamento do sinal de fundo do dossel e uma redução nas influências da atmosfera. A equação assume a forma da Equação (1).

$$EVI = G\rho_{NIR} - pred\rho_{NIR} + C1 \times pred - C2 \times \rho_{blue} + L \quad (Eq.1)$$

onde  $\rho$  são *refletores de superfície* corrigidos atmosféricamente ou parcialmente corrigidos pela atmosfera (Rayleigh e absorção de ozônio),  $L$  é o ajuste do fundo do dossel que aborda a transmissão não linear e diferencial NIR e vermelho através de um dossel e  $C1$ ,  $C2$  são os coeficientes do aerossol termo de resistência, que usa a banda azul para corrigir as influências do aerossol na banda vermelha. Os coeficientes adotados no algoritmo EVI são:  $L = 1$ ,  $C1 = 6$ ,  $C2 = 7,5$  e  $G$  (fator de ganho) = 2, (HUETE, 2002).

O conceito de "resistência atmosférica" é baseado na dependência do comprimento de onda dos efeitos de aerossol, utilizando a banda azul mais sensível à atmosfera para corrigir a banda vermelha para influências de influências de aerossóis. O objetivo é encontrar uma função,  $f(\lambda)$  ( $\rho_{\lambda}$ ,  $\rho_{\text{azul}}$  -  $\rho_{\text{vermelho}}$ ), que é estável contra variações de condição de aerossol atmosférico. Verificou-se que o EVI funcionou bem no aerossol pesado, condições de queima de biomassa no Brasil (MIURA, 1998).

### 2.4.2 NDVI

Em inglês, esse índice é denominado “Normalized Difference Vegetation Index” (NDVI). As informações sobre conteúdo de umidade interno à vegetação são de grande utilidade na agricultura, floresta e hidrologia e assume a forma da Equação (2).

$$NDVI = \frac{Nir - Red}{Nir + Red} \quad (\text{Eq. 2})$$

O índice de vegetação diferenciada normalizada (NDVI), uma transformação não-linear da relação de reflectância do infravermelho próximo (Nir) para o vermelho (Red), foi desenvolvido para melhorar o sinal de vegetação em baixas condições de biomassa. Infelizmente, esse aprimoramento é alcançado à custa de sensibilidade reduzida na parte superior, ou alta faixa de condições de biomassa. Como resultado, o NDVI muitas vezes "satura" em áreas de alta biomassa, como em biomas florestados (Huete, et al., 1997).

### 2.4.3 SAVI

Em inglês, esse índice é denominado “Soil Adjusted Vegetation Index – SAVI”. O índice de vegetação ajustado por solo é obtido através da introdução de um fator no NDVI para incorporar o efeito da presença do solo, mantendo-se o valor NDVI dentro de -1 a +1 (HUETE, 1988).

$$SAVI = \frac{NIR - Red}{NIR + Red + L} \times (1 + L) \quad (\text{Eq. 3})$$

Os três valores de L, representando as três condições da superfície – solo nu, cobertura média da vegetação e cobertura da vegetação, são listados a seguir:

L = 0,5 para cobertura média de vegetação;

L = 1,0 para a superfície do solo nu ou pouca vegetação;

L = 0,25 para a cobertura completa de vegetação.

O SAVI com o valor de L de 0,5 representa melhor que NDVI nas várias condições da vegetação. (LIU, 2006)

### 2.3.4 STI

Em inglês, Soil Tillage Index (STI) (VAN DEVENTER, 1997) é um índice empírico construído com diferentes bandas Landsat TM (Eq 4)

$$STI = \rho M6 / \rho M7 \quad (Eq. 4)$$

O STI pode ser aplicado para monitorar massa de tecidos de vegetação seca em muitas áreas semi-áridas. A influência de pastagem na estação seca na decomposição da massa seca, a influência do clima na variabilidade na produção vegetal, bem como as mudanças abruptas causadas pelo fogo (JACQUES *et al.*, 2014).

### 2.5. Técnicas de Classificação Supervisionada

Classificadores paramétricos supervisionados, como Máxima Verossimilhança (MLC), oferece excelentes resultados ao lidar com dados unimodais. No entanto, eles têm limitações ao lidar com conjuntos de dados de entrada multimodais porque esses classificadores assumem um padrão de distribuição de dados (Liu & Zhang, 2011). Os classificadores supervisionados não parametrizados como a Árvore de Classificação e Regressão (CART), Support Vector Machine (SVM) e Artificial Rede Neural (ANN) não fazem quaisquer suposições quanto à distribuição de frequência e tornaram-se cada vez mais populares para classificação de dados detectados remotamente, que raramente possuem distribuições normais (BELGIU & DRAGURT, 2016)

Com classificação simples a natureza e as causas da variação espacial nas imagens não são entendidas, e a análise das imagens fica limitada a associação empírica entre fenômeno superficial e padrões em imagens, com a suposição implícita de que a realidade tem uma resposta espectral consistente em imagens. Esta suposição é muitas vezes violada, no entanto, como consequência da complexa interação entre fatores como a complexidade da cena, a escala e a agregação. Portanto, os classificadores simples podem atingir seus limites em muitas aplicações (BELGIU & DRAGURT, 2016).

Nos últimos anos, a atenção da comunidade de sensoriamento remoto voltou-se para classificadores de conjuntos (GISLASON & SVEINSSON, 2006). Esses classificadores podem ser baseados em um indivíduo classificador supervisionado ou em vários classificadores supervisionados diferentes que são treinados usando ensacamento ou ampliando abordagens ou aproximadas variações destes. Na abordagem de ensacamento também conhecido como *bootstrap - abordagem de agregação* - cada classificador no conjunto é treinado em um subconjunto aleatório de um conjunto de

amostras de treinamento, enquanto que no impulso aborda os classificadores de conjuntos são treinados iterativamente usando todas as amostras de treinamento, aumentando os pesos para amostras classificadas incorretamente durante o procedimento de treinamento. O trabalho mostrou que o uso de métodos de conjunto de empacotamento e ensacamento obteve maior precisão do que usar classificadores únicos, como classificadores de árvores de decisão bem como, sendo mais estável e robusto ao ruído nos dados de treinamento (BELGIU & DRAGURT, 2016).

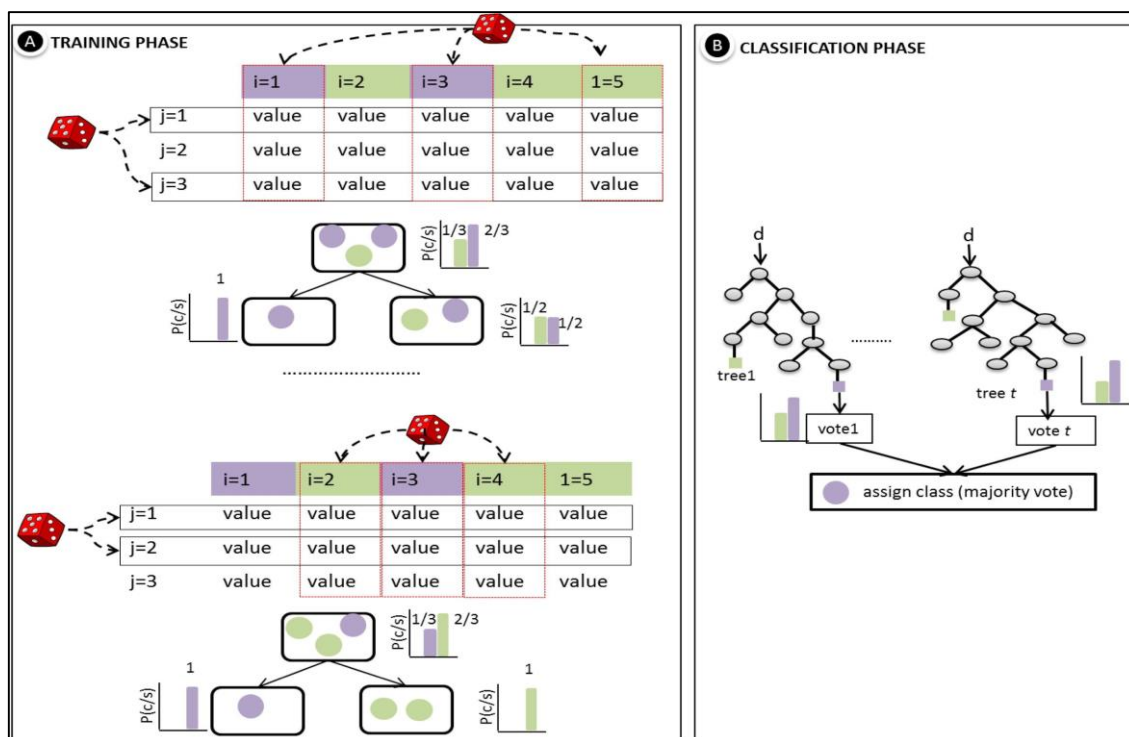
### **2.5.1 Método *Random Forest***

O classificador de Random Forest(RF) são uma combinação de preditores das árvores, de tal forma que cada árvore depende dos valores de um vetor aleatório da amostra de forma independente e com a mesma distribuição para todas as árvores na floresta (BREIMAN, 2001). As árvores são criadas desenhando um subconjunto de amostras de treinamento através da substituição (uma abordagem de ensacamento). Isso significa que a mesma amostra pode ser selecionado várias vezes, enquanto outros não podem ser selecionados (Fig. 3A)(BELGIU; DRAGURT, 2016).

Cerca de dois terços das amostras (referidas como amostras no saco) são usados para treinar as árvores (Fig. 3A) com o terceiro restante (referidas como amostras fora do saco) são usadas em uma técnica interna de validação cruzada para estimar o desempenho do modelo RF executado (BREIMAN, 2001).

Essa estimativa de erro é conhecida como erro de saída. Cada árvore de decisão é produzida independentemente sem qualquer poda e cada nó é dividido usando um número de recursos definido pelo usuário, selecionado ao acaso. Ao crescer a floresta até um usuário definido número de árvores (Ntree), o algoritmo cria árvores que têm alta variância e baixa tendência (BREIMAN, 2001).

A decisão de classificação final é tomado pela média (usando a média aritmética) a classe probabilidades de atribuição calculadas por todas as árvores produzidas. Uma nova entrada de dados não marcados é, portanto, avaliada em relação a todas as árvores de decisão criadas no conjunto e cada árvore vota para uma associação de classe. A classe de membros com o máximo de votos será finalmente selecionado (Fig. 3B) (BELGIU & DRAGURT, 2016)

Figura 3: Árvore de Decisões: Método *Random Forest*

Fonte: (BELGIU & DRAGURT, 2016)

O classificador de RF foi bem sucedido quando usado para o mapeamento de derrames de óleo de dados SAR (*Synthetic Aperture Radar*), para classificações *land cover* (LC), para mapear biomassa usando dados temporais Landsat, para classificar superfícies impermeáveis urbanas, entre outras aplicações (BELGIU & DRAGURT, 2016).

Uma árvore de classificação consiste em gerar uma classificação gradual ao dividir a variância explicada de forma recursiva, em uma sucessão de divisões binárias de limiares de preditores determinados de acordo com o critério de minimização residual de mínimos quadrados (VALBUENA, et al., 2016).

Vários estudos analisaram o potencial da classificação de RF para melhorar a classificação de objetos urbanos através de dados aéreos- LiDAR (*Light Detection And Ranging*) e também para o mapeamento de deslizamentos de reflorestamento usando variáveis calculadas para cada pixel ou para objetos de imagem delineados por segmentação. Alguns estudos também exploraram o uso de florestas aleatórias na classificação do veículo aéreo não tripulado (UAV) e em sensoriamento remoto térmico (SUN & SCHULZ, 2015).



A classificação dos dados hiperespectrais é desafiada por sua alta dimensionalidade. Esse problema é amplificado pela quantidade relativamente pequena dos dados de treinamento disponíveis, o que resulta em um aumento no número de classificações erradas. As aplicações de sensoriamento remoto, portanto, requerem soluções que são capazes de remover bandas de ondas redundantes e ruidosas. O classificador de RF foi testado como uma otimização de espaço de recursos técnica com resultados satisfatórios em inúmeras aplicações, tais como o mapeamento de espécies arbóreas (Cavallaro, et al., 2015), pinheiros atingido por relâmpagos, espécies de capim e níveis de defoliação de insetos em Árvore mopane (BELGIU & DRAGURT, 2016).

### **2.5.2. Classificador de floresta aleatório e dados multi-fonte**

Os dados de sensoriamento remoto nem sempre são suficientes para classificar objetos alvo. Um número de estudos avaliou o impacto do uso qualquer de uma imagem adquirida por diferentes sensores ou removida dados combinados com dados geográficos auxiliares, como informações derivado de dados de elevação digital (GISLASON & SVEINSSON, 2006) ou classes de solo (CORCORAN, et al., 2013) sobre a precisão da classificação.

Dada a diversidade de conjuntos de dados de entrada que poderiam ser usados para melhorar a classificação, é importante incluir apenas os conjuntos de dados mais relevantes a fim de reduzir a carga computacional sem sacrificar a precisão dos resultados (CORCORAN, et al., 2013).

O classificador *Random Forest* pode ser usado com sucesso para classificar a detecção remota multisource e dados geográficos, principalmente por causa de sua velocidade computacional (CORCORAN, et al., 2013) e porque ajuda os analistas de imagens a otimizar o modelo de classificação usando apenas os conjuntos de dados de entrada identificados como importantes para a classificação do objeto.

## **2.6 Acurácia**

Acurácia está relacionada à veracidade de obtenção de dados em um intervalo de probabilidade especificado (LOBÃO, 2003). Segundo (WATZLAWICK, et al., 2002) a acurácia geral representa a relação entre o número de pontos classificados corretamente e o número total de pontos amostrais utilizados. É expressado por:

$$P_0 = \frac{\sum_{i=1}^K x_{ii}}{N}, \quad (\text{Eq } 5)$$

$P_0$ = Acurácia Global

$x_{ii}$ = elementos da diagonal principal da matriz de confusão;

$N$ = número total de pontos da matriz de confusão

### 2.6.1 Matriz de Confusão

A Matriz de Confusão (Figura 4) é a ferramenta mais utilizada em avaliações de acurácia de mapeamentos, além dos coeficientes dela derivados. Segundo (SPIEGEL, 1993) cada frequência observada na matriz corresponde, também, a frequência esperada a qual é calculada sob um determinada hipótese conforme as regras de probabilidade

Figura 4: Matriz de Confusão

		Referência de campo					Total	
		Classes	1	2	3	4		5
Mapeamento	1	$P_{1,1}$					$P_{1,5}$	$\sum_{j=1}^5 P_{j,i}$
	2		$P_{2,2}$					
	3			$P_{3,3}$				
	4				$P_{4,4}$			
	5	$P_{5,1}$					$P_{5,5}$	
	Total	$\sum_{i=1}^5 P_{i,j}$						$\sum_{i=1}^5 P$

Fonte: (BRITO, 2009 )

A diagonal da matriz ( $x_{ii}$ ) apresenta a frequência observada e representa a concordância entre o esperado e observado em cada classe temática. A coluna que margeia a matriz representa o total de pontos da imagem para cada classe ( $x_{i+}$ ), a linha marginal representa o número de pontos encontrados para cada classe ( $x_{+i}$ ) em campo.

Segundo (ANTUNES & LINGNAU, 1997) os coeficientes de concordância total quanto os de concordância, para classes individuais, podem ser utilizados para avaliar a acurácia de um mapeamento.

### 2.6.2 Acurácia do Produtor

Acurácia do produtor refere-se aos erros de exclusão ou omissão, não sendo atribuídos à determinada classe os erros de classificação das demais classes. Segundo (Antunes, 2003) a acurácia do produtor refere-se à probabilidade de determinada classe ter sido corretamente classificada de acordo com os pontos de referência de campo. Pode variar entre 0 e +1 sendo expresso por:

$$Ap_i = \frac{x_{ii}}{x_{i+}}, \quad (Eq. 6)$$

onde:

$x_{ii}$  = elementos da diagonal principal da matriz de confusão;

$x_{i+}$  = somatório da linha para dada classe temática

### 2.6.3 Acurácia do Usuário

A acurácia do usuário refere-se aos erros de inclusão ou comissão (CONGALTON, 1991). Os de inclusão relacionam-se a todos os pontos atribuídos a uma determinada classe, oriundos de erros de classificação de outras classes, indicando a probabilidade de uma classe classificada representar a mesma em campo.

Este índice pode assumir valores entre 0 e + 1, sendo expresso por (ANTUNES & LINGNAU, 1997).

$$Au_i = \frac{x_{ii}}{x_{+i}}, \quad (Eq.7)$$

$x_{ii}$  = elementos da diagonal principal da matriz.

$x_{+i}$  = somatório da coluna para dada classe temática.

### 2.6.4 Índice Kappa

O índice Kappa é um coeficiente de concordância para escaladas nominais que mede a proporção de concordância, depois que a concordância atribuída à casualidade é

retirada da consideração. Ele é expresso pela seguinte equação segundo (CONGALTON, 1991).

$$K = \frac{P_o - P_{co}}{1 - P_{co}} \quad ; \quad (\text{Eq.8})$$

onde :

$K$  = coeficiente de concordância Kappa;

$$P_o = \frac{\sum_{i=1}^M n_{ii}}{N} \quad ; \quad \text{, representa a proporção de pontos de verdade terrestre concordantes;}$$

$$P_{co} = \frac{\sum_{i=1}^M n_{i+} n_{+i}}{N^2} \quad ; \quad \text{, representando a proporção de pontos atribuído à classe ao acaso;}$$

$N$  = número total de pontos da matriz de confusão;

$n$  = elemento da matriz de confusão;

$n_{ii}$  = elementos da diagonal principal da matriz de confusão;

$n_{i+}$  = somatório da linha para dada classe temática;

$n_{+i}$  = somatório da coluna para dada classe temática.

### 3. MATERIAIS E MÉTODOS

#### 3.1 Materiais

- Série de Imagens de satélite Landsat 7 e 8 (no período entre outubro de 2014 a setembro de 2016);
- Software Qgis 2.8;
- Programa Google Earth® (Google Systems);
- Software R *statistical environment*;

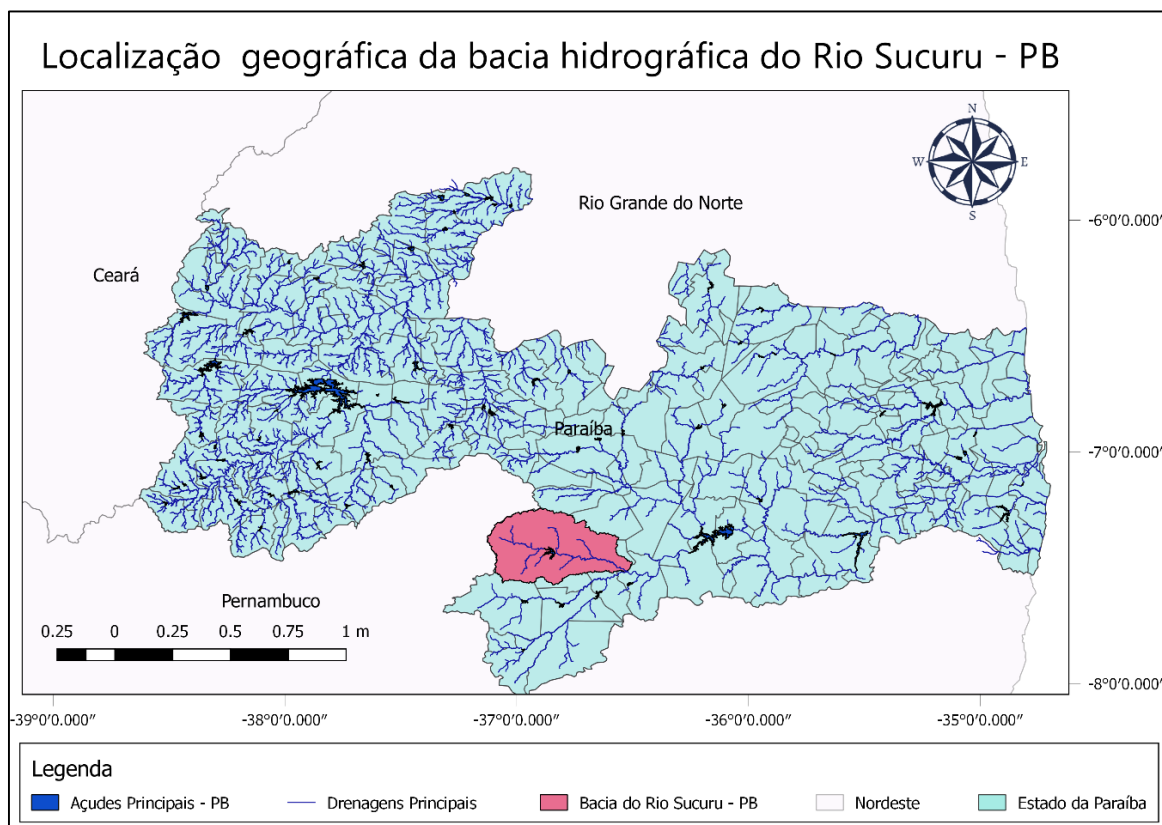
#### 3.2 Localização e descrição da área de estudo

A área de abrangência desta pesquisa corresponde à bacia do Rio Sucuru (Figura 5), com área territorial de 1.682,87 Km<sup>2</sup>, sendo o principal rio o Sucuru, localizada na mesorregião da Borborema e microrregião do Cariri da Paraíba, entre as coordenadas geográficas 7°28'30'' e 7°49'30'' Sul e 36°34'00'' e 37°12'00'' Oeste. Os limites dessa bacia hidrográfica são a Leste e ao Sul, com a bacia do Rio Paraíba em seu alto curso; a Oeste, com a bacia do Rio Pajeú, integrada ao Estado de Pernambuco; e ao Norte, com a bacia do Rio Taperoá.

À nascente do rio principal dessa bacia de regime intermitente, o Sucuru, situa-se na serra dos Cariris Velhos, com altimetria aproximada de 591 metros, nos limites da Paraíba com Pernambuco, drenando águas de forma parcial e total da área de oito municípios paraibanos, sendo eles: Amparo, Congo, Coxixola, Monteiro, Ouro Velho, Prata, Serra Branca e Sumé. Esses municípios contribuem com um fluxo hídrico para o açude público de Sumé e o Epitácio Pessoa, que é conhecido popularmente como “Boqueirão” (SILVA, 2017).

Assim, essa bacia apresenta grande relevância pois contribui para o abastecimento de forma direta ou indireta para esses oito municípios paraibanos, apresentando grande importância para esta região que se situa no semiárido paraibano (Cariri) que faz parte da diagonal mais seca do Brasil, onde é comum, em determinados setores, o registro pluviométrico médio anual inferior a 300 mm. (NASCIMENTO; ALVES, 2008)

Figura 5: Localização geográfica da bacia hidrográfica do Rio Sucuru - PB



### 3.3 Aquisição de dados

Foram utilizadas imagens selecionadas dos sensores Enhanced Thematic Mapper, plus - ETM+, satélite Landsat 7, e Operational Land Imager - OLI, satélite Landsat 8, e índices de vegetação EVI, NDVI e SAVI disponibilizadas pelo Serviço Geológico Americano - U.S. Geological Survey – USGS, de órbita 215 e ponto 65, correspondendo ao período de outubro de 2014 a setembro de 2016. A obtenção do índice STI foi realizado através do processamento das bandas Swir-I e Swir-II realizado no software R (R Development Core Team, 2018). As etapas metodológicas envolvidas podem ser observadas na Figura 6.

As visitas de campo foram realizadas no início do período chuvoso, entre os dias 28/09 a 30/09, 06 e 07/10 de 2016, perfazendo um total de cinco dias de atividades (SILVA, 2017). A coleta dos dados correspondeu ao período seco, estação em que a vegetação de Caatinga está desprotegida de suas folhagens em decorrência das condições naturais, tornando-se uniforme em termos de comparação.

Foram selecionados 65 pontos (Figura 7), em que 41 pontos foram obtidos em campo (SILVA, 2017), no qual foram utilizado 7 pontos como base para gerar o modelo

de classificação, e 24 pontos obtidos através do Google Earth 2018 (GOOGLE, 2018). Obedecendo a classificação proposta por (CHAVES, et al., 2008) e adaptada por (SILVA, 2017).

Figura 6: Fluxograma simplificado das etapas metodológicas da classificação do uso do solo

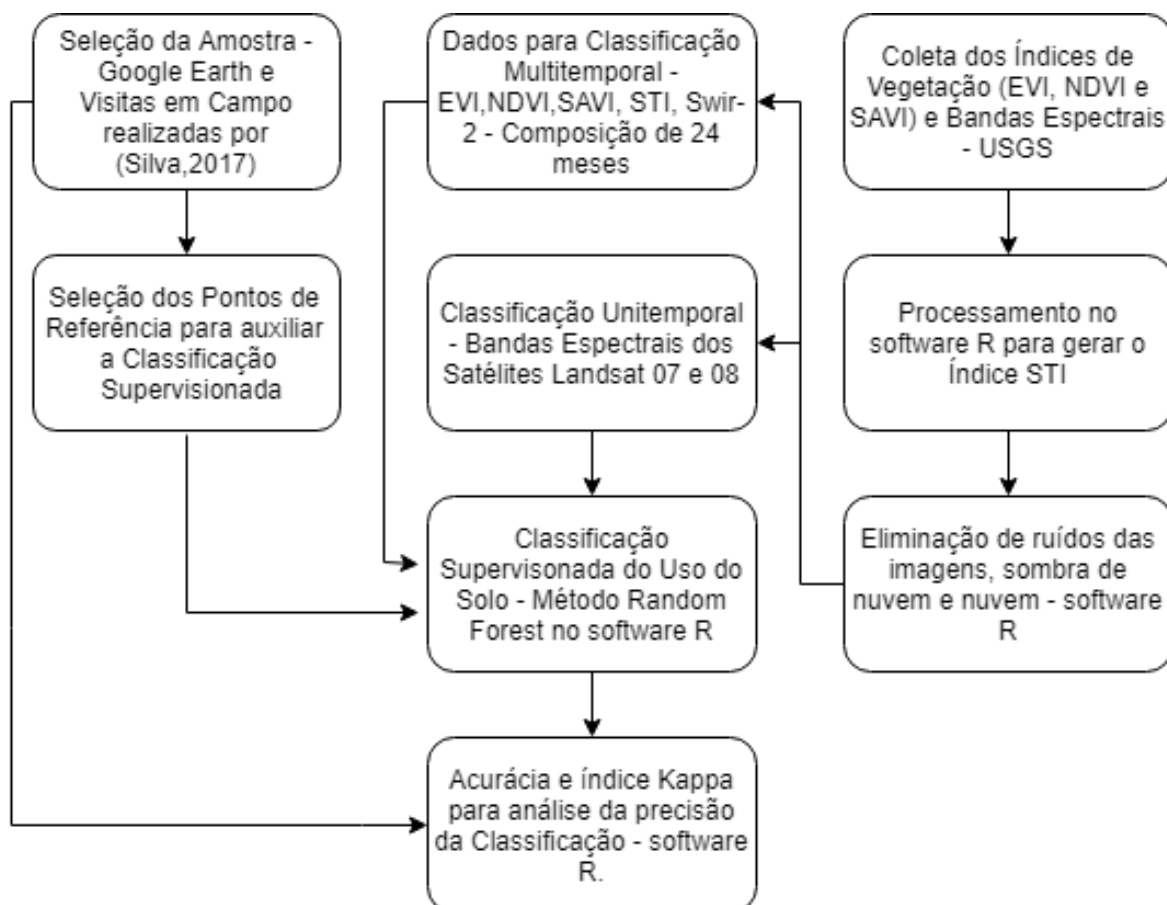
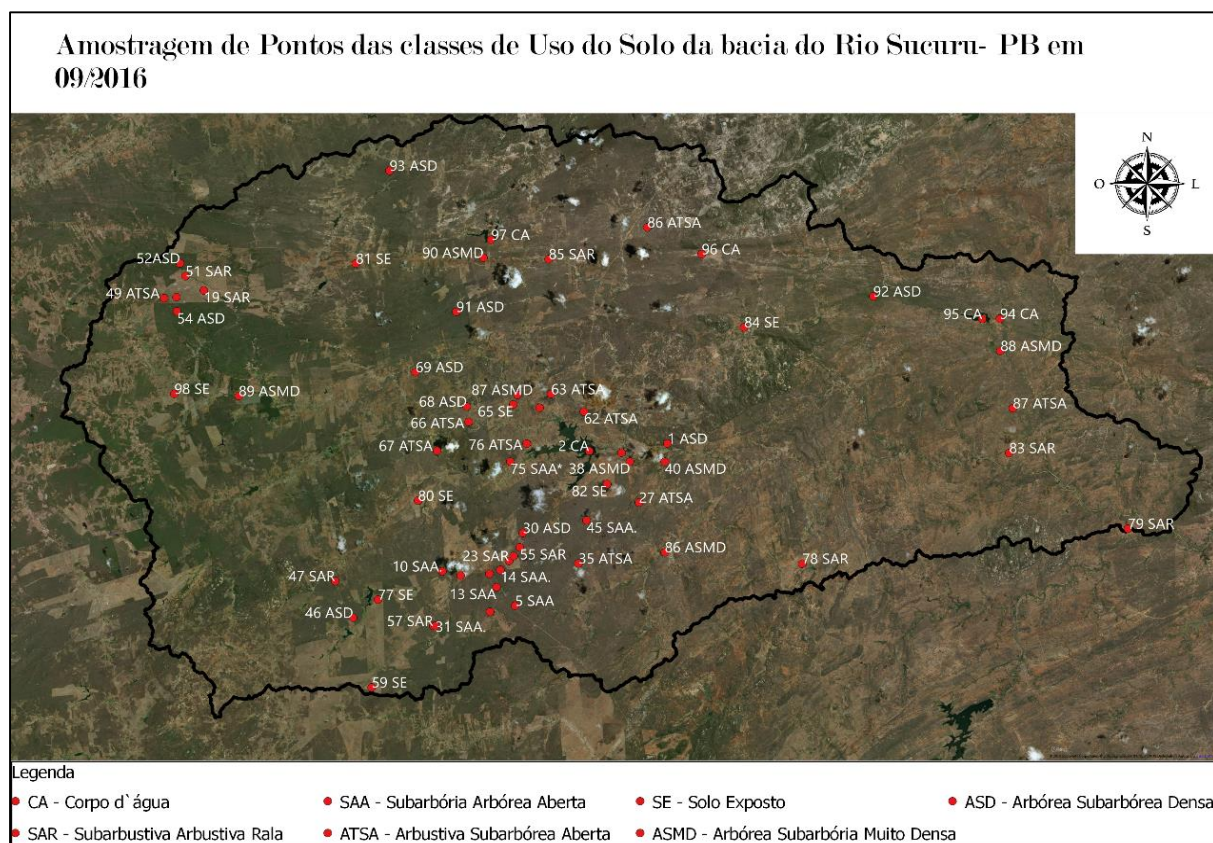


Figura 7: Amostragem dos pontos das classes do uso do solo da bacia do Rio Sucuru

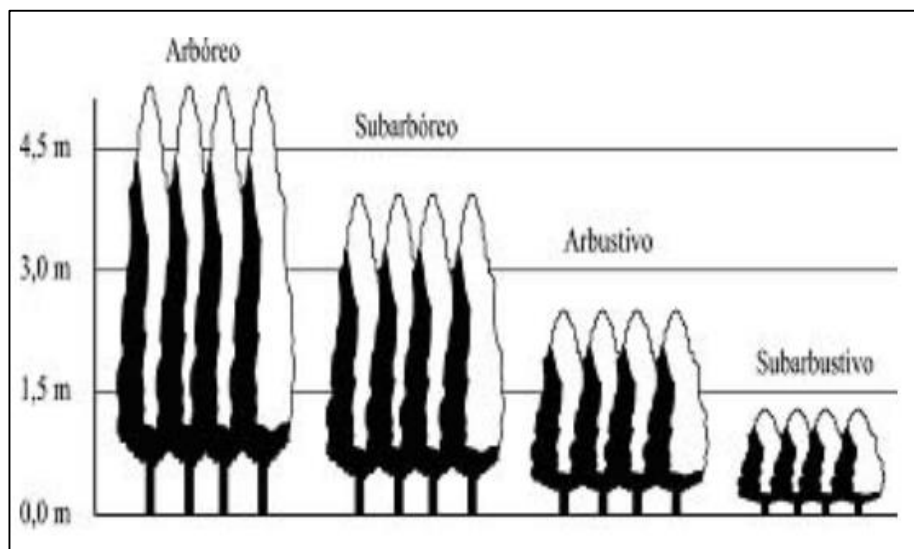


A classificação se compõe de quatro níveis categóricos: Tipo, Grupo, Classe e Sub-Classe. O “Tipo”– separa a vegetação natural da vegetação cultivada. Compreende os cultivos anuais, semi-perenes e perene; as pastagens e forrageiras; os bosques florestais e a vegetação de Caatinga, com suas espécies nativas. Incluem-se, nesta categoria, as superfícies compreendidas por estradas, edificações, afloramento de rocha, mineração e corpos d’água. (CHAVES, et al., 2008).

Assim como as demais categorias, o “grupo” refere-se à vegetação da Caatinga, a qual separa as comunidades vegetais que têm, em comum, a dominância de um dado porte (Figura 8). São assim identificados os grupos: arbóreo (> 4,5 m de altura); subarbóreo (> 3 m e < 4,5 m); arbustivo (> 1,5 m e < 3 m) e subarbustivo (< 1,5 m). (CHAVES, et al., 2008).



Figura 8: Classificação da vegetação quanto ao porte

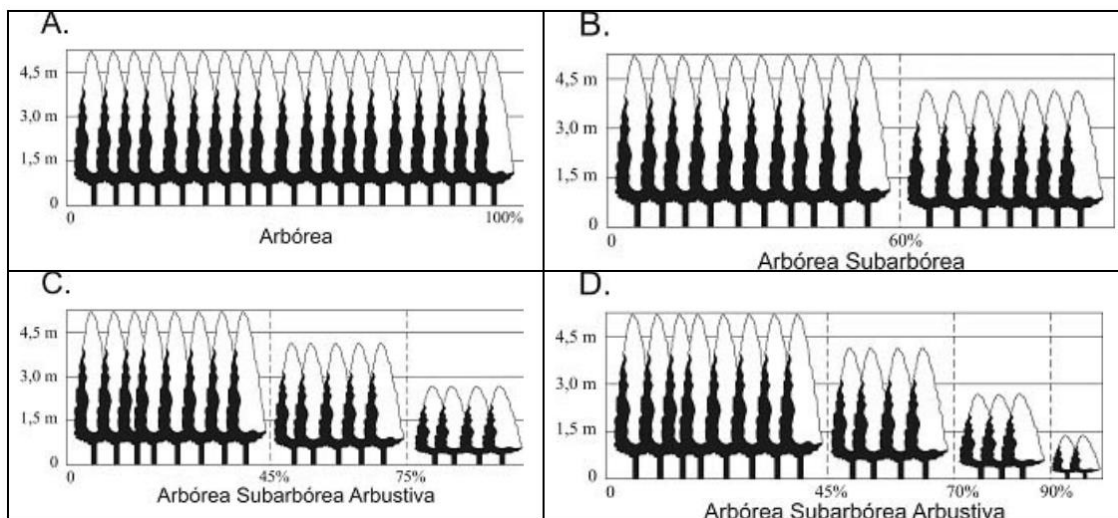


Fonte: (CHAVES, et al., 2008)

Dentro de cada grupo as “classes” separam as comunidades vegetais de diferentes feições morfológicas, sendo descritas pela ordem decrescente dos portes dominantes, quando identificáveis, até o terceiro estrato, Figura 9 (CHAVES *et al.*, 2008).

Para comunidades com porte homogêneo, o nome da classe se confunde com o do próprio grupo; assim, a classe arbórea pertence ao grupo arbóreo, Figura 9A. Cada grupo apresenta dez classes, descritas de acordo com a sequência dos três estratos vegetais predominantes; surgem assim, nomes compostos formados pela descrição desses estratos, exemplo: classe arbórea subarbórea, Figura 9B, classe arbórea subarbórea arbustiva, Figura 9C. A inclusão dos segundo e terceiro estratos é feita quando a sua área percentual de ocorrência ultrapassa 15% do conjunto da vegetação da área observada (Figura 9D)(CHAVES *et al.*, 2008).

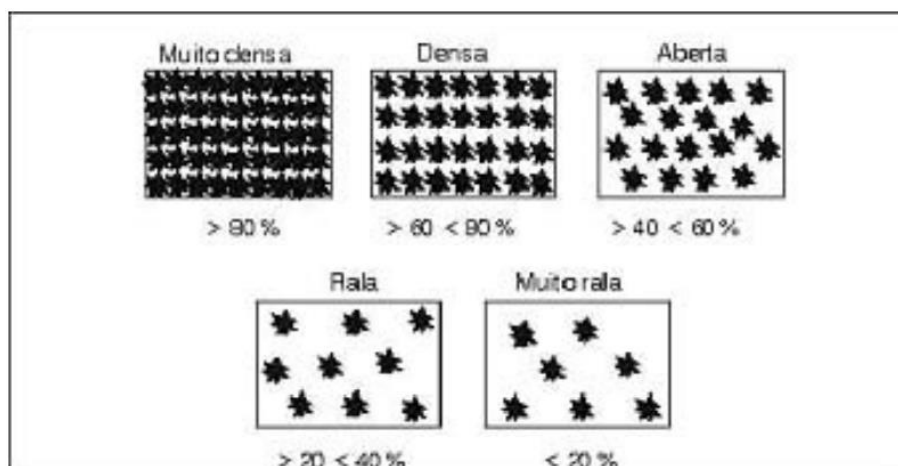
Figura 9: Classificação das plantas quanto à classe.



Fonte: (CHAVES, et al., 2008)

A “subclasse” representam a cobertura vegetal, em termos de percentagem de recobrimento da superfície do solo, sendo avaliada pelos seguintes graus de recobrimento: muito densa (> 80% de cobertura); densa (> 60 e < 80%); aberta (> 40 e < 60%); rala (> 20 e < 40%) e muito rala (< 20%), como se apresenta na Figura 10 (CHAVES *et al.*, 2008).

Figura 10: Classificação de comunidades vegetais quanto ao grau de recobrimento da superfície do solo



Fonte: (CHAVES, et al., 2008)

Foi utilizada a classificação da vegetação adotada por (SILVA, 2017) e adaptada para obtenção das seguintes classes: Arbórea subarbórea muito densa, Arbórea subarbórea densa, Subarbórea arbustiva aberta, Arbustiva subarbórea aberta, Subarbustiva arbustiva rala, Solo exposto, Corpos d'água.

### 3.3 Processamento das Imagens de Satélite Landsat

O processamento digital das imagens ocorreu através do software R (R Development Core Team, 2018), com a eliminação de ruídos das imagens, sombra de nuvem e nuvem.

Os dados supervisionados utilizados para gerar a classificação das diferentes concepções foram os 7 pontos (Tabela 2) que apresentaram melhor qualidade acerca do tipo de cobertura do solo. Foi realizado a verificação do comportamento desses pontos ao longo de uma série temporal para os índices EVI, NDVI, SAVI, STI e banda espectral Swir-II e realizado um análise unitemporal para o período de Março e Setembro de 2016.

É utilizando a pacote *random forest* (LIAW; WIENER, 2002) implementado no R, foram realizadas a classificação multitemporal (período de Novembro de 2014 à Setembro de 2016) para os índices EVI, NDVI, SAVI, STI e banda espectral Swir-II. A classificação unitemporal foi realizada no mês de setembro de 2016, utilizando as bandas espectrais (azul, verde, vermelho, NIR, Swir-I e Swir-2).

Tabela 2: Pontos base para processamento dos dados para obtenção da classificação da vegetação da Bacia do Sucuru.

Ponto	Classificação	Longitude (x)	Latitude (y)
2	Corpo D'água	-36,903470	-7,672449
59	Solo Exposto	-37,025029	-7,806065
53	Subarbustiva Arbustiva Rala	-37,134091	-7,585550
76	Arbustiva Subarbórea Aberta	-36,938389	-7,668392
13	Subarbórea Arbórea Aberta	-36,955257	-7,749482
54	Arbórea Subarbórea Densa	-37,133759	-7,593610
38	Arbórea Subarbórea Muito Densa	-36,880703	-7,678481

Fonte: (SILVA, 2017)

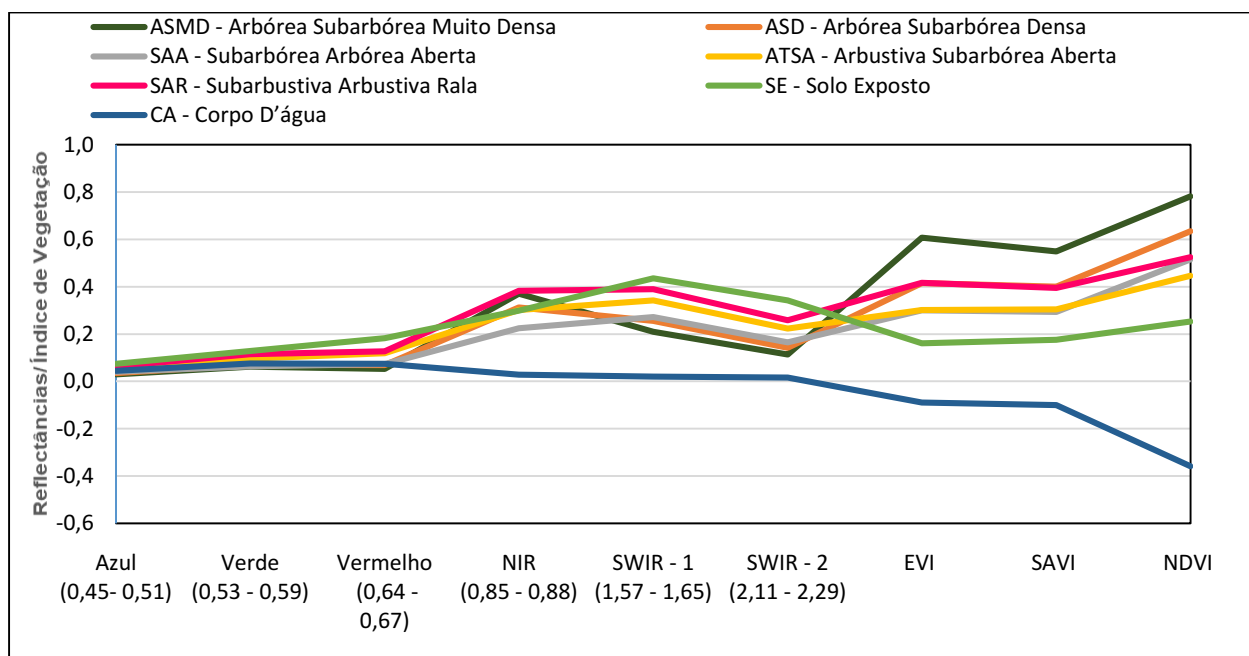
### 3.5 Análise da Acurácia

A matriz de confusão foi gerada por meio do pacote *caret* implementado no R a partir dos 65 pontos de referência. A acurácia geral, do usuário, do produtor e o índice kappa foram calculados através do pacote *caret* (KUHN *et al.*, 2018).

## 4. RESULTADOS E DISCUSSÕES

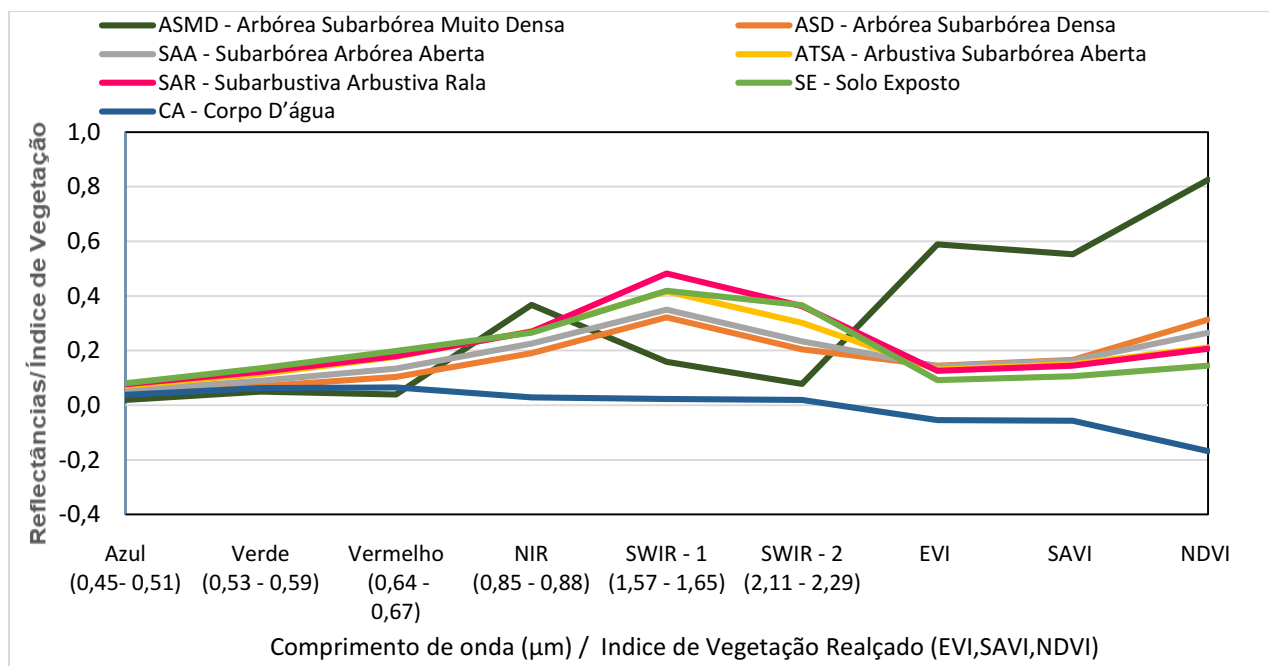
Para as duas situações unitemporais (março e setembro de 2016) analisadas por meio dos sete pontos referentes as classes é possível observar a dificuldade de distinguir determinadas classes de vegetação, pois as mesmas apresentam respostas espectrais semelhantes (Figura 11 e 12). Na Figura 11 são apresentadas a resposta espectral e o valor dos índices de vegetação EVI, NDVI e SAVI para os sete tipos de cobertura do solo investigados para estação de chuva.

Figura 11: Reflectância e índices de vegetação EVI, NDVI e SAVI os tipos de cobertura do solo para uma imagem no período chuvoso (Março de 2016).



Observa-se a similaridade das respostas para os tipos de cobertura do solo SAR e ASD, também para ATSA e SAA. Na Figura 12 são apresentados os valores das respostas espectrais e índice de vegetação para os sete tipos de cobertura do solo para a estação seca. Observa-se que houve comportamento semelhante ao identificado para a estação chuvosa, existindo tipos de cobertura do solo com respostas semelhantes, a saber: os tipos de cobertura do solo ASD, SAA e ainda SAR, ATSA e SE.

Figura 12: Reflectância e índices de vegetação EVI, NDVI e SAVI os tipos de cobertura do solo para uma imagem no período de seca (setembro de 2016).



A análise do comportamento unitemporal dos sete tipos de cobertura do solo tanto no período de chuva ou estiagem indica que é difícil identificar diferenças das bandas espectrais (com ênfase para as bandas azul, verde e vermelho) ou dos índices de vegetação (EVI, SAVI e NDVI) entre alguns tipos de cobertura monitorados. No período de chuvas, a região é tomada por uma vegetação rasteira que surge apenas neste período, o que impossibilita a diferenciação entre os padrões vegetados, especialmente os que foram classificados como “abertas”. Para época de estiagem, há perda de folhagem, o que diminui a distinção dos padrões. Os dois momentos representam duas situações em que classificações de cobertura do solo poderiam distinguir padrões, sejam por métodos de classificação supervisionada ou não supervisionada em processamento digital de imagem de satélite. Entretanto, em ambas, haverá limites para detectar tipos de cobertura do solo distintos que apresentam resposta semelhantes em algum período do ano.

Todavia, avaliando o comportamento dos alvos ao longo de uma série temporal de 24 meses no período de outubro de 2014 à setembro de 2016, utilizando Índices de Vegetação (Figura 13) ou bandas espectrais (Figura 14) é possível distinguir melhor as classes de cobertura do solo.

Figura 13: Valor de NDVI (série temporal) para as classes de cobertura do solo

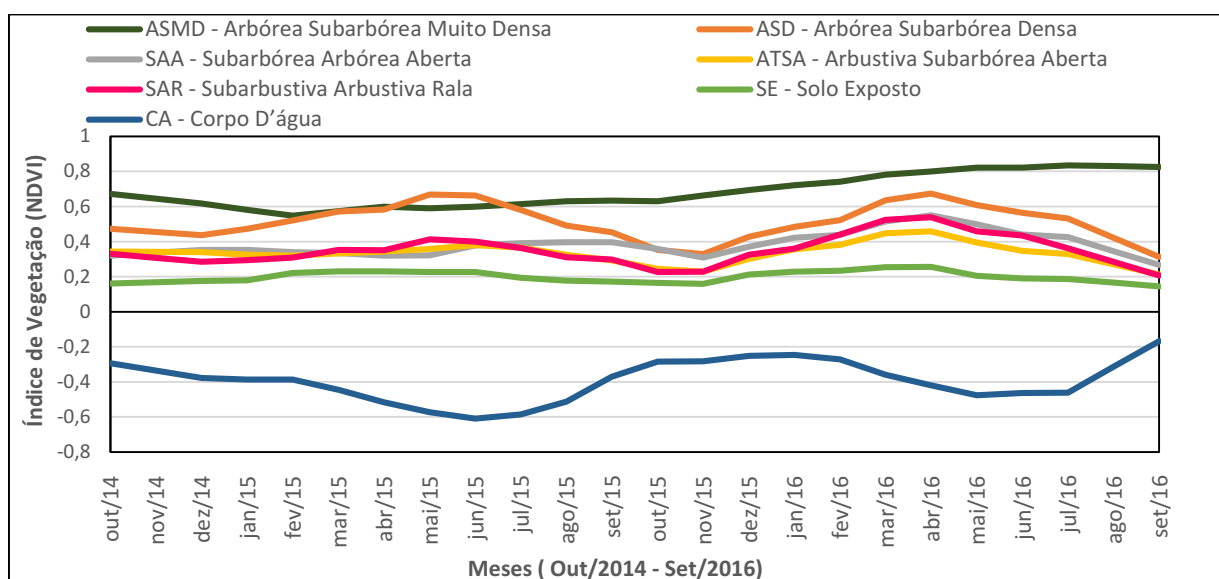
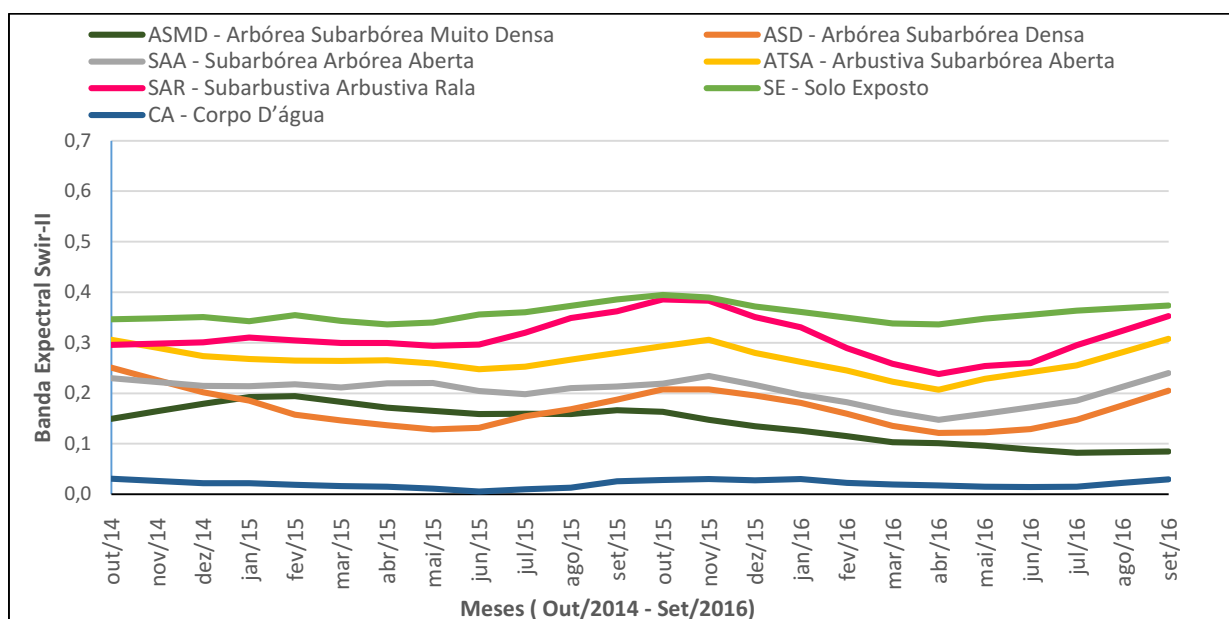


Figura 14: Valor de Swir-2(série temporal) para as classes de cobertura do solo



Por meio da análise multitemporal é possível identificar um padrão de comportamento para cada tipo de cobertura do solo, obtendo uma assinatura multitemporal para as classes de cobertura, foi observado esse padrão para os índices de vegetação EVI, SAVI, STI que se encontram em anexos. A distinção das assinaturas temporais pode ser explicada pelas características morfológicas de cada classe, principalmente as relacionadas à vegetação, já que a Caatinga é caracterizada pela grande perda de biomassa na estação seca e, por isso, torna-se importante avaliar a sazonalidade

do comportamento das classes. Portanto, pode-se distingui-las de maneira mais eficaz avaliando suas características ao longo dos ciclos de chuvas e secas aos quais o ecossistema é submetido.

Assim, realizando a classificação multitemporal da cobertura do solo da bacia, o índice de vegetação EVI (Figura 15) apresentou pior desempenho na classe SAR com nível de acurácia do usuário de 40,0%, entretanto para as classes ATSA, ASMD, CA, SE apresentou acurácia do usuário acima de 80%. De modo geral o índice apresentou acurácia geral de 66,2 % e índice Kappa de 0,601. (Tabela 3).

Figura 15: Classificação do uso do Solo - Série temporal 24 meses EVI

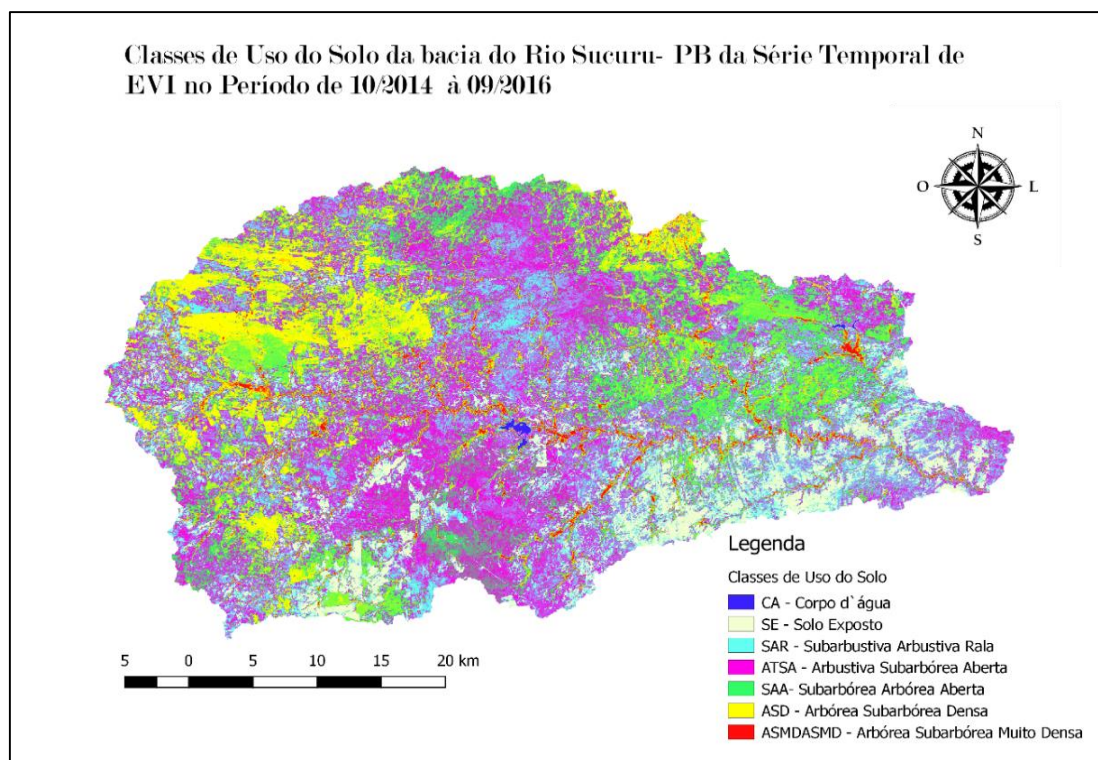


Tabela 3: Acurácia do índice de vegetação EVI

Classes de Cobertura	Classes Observadas							Acurácia do Usuário
	CA	SE	SAR	ATSA	SAA	ASD	ASMD	
CA	4			1				80,0%
SE		8	2					80,0%
SAR		4	4	2				40,0%
ATSA			1	9		1		81,8%
SAA				5	5	0		50,0%
ASD				1	4	5		50,0%
ASMD					0	1	8	88,9%
Acurácia do Produtor	100,0%	66,7%	57,1%	50,0%	55,6%	71,4%	100,0%	65

Já o índice de vegetação NDVI (Figura 16) a classe que apresentou pior desempenho foi a SAR com 40,0% de precisão do usuário, entretanto, as demais classes apresentam acurácia do usuário acima de 80%. O índice apresentou acurácia geral 78,4% e índice Kappa 0,746. (Tabela 4).

Figura 16: Classificação do uso do Solo - Série temporal 24 meses NDVI

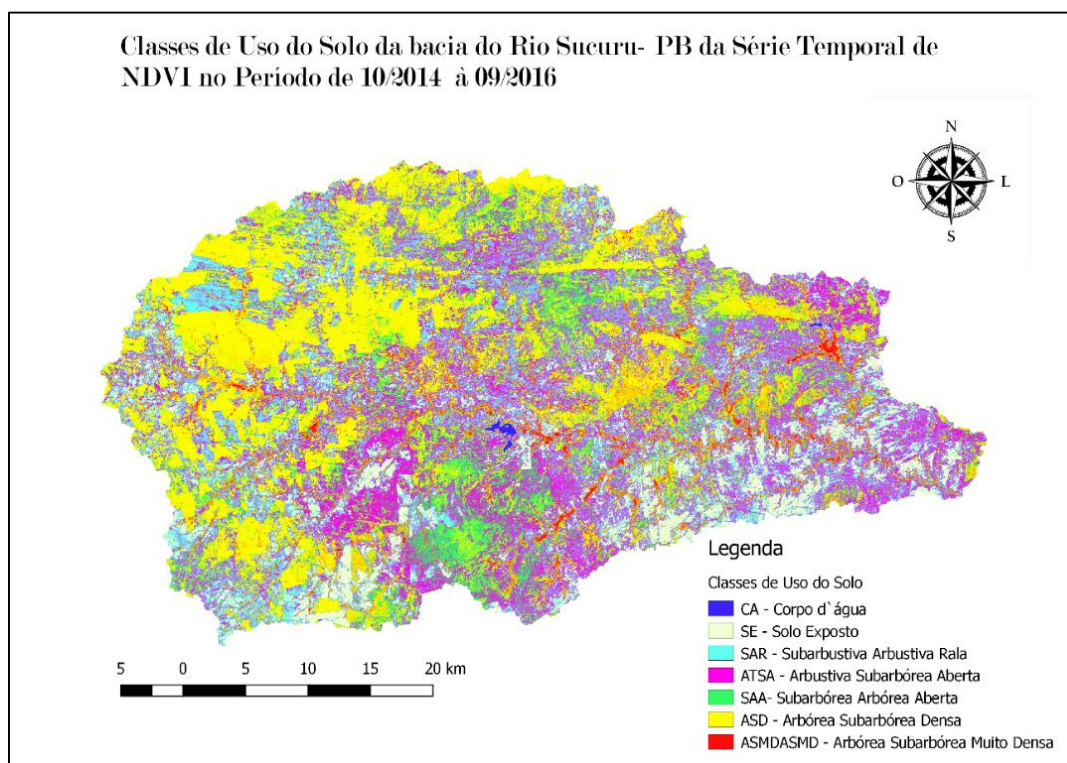


Tabela 4: Acurácia do índice de vegetação NDVI

Classes de Cobertura	Classes Observadas							Acurácia do Usuário
	CA	SE	SAR	ATSA	SAA	ASD	ASMD	
CA	4				1			80,0%
SE		8						80,0%
SAR		6	4					40,0%
ATSA				10		1		90,9%
SAA				2	8			80,0%
ASD					2	8		80,0%
ASMD							9	100,0%
Acurácia do Produtor	100,0%	57,1%	66,7%	83,3%	72,7%	88,9%	100,0%	65



Para o índice de vegetação SAVI (Figura 17) de acordo com a Tabela 5 pode-se observar que o índice que apresentou menor acurácia do usuário foi o SAR, e apenas os índices CA, SE e ASMD apresentaram acurácia do usuário acima de 80 %. Esse índice apresentou acurácia global de 69,2 % e índice Kappa 0,638.

Figura 17: Classificação do uso do Solo - Série temporal 24 meses SAVI

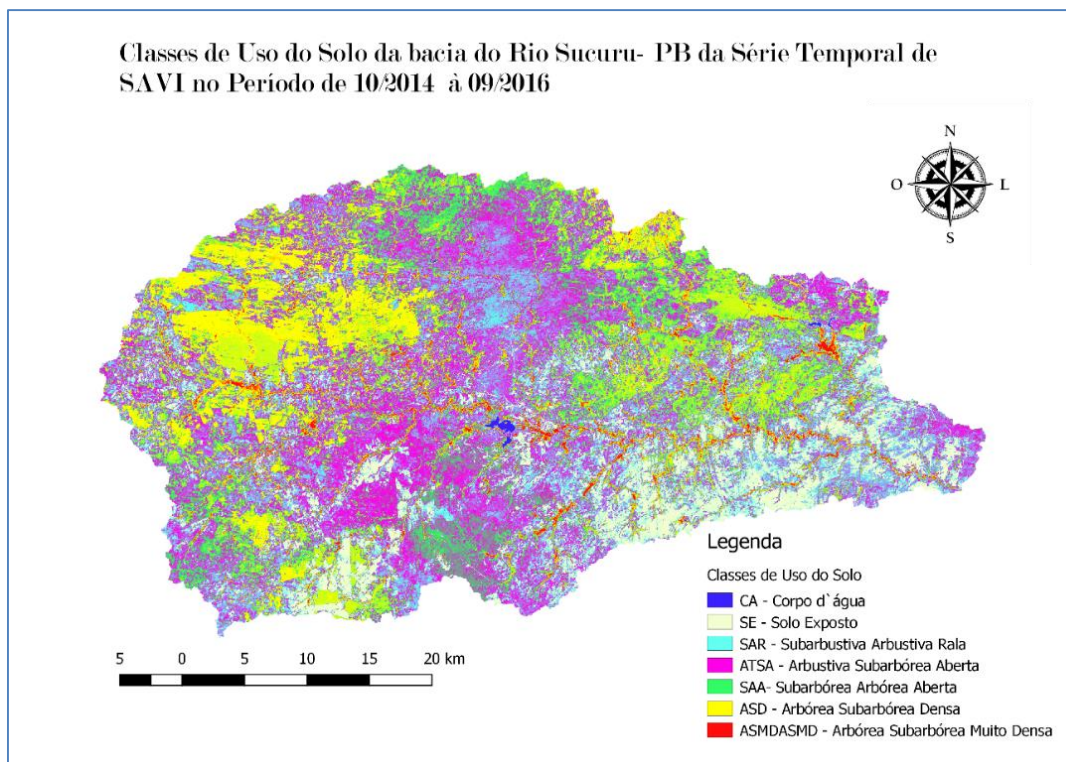


Tabela 5: Acurácia do índice de vegetação SAVI

Classes de Cobertura	Classes Observadas							Acurácia do Usuário
	CA	SE	SAR	ATSA	SAA	ASD	ASMD	
CA	4			1				80,0%
SE		8	2					80,0%
SAR		4	4	2				40,0%
ATSA			2	8		1		72,7%
SAA				4	6			60,0%
ASD				1	2	7		70,0%
ASMD						1	8	88,9%
Acurácia do Produtor	100,0%	66,7%	50,0%	50,0%	75,0%	77,8%	100,0%	65

Utilizando o índice STI (Figura 18) pode-se observar as classes SAR, ATSA apresentaram acurácia do usuário menores que 40 % e apenas a classes ASM apresentou acurácia do usuário de 90,0 % (Tabela 6). Sua acurácia global foi de 61,5 % e índice kappa de 0,549.

Figura 18: Classificação do uso do Solo - Série temporal 24 meses STI

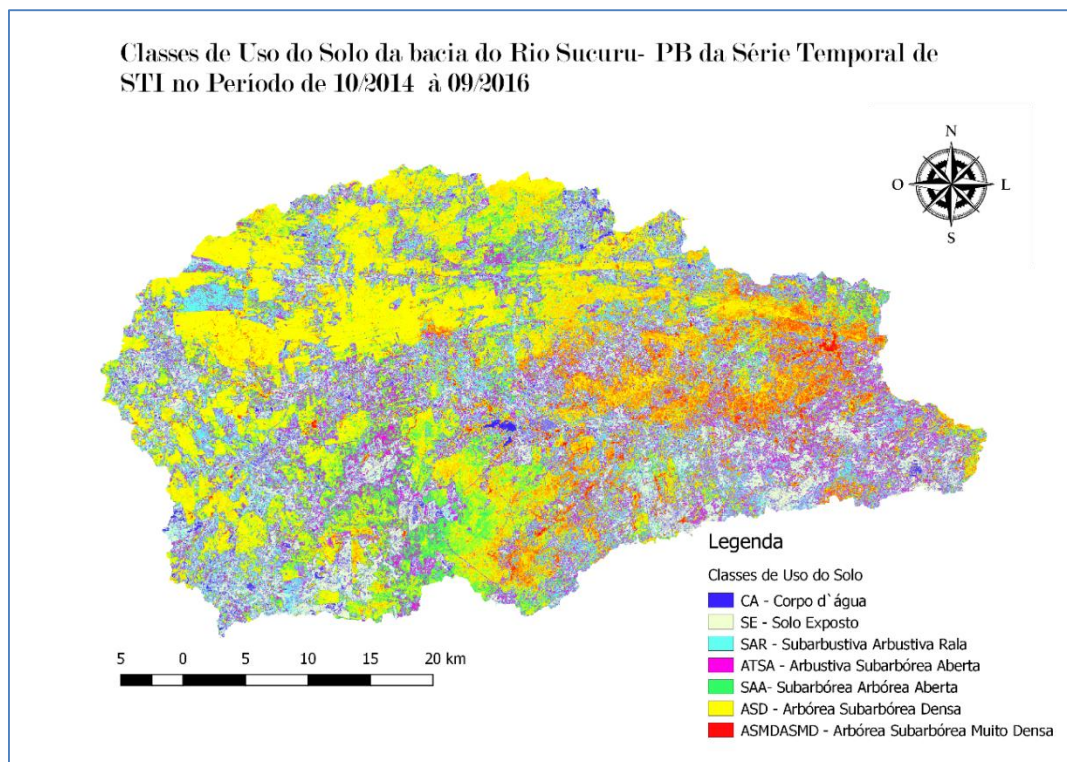


Tabela 6: Acurácia do índice de vegetação STI

Classes de Cobertura	Classes Observadas							Acurácia do Usuário
	CA	SE	SAR	ATSA	SAA	ASD	ASMD	
CA	3	1				1		60,0%
SE	1	7	2					70,0%
SAR	2	5	3					30,0%
ATSA		1	2	4	2	2		36,4%
SAA					7	3		70,0%
ASD				1		9		90,0%
ASMD						2	7	77,8%
Acurácia do Produtor	50,0%	50,0%	42,9%	80,0%	77,8%	52,9%	100,0%	65

Realizando a classificação multitemporal para a banda Swir-II (Figura 19) pode-se verificar que as classes que apresentaram pior desempenho foram SAR e ATSA, respectivamente com acurácia do usuário de 50% e 30%. A classe que apresentou melhor

desempenho foi a ASMD com 88,9%, a banda apresentou acurácia global de 64,6% e índice kappa de 0,584.

Figura 19: Classificação do uso do Solo - Série temporal 24 meses Swir-II

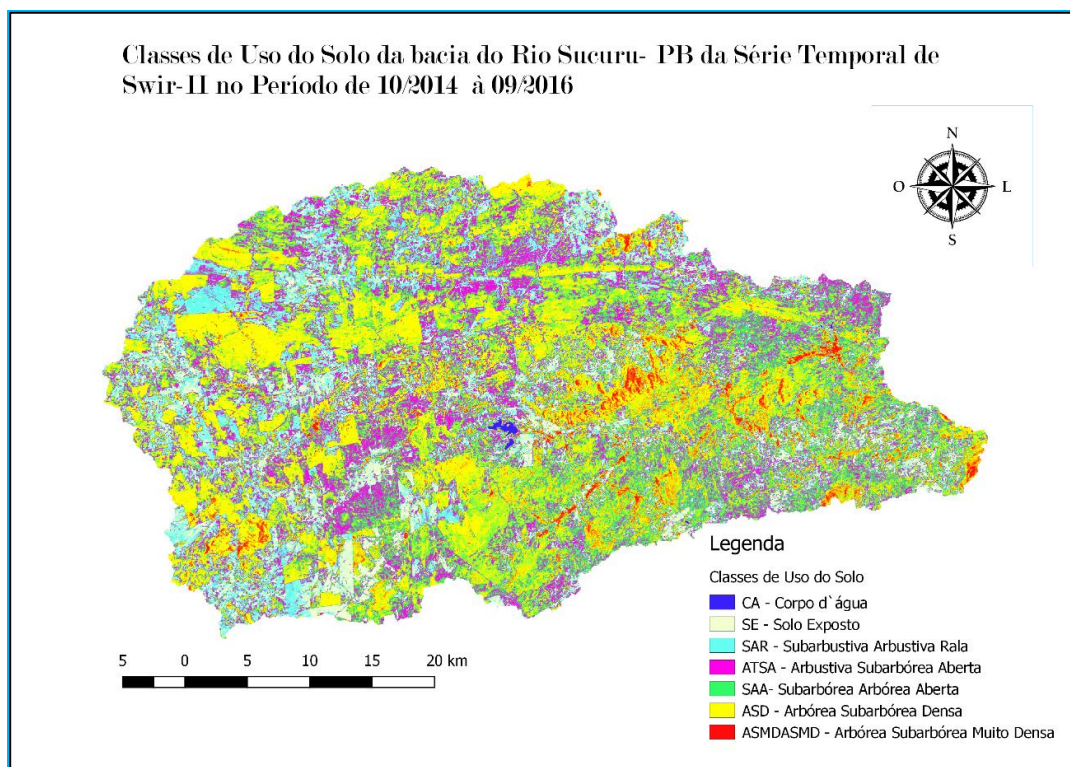


Tabela 7: Acurácia do índice de vegetação Swir-II

Classes de Cobertura	Classes Observadas							Acurácia do Usuário
	CA	SE	SAR	ATSA	SAA	ASD	ASMD	
CA	3		1			1		60,0%
SE		8	2					80,0%
SAR		5	5					50,0%
ATSA			1	4	5	1		36,4%
SAA				2	6	2		60,0%
ASD						8	2	80,0%
ASMD						1	8	88,9%
Acurácia do Produtor	100,0%	61,5%	55,6%	66,7%	54,5%	61,5%	80,0%	65

Para realizar a classificação unitemporal (Figura 20) no mês de setembro de 2016, foram utilizadas as bandas espectrais (azul, verde, vermelho, NIR, Swir-I e Swir-2). Foi observado que a classe que apresentou pior desempenho foi ATSA com acurácia do usuário de 36,4 %, única classe que apresentou acurácia do usuário acima de 80% foi a

ASD com 90,0 % de precisão. Sua acurácia geral foi de 64,6% e índice kappa de 0,584 (Tabela 8).

Figura 20: Classificação Unitemporal do uso do Solo em Setembro de 2016

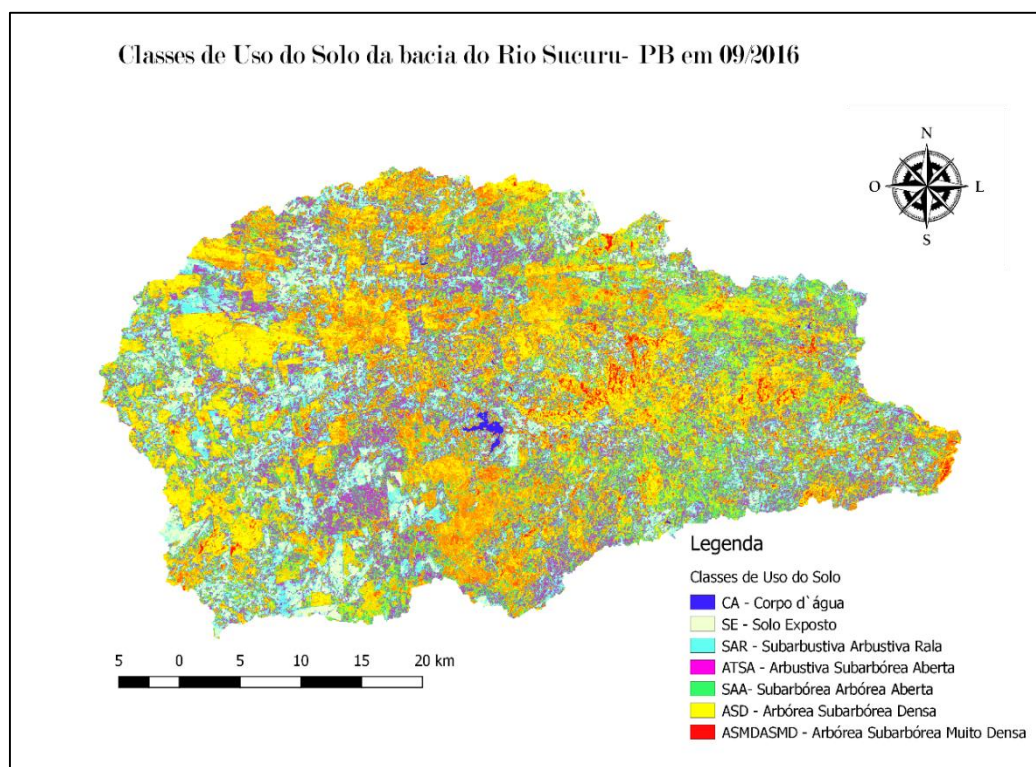


Tabela 8: Acurácia da classificação unitemporal – Bandas espectrais

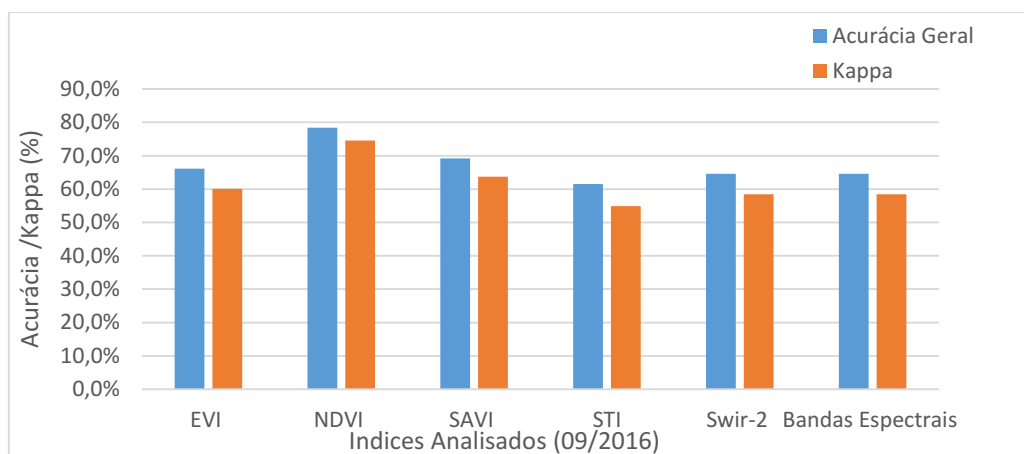
Classes de Cobertura	Classes Observadas							Acurácia do Usuário
	CA	SE	SAR	ATSA	SAA	ASD	ASMD	
CA	3					2	0	60,0%
SE		7	3					70,0%
SAR		3	7					70,0%
ATSA				4	5	2		36,4%
SAA					6	3	1	60,0%
ASD						9	1	90,0%
ASMD						3	6	66,7%
Acurácia do Produtor	100,0%	70,0%	70,0%	100,0%	54,5%	47,4%	75,0%	65

Avaliando o comportamento das classes em relação a classificação multitemporal com os índices de vegetação e a banda Swir-II pode-se observar que a classe que apresentou menor acurácia do usuário foi a SAR - Subarbustiva Arbustiva Rala, em que muitos de seus pontos se confundiam com a classe SE – Solo exposto. Isso pode ser

explicado pela dificuldade de distinção que ocorre entre estas classes, já que as mesmas se assemelham em diferentes pontos do ano. Devido ao fato da vegetação da caatinga apresentar variabilidade em função dos períodos de chuva e estação de estiagem. Este fato não ocorreu para a análise unitemporal, em que a classe SAR apresentou 70% de precisão do usuário, e pode-se avaliar tal precisão devido ao período de obtenção da amostra ser no período de setembro de 2016, o que facilita a distinção entre as classes, já que a classificação unitemporal foi referente a este mês.

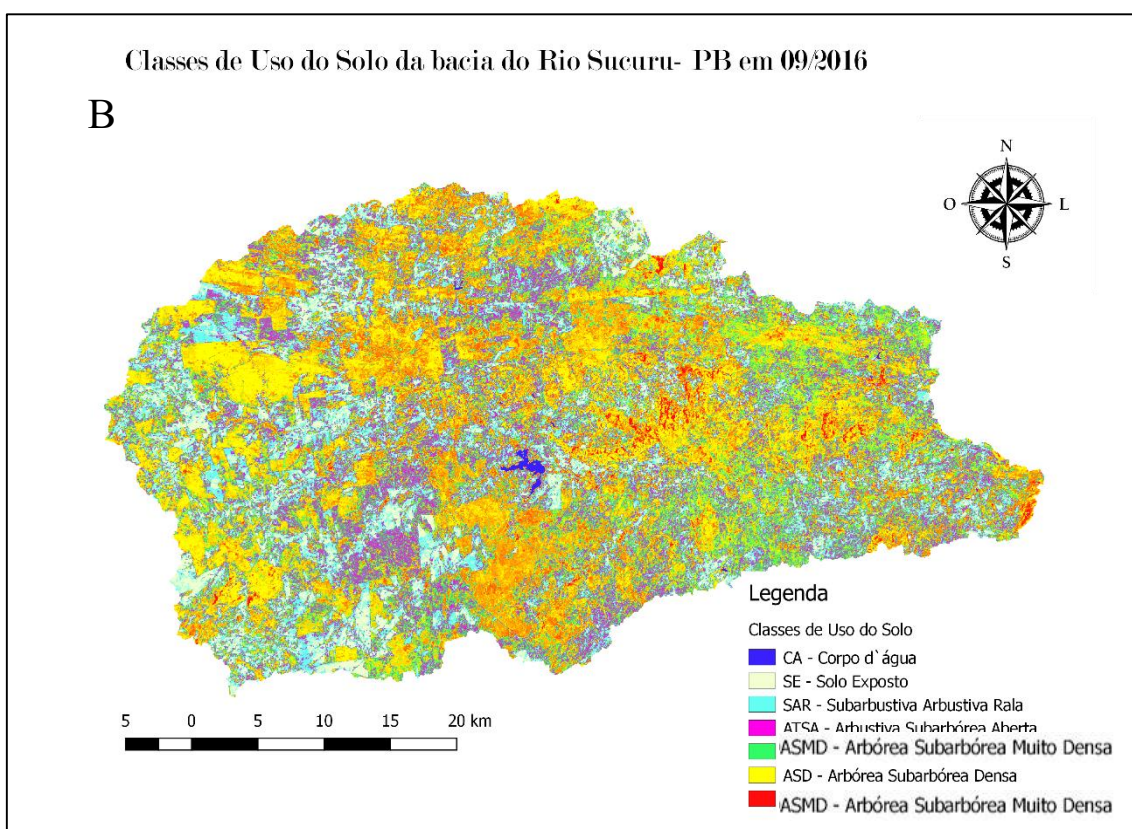
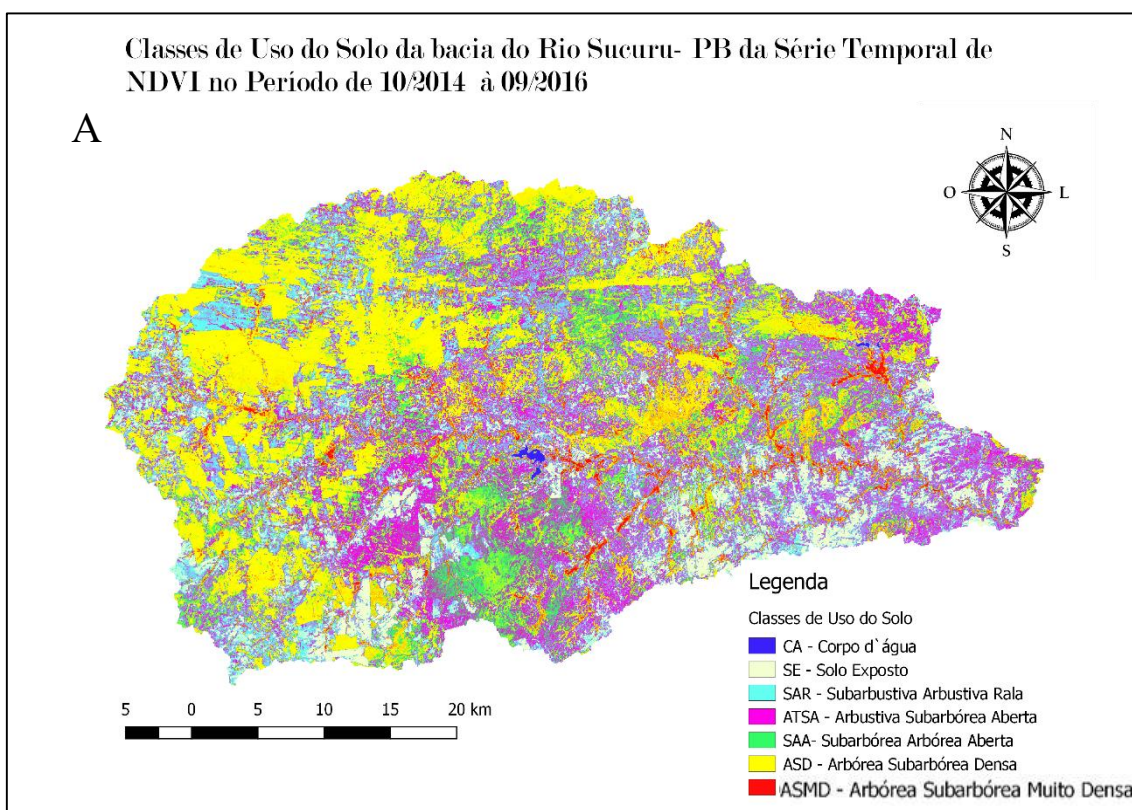
De forma geral os índices de vegetação em que foram realizados a análise multitemporal apresentam melhor desempenho que a análise unitemporal como mostra a Figura 19. O índice de vegetação que apresentou melhor acurácia geral e índice kappa foi o NDVI, com 78,4 %, (0,746 respectivamente). Apresentando desempenho melhor que a classificação unitemporal com acurácia geral de 64,6% e índice kappa 0,584.

Figura 159: Resumo de acurácia geral e índice kappa



Pode-se observar a distinção entre as classes de cobertura na análise multitemporal (NDVI) e análise unitemporal (Bandas espectrais), ao espacializarmos as classes apresentado na Figura 21.

Figura 21: Classificação do uso do Solo: A- Série temporal 24 meses NDVI, B - Unitemporal com as bandas espectrais (09/2016)



Realizando os quantitativos das áreas de cobertura (Tabela 9) pode-se notar que a classe que apresentou maior diferença percentual entre a análise multitemporal e unitemporal foi a ASMD - Arbórea Subarbórea Muito Densa com crescimento de 241,03% em relação a análise multitemporal, e a classe que apresentou maior decréscimo foi a ATSA - Arbustiva subarbórea Aberta com 44,32 %. Diferenças acentuadas podem interferir nos resultados de simulações realizadas com modelos hidrológicos ou hidrosedimentológicos que necessariamente simulam cenários de uso e cobertura do solo, tomando como referência cenários reais, implicando em aproximações equivocadas da realidade. Neste sentido, considera-se muito relevante a avaliação do espectro temporal da cobertura do uso do solo para que haja uma maior aproximação da realidade simulada.

Tabela 9: Áreas das classes de uso do solo - Análise Multitemporal x Análise Unitemporal

Classes de Uso do Solo	Áreas das Classes de Uso do Solo				Diferença (%)
	Análise Multitemporal (NDVI)		Análise Unitemporal (Bandas Espectrais)		
	(Km <sup>2</sup> )	(%)	(Km <sup>2</sup> )	(%)	
CA - Corpo d'água	2,34	0,14%	3,32	0,19%	41,94%
SE - Solo Exposto	151,24	8,85%	127,96	7,49%	-15,39%
SAR - Subarbustiva Arbórea rala	295,95	17,33%	193,49	11,33%	-34,62%
ATSA - Arbustiva subarbórea Aberta	436,02	25,53%	242,79	14,21%	-44,32%
SAA - Subarbórea Arbórea Aberta	256,70	15,03%	391,42	22,92%	52,48%
ASD - Arbórea Subarbórea Densa	536,01	31,38%	647,42	37,90%	20,78%
ASMD - Arbórea Subarbórea Muito Densa	29,82	1,75%	101,69	5,95%	241,03%
<b>Área Total</b>	<b>1708,08</b>	<b>100%</b>	<b>1708,08</b>	<b>100,00%</b>	<b>-</b>

## 5. CONCLUSÃO

A caatinga apresenta grande variação de biomassa em um curto período de tempo, fato esse que está relacionado ao regime de precipitação ao qual esta é submetida. Tal característica deve ser levada em consideração, surgindo então a necessidade de avaliar o comportamento dos tipos de cobertura do solo ao longo de uma série temporal.

Assim, a análise multitemporal dos tipos de cobertura do solo, por meio de índices de vegetação, como NDVI, se mostrou mais eficaz que o estudo unitemporal, possibilitando a identificação de um padrão para cada classe de vegetação – análise esta, mais complexa que a avaliação de um único instante (regime de chuvas ou secas).

Portanto, para se obter uma classificação da cobertura do solo mais próxima da realidade, através do processamento de imagens de satélites Landsat, é necessário ponderar e avaliar o comportamento das classes de cobertura do solo ao longo de um regime de chuvas e secas, melhorando assim a qualidade e a precisão dos resultados obtidos.

Nessa perspectiva, a utilização de técnicas de *machine learning*, que utilizam dados supervisionados ou não supervisionados, surge como alternativa para reconhecimento das assinaturas multitemporais dos tipos de cobertura do solo ao longo de uma série temporal. A técnica de *Random Forest* (RF) se mostrar como um caminho para classificação de séries multitemporais, podendo, assim, auxiliar a construção de modelos hidrológicos para região da caatinga, com a elaboração de classificação do uso do solo desse ecossistema.



## **6. RECOMENDAÇÕES**

A caracterização da cobertura do solo pode auxiliar a construção de modelos de cobertura do solo da Caatinga, auxiliando a construção de modelos hidrológicos mais fidedignos, contribuindo para gestão de bacias hidrográficas e identificação das mudanças da cobertura do solo podem causar na bacia hidrográfica.

A simulação de cenários, a partir de análises de cobertura do solo, de escoamento superficial pode-se criar estratégias viáveis de exploração e manejo de bacias hidrográficas entendendo seu comportamento, e verificando a influência da cobertura do solo nas bacias hidrográficas. Podendo, assim, avaliar os processos de erosão intensos, as inundações, os assoreamentos desenfreados de reservatórios e cursos d'água causados decorrentes da cobertura do solo.

## REFERÊNCIAS

- ANTUNES, A. F. B., e C. LINGNAU. “Uso dos índices de acurácia para avaliação de mapas temáticos obtidos por meio de classificação digital.” **CONGRESSO E FEIRA PARA USUÁRIO DE GEOPROCESSAMENTO**, 1997.
- ASSIS, J.M.O., Calado, L.O., Souza, W.M., Sobral, M.C. “O MUNICÍPIO DE BELÉM DE SÃO FRANCISCO – PE NOS ANOS DE 1985 E 2010.” **Revista Brasileira de Geografia Física**, 2014.
- BELGIU, Mariana, e Lucian DRAGURT. “Random forest in remote sensing: A review of applications and future directions.” **ISPRS Journal of Photogrammetry and Remote Sensing**, 20 de Fevereiro de 2016: 24-31.
- BORGES, Ck., R.M.de MEDEIROS, Rep RIBEIRO, e Santos. “Study of biophysical parameters using remote sensing techniques to Quixeré-CE region.” **Journal of Hyperspectral Remote Sensing** 6, 2016.
- BREIMAN, L. “Random forest.” **Mach.Learn.** 45 (2001).
- BRITO, A. “Acurácia e precisão do mapeamento da flora nativa e dos reflorestamentos de Minas Gerais.” Lavras: Dissertação de Mestrado - Universidade Federal de Lavras, 2009 .
- BELGIU, M.; DRĂGU, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 114, p. 24–31, 2016.
- CHAVES, I. DE B. *et al.* Uma Classificação Morfo-Estrutural Para Descrição E a Morpho-Structural Classification for Description and Evaluation of the Biomass of the Caatinga Vegetation. *Revista Brasileira de Geografia Física*, v. 06, n. 2, p. 225–228, 2008. Disponível em:  
<<http://www.bibliotecaflorestal.ufv.br:80/handle/123456789/9039>>.
- CONGALTON, R.A. “A review of assessing the accuracy of classification remotely sensed data.” **Remote Sensing of Environment**, Jan de 1991: p. 35-36.
- CORCORAN, Jennifer M, Joseph F Knight, e Alisa L. Gallant. “Influence of Multi-Source and Multi-Temporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota.” **Remote Sensing**, 5 de 2013: 3212-3238.
- CUNHA, J. E. de B. L., I. A. A. RUFINO, C. de O. GALVÃO, T. E. PEREIRA, F. V. BRASILEIRO, e E. V. PEREIRA. “Análise e processamento automático de grandes volumes de dados ambientais (Big Earth Observation Data Sets).” **Anais... XVIII Simpósio Brasileiro de Sensoriamento Remoto (SBSR)**, 2017: p. 7459-7466.
- DEMARCHI, J.C, E.L PIROLI, e C, R, L ZIMBACK. “Temporal analysis of the land use and comparison between NDVI and SAVI indexes in the city of Santa Cruz do Rio Pardo - SP using Landsat-5 image.” **RAOEGA**, 2011: p.234-271.

- EMBRAPA. **Landsat**. 2008. <https://www.cnpm.embrapa.br/projetos/cdbrazil/txt/landsat.htm> (acesso em 20 de Fevereiro de 2018).
- FITZ, P. R. **Geoprocessamento sem complicação**. São Paulo: Oficina de textos, 2008.
- FLORENZANO, G,F. **Iniciação em Sensoriamento Remoto** . São Paulo : Oficina de Textos, 2007.
- GISLASON, P,O.,BENEDIKTSSON,J,A, e J.R. SVEINSSON. “Random forests for land cover classification.” **Pattern Recogn**, 2006: 294-300.
- GOOGLE. **Google Earth website**. 2018. <http://earth.google.com/>,.
- HANSEN, J., SATO, M., & Ruedy, R. (2012). “Perception of climate change.” **Proceedings of the National Academy of Sciences of the United States of America**, 2012, 109 ed.: E2515-E2423.
- HUETE, A, et al. “Overview of the radiometric and biophysical performance of the MODIS vegetation indices.” **Remote Sensing Of Environment**, Nov de 2002, 1-2 ed.: 195-213.
- HUETE, A,R. “" A Soil-adjusted vegetation index (SAVI).” "**Remote sensing of environment**", 25 de 3 de 1988: 295-309.
- HUETE, A. R.; HUIQING LIU; VAN LEEUWEN, W. J. D. The use of vegetation indices in forested regions: issues of linearity and saturation. *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development*, v. 4, n. 1, p. 1966–1968, 1997. Disponível em: <<http://ieeexplore.ieee.org/document/609169/>>.
- JACQUES, D. C. *et al.* Monitoring dry vegetation masses in semi-arid areas with MODIS SWIR bands. **Remote Sensing of Environment**, v. 153, p. 40–49, 2014.
- JENSEN, J. R. **Sensoriamento Remoto do Ambiente: uma perspectiva em Recursos Terrestres**. Tradução: Tradução de J. C. N. Epiphanyo. São José dos Campos, SP: Ed. Parênteses, 2011.
- KUHN, M. *et al.* **Package “caret”** Title Classification and Regression Training Description Misc functions for training and plotting classification and regression models. p. 2, 2018. Disponível em: <<https://cran.r-project.org/web/packages/caret/caret.pdf>>.
- LEITE, A P, R SANTOS G, e J ,E,O SANTOS. “ANÁLISE TEMPORAL DOS ÍNDICES DE VEGETAÇÃO NDVI E SAVI NA ESTAÇÃO EXPERIMENTAL DE ITATINGA UTILIZANDO IMAGENS.” **Revista Brasileira de Energias Renováveis**, 2017, 4 ed.: 606-623.
- LIAW, A.; WIENER, M. Classification and Regression by randomForest. **R news**, v. 2, n. December, p. 18–22, 2002.
- LIU, K,Shi,W, e H. Zhang. “A fuzzy topology-based maximum likelihood classification.” **J.Photogramm Remote Sensing**, 2011: 103-104.

LOBÃO, J.S.B. et al. “Uso de imagem de satélite para a discriminação de alvos urbanos e não-urbanos no município de Feira de Santana.” **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO** 11 (2003): 1867-1874.

LOYD, CHARLIE. **Landsat 8 Bands**. 14 de Jun de 2013. <https://landsat.gsfc.nasa.gov/landsat-8/landsat-8-bands/> (acesso em 20 de Janeiro de 2018).

MENESES, P,R, e T. ALMEIDA. **Introdução ao processamento de imagens sensoriamento remoto**. BRASÍLIA: Cnpq, 2012.

MILLER, H. M., RICHARDSON, LESLIE, KOONTZ, S. R., LOOMIS, JOHN, & KOONTZ, LYNNE. “Users, uses, and value of Landsat satellite imagery—Results from the 2012 survey of.” **U.S. Geological Survey Open-File Report**, 2013: 51.

Ministério do Meio Ambiente. **Ministério do Meio Ambiente**. 2017. <http://www.mma.gov.br/biomas/caatinga> (acesso em 10 de Janeiro de 2018).

MIURA, T., HUETE, A. R., VAN LEEUWEN, W. J. D., & DIDAN, K. “Vegetation detection through smoke-filled AVIRIS images: an assessment using MODIS band passe.” **Journal of Geophysical Research**, 1998: 32001-32011.

NASCIMENTO, S. S.; ALVES, J. J. . ECOCLIMATOLOGIA DO CARIRI PARAIBANO. *Rev. Geogr. Acadêmica*, v. 2, n. December, p. 28–41, 2008. Disponível em: <[www.rga.ggf.br](http://www.rga.ggf.br)>.

OLIVEIRA, Michele Laurentino De. “**Estimativa Da Evapotranspiração Utilizando Sebal Automatizado Em Clima Mediterrâneo**.” Dissertação de Mestrado Programa de Engenharia Civil - UFCG. Campina Grande, 2018.

R Development Core Team. **R: A Language and Environment for Statistical Computing**. 2018. <https://www.r-project.org/> (acesso em 20 de Fevereiro de 2018).

RODAL, M. J. N.; COSTA, K. C. C.; SILVA, A. C. B. L. E. Estrutura da vegetação caducifólia espinhosa (caatinga) de uma área do sertão central de Pernambuco. *Hoehnea*, v. 35, n. 2, p. 209–217, 2008. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2236-89062008000200004&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2236-89062008000200004&lng=pt&tlng=pt)>.

ROY, D.p. et al. “Landsat-8: Science and product vision for terrestrial global change research.” **Remote Sensing Of Environment**, abr de 2014: 154-172.

SALLES, A,T, M,F VALÉRIO, e T, G FLORENZANO. “MONITORAMENTO DA COBERTURA VEGETAL E DO USO DO SOLO DA RESERVA INDÍGENA CAARAPÓ - MS.” **MULTITEMAS**, Novembro de 1998.

SILVA, JOÃO NAILSON DE CASTRO. “**Análise temporal de índices de vegetação no apoio à classificação de imagens: Cobertura do solo na bacia hidrográfica do rio Sucuro**.” 107f. (Dissertação de Mestrado) , Universidade Federal de Campina Grande, Campina Grande-PB, 2017.

SILVA, M. DE S. *et al.* Avaliação da cobertura do solo como indicador de gestão de recursos hídricos: um caso de estudo na sub-bacia do Córrego dos Bois, Minas Gerais.

*Engenharia Sanitaria e Ambiental*, v. 22, n. 3, p. 445–452, 2017. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-41522017000300445&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-41522017000300445&lng=pt&tlng=pt)>.

SPIEGEL, M. R. **Estatística**. São Paulo: Makron Books, 1993.

SUN, Liya, e Karsten SCHULZ. “The Improvement of Land Cover Classification by Thermal Remote Sensing.” **Remote sensing**, 26 de Junho de 2015: 8368-8390.

TEODORO, Valter Luiz Lost, Denilson TEIXEIRA, e Daniel Jadyr Leite COSTA. “O Conceito de Bacia Hidrográfica e a Importância da Caracterização Morfométrica Para o Entendimento da Dinâmica Ambiental Local.” Revista **UNIARA** N.20, 2007.

US Geological Survey. USGS. 25 de Dezembro de 2018. <https://landsat.usgs.gov>.

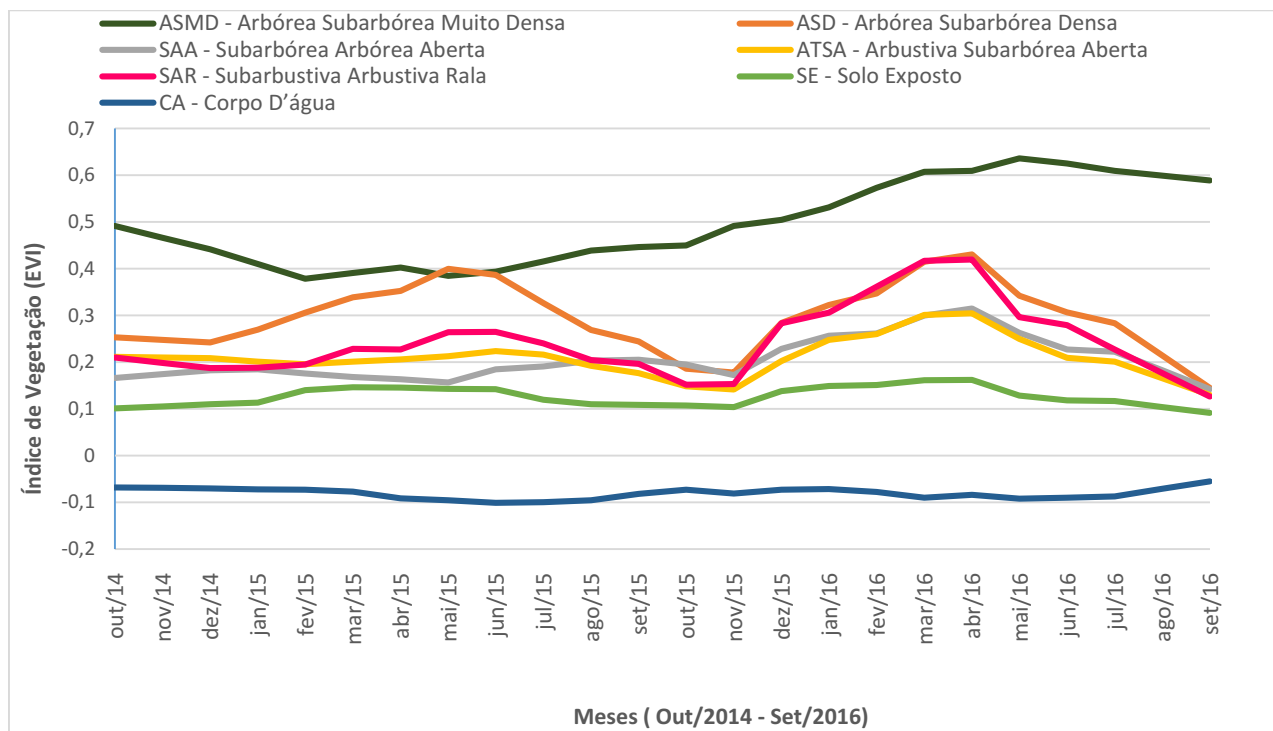
VALBUENA, R, M MALTAMO, e P. PACKALEN. “Classification of forest development stages from national low-density lidar datasets: a comparison of machine learning methods.” **REVISTA DE TELEDETECCIÓN** - Spanish Association of Remote Sensing, 2016: 15-25.

VAN DEVENTER, A., WARD, A., GOWDA, P., & LYON, J. (1997). “Using thematic mapper data to identify contrasting soil plains and tillage practices.” **Photogrammetric Engineering & Remote Sensing**, January de 1997: 87-93.

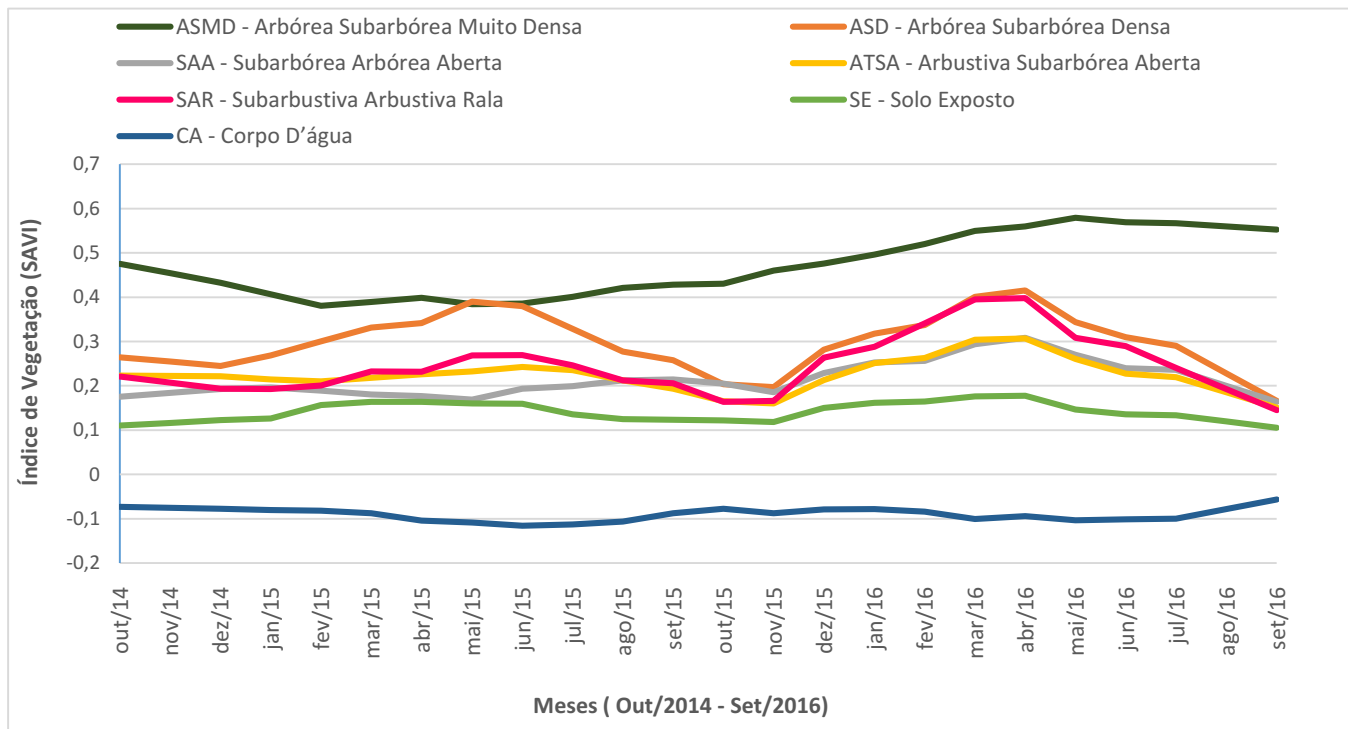
WATZLAWICK, L.F., F.F. KIRCHNER, C.R. SANQUETTA, e M.V. SCHUMACHER. “Fixação de carbono em floresta ombrófila mista em diferentes estágios de regeneração.” *As Floresta e o Carbono.*, 2002: p. 418-422.

## APÊNDICES

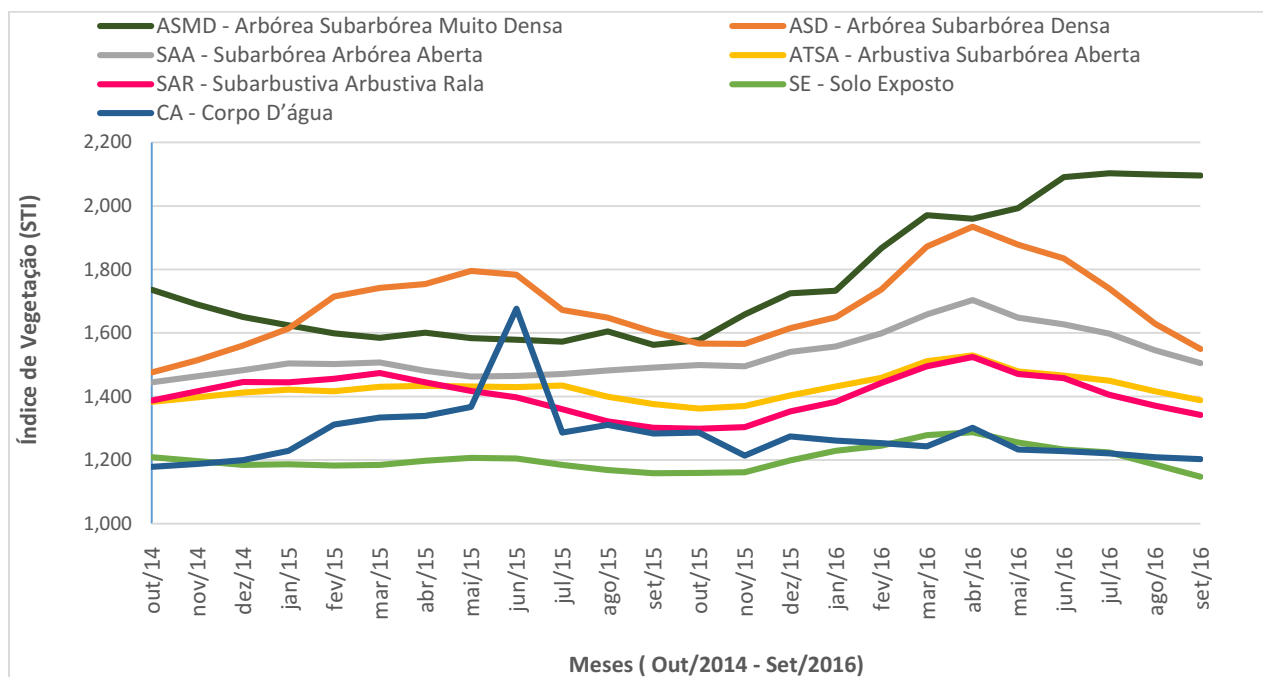
Valor de EVI para as classes de cobertura do solo



Valor de SAVI para as classes de cobertura do solo



## Valor de STI para as classes de cobertura do solo



## Pontos Utilizados

Classes	Pontos	Longitude (x)	Latitude (y)	Classes	Pontos	Longitude (x)	Latitude (y)
CA - Corpo D'água	2	-36,903470	-7,672449	SAA - Subarbórea Arbórea Aberta	13	-36,955257	-7,749482
	94	-36,674007	-7,597753		5	-36,944905	-7,759952
	95	-36,683846	-7,597871		10	-36,985557	-7,740523
	96	-36,840690	-7,561248		14	-36,953141	-7,739744
	97	-36,958301	-7,553139		21	-36,911594	-7,786399
SE - Solo Exposto	59	-37,025029	-7,806065		31	-36,989389	-7,771298
	65	-36,945951	-7,646226		33	-36,959258	-7,741898
	60	-36,975267	-7,742947		45	-36,904988	-7,711779
	77	-37,021485	-7,756457		61	-36,958929	-7,763349
	80	-36,999109	-7,700562		75	-36,947547	-7,678581
	81	-37,033935	-7,566597		54	-37,133759	-7,593610
	82	-36,893456	-7,691217		1	-36,859800	-7,668297
	84	-36,816951	-7,602589		30	-36,940591	-7,718864
	98	-37,135343	-7,640224	46	-37,035430	-7,766737	
	23	-36,948260	-7,734651	52	-37,132102	-7,566426	
SAR - Subarbusativa Arbustiva Rala	53	-37,134091	-7,585550	ASD - Arbórea Subarbórea Densa	68	-36,971917	-7,647373
	47	-37,045426	-7,746080		69	-37,000527	-7,627991
	83	-36,668892	-7,673643		91	-36,977760	-7,593883
	79	-36,602558	-7,716384		92	-36,744813	-7,585110
	78	-36,784551	-7,736231		93	-37,015094	-7,514116
	19	-37,118819	-7,581788		38	-36,880703	-7,678481
	57	-36,990610	-7,771365		39	-36,885406	-7,673497
	55	-36,942250	-7,726967		40	-36,861216	-7,678559
ATSA - Arbustiva Subarbórea Aberta	85	-36,926146	-7,564059	ASMD - Arbórea Subarbórea Muito Densa	85	-36,931200	-7,648128
	51	-37,129083	-7,573729		86	-36,861297	-7,729774
	67	-36,988542	-7,672419		87	-36,943610	-7,640820
	76	-36,938389	-7,668392		88	-36,673797	-7,615991
	35	-36,909461	-7,736079		89	-37,099330	-7,641433
	49	-37,141000	-7,586134		90	-36,962555	-7,563310
	62	-36,906473	-7,650227				
	56	-36,945610	-7,731857				
	63	-36,924831	-7,640370				
	66	-36,970835	-7,656102				
	27	-36,875677	-7,701662				
	86	-36,871073	-7,546137				
87	-36,666855	-7,648311					
				<b>Legenda</b>			
				Pontos Coletados do Google Earth			
				Pontos Coletados em Campo	Refêrencia		
					Demais Pontos		

