



**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOLOGIA E GEOQUÍMICA**

TESE DE DOUTORADO Nº 156

**TRÊS DÉCADAS DE MUDANÇAS NA PLANÍCIE COSTEIRA
BRASILEIRA: O STATUS DOS MANGUEZAIS, DA
AQUICULTURA E SALICULTURA A PARTIR DE SÉRIES
TEMPORAIS LANDSAT E TÉCNICAS DE APRENDIZADO
DE MÁQUINA**

Tese apresentada por:

CESAR GUERREIRO DINIZ

Orientador: Prof. Dr. Pedro Walfir Martins e Souza Filho (UFPA)

**BELÉM - PARÁ
2021**

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

D585t Diniz, Cesar Guerreiro.

Três décadas de mudanças na planície costeira brasileira: O status dos manguezais, da aquicultura e salicultura a partir de séries temporais Landsat e técnicas de aprendizado de máquina / Cesar Guerreiro Diniz. — 2021.

xvi, 92 f. : il. color.

Orientador(a): Prof. Dr. Pedro Walfir Martins e Souza Filho

Tese (Doutorado) - Universidade Federal do Pará, Instituto de Geociências, Programa de Pós-Graduação em Geologia e Geoquímica, Belém, 2021.

1. Aprendizado de Máquina. 2. Mudança de Uso e Cobertura da Terra. 3. Zona Costeira. 4. Manguezais. 5. Aquicultura e Salicultura. I. Título.

CDD 621.3678



Universidade Federal do Pará
Instituto de Geociências
Programa de Pós-Graduação em Geologia e Geoquímica

**TRÊS DÉCADAS DE MUDANÇAS NA PLANÍCIE COSTEIRA
BRASILEIRA: O STATUS DOS MANGUEZAIS, DA
AQUICULTURA E SALICULTURA A PARTIR DE SÉRIES
TEMPORAIS LANDSAT E TÉCNICAS DE APRENDIZADO
DE MÁQUINA**

Tese apresentada por

CESAR GUERREIRO DINIZ

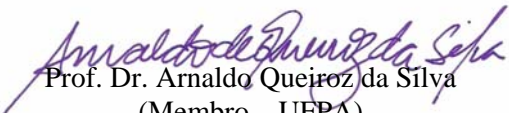
**Como requisito parcial à obtenção de Grau de Doutor em Ciências na Área de GEOLOGIA e
Linha de Pesquisa GEOLOGIA MARINHA E COSTEIRA.**


Data de Aprovação: 31 / 03 / 2020


Banca Examinadora:


Prof. Dr. Pedro Walfir H. de Souza Filho
(Orientador – UFPA)


Prof. Dr. David Lagomasino
(Membro – Universidade da Carolina do
Leste - ECU)


Prof. Dr. Arnaldo Queiroz da Silva
(Membro – UFPA)


Prof. Dr. Carlos Souza Junior
(Membro – Imazon)


Prof. Dr. Nils Asp Neto
(Membro – UFPA)

“Você provavelmente não mudará o mundo
com seu Mestrado ou Doutorado.
Mas, uma vez que os tenha começado,
termine-os.”
(Cesar Diniz, 2016)

"You probably won't change the World
with your Master's or Doctorate thesis.
But once you've started,
finish them."
(Cesar Diniz, 2016)

AGRADECIMENTOS

A ciência é um eterno exercício de cooperação. O clichê Newtoniano “Se enxerguei mais longe é porque me apoiei em ombros de gigantes” é, de fato, a representação do mecanismo básico de funcionamento de todas as ciências. Assim, deixo aqui meus agradecimentos a um punhado de amigos, professores, técnicos e familiares;

Absolutamente privilegiado, sou filho de dois professores e irmão de mais dois. Claro que todos os pais, mães e irmãos, independentemente de suas formações, são professores por nascimento. Em meu caso, porém, sou cercado de 4 professores de fato. Todos treinados para tal, tendo exercido ou exercendo a profissão em momentos distintos de suas vidas. Minha mãe, é especialista da educação de infantil, meu pai e irmãos, especialistas no treinamento acadêmico-universitário. Eu, portanto, sou um eterno e afortunado aluno, dentro ou fora de casa. Assim, muito obrigado, Pai, Mãe e meus Irmãos.

Deixo um agradecimento ao Programa de Pós-graduação em Geologia e Geoquímica (PPGG) e a Biblioteca do Instituto de Geociências (IG) da Universidade Federal do Pará (UFPA) pela infraestrutura, pelo excelente quadro de funcionários e pela conceituada equipe docente. Mais uma vez privilegiado, pude passar boa parte de minha vida acadêmica vinculado a esta que, de longe, é a melhor Universidade do Norte do país.

Não poderia deixar de agradecer as agências nacionais de fomento à pesquisa CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), durante os 48 meses desta empreitada não houve um único mês em que não tenha contado com o apoio de ao menos umas destas duas agências. Juntas estas duas agências representam a esperança por dias melhores ao desenvolvimento científico brasileiro. Essa pesquisa recebeu o suporte dos projetos CNPq-870005/1997-9, 310283/2019-1, 306334/2020-8 e CAPES-001.

Um dos resultados mais inusitados deste doutorado, é, sem dúvida, a consolidação de uma empresa de C&T, lotada no Parque de Ciência e Tecnologia do Guamá (PCT-Guamá), especializada no uso da inteligência artificial (IA) para descoberta de padrões em imagens orbitais. A Solved – Soluções em Geoinformação, surgiu de um desafio técnico, proposto por Pedro Walfir (Orientador) e Tasso Azevedo (Coordenador da Rede MapBiomias), em busca de soluções inovadoras para o mapeamento dos ambientes costeiros no Brasil. Aos dois, expresso minha gratidão pela confiança na capacidade de solucionar os desafios propostos.

Assim, deixo aqui meus agradecimentos nominais a Luiz Cortinhas, coautor de todas as publicações científicas deste período de doutoramento, além de sócio fundador da empresa.

Agradeço igualmente a Luis Sadeck, Maria Luize, Bruno Haick, Alexandre Filho, Gilberto Nerino e Jhonatans Rodrigues, todos colaboradores Solved. Sem o auxílio prestado por cada um de vocês, este doutorado não teria chegado ao seu fim, tão pouco existiria a nossa empresa.

Por fim, agradeço a Nonata Trévia, pelos mais de 10 anos de companheirismo, amizade e paciência. Ser pesquisador por si só já é um desafio, é muito longa a jornada acadêmica entre graduação, mestrado, doutorado. Além da jornada por si própria, fazer da pesquisa, na Amazônia brasileira, a atividade principal de sustentação de uma empresa de tecnologia, é outro desafio gigantesco. Porém, nenhum destes é maior que o desafio que se iniciou semanas atrás. Poucos dias antes desta defesa, ganhei o maior de meus títulos, o de Pai. Assim, agradeço a meu filho, Gael, e sua mãe, pelos desafios dos próximos capítulos.

RESUMO

Desde a década de 80, o mapeamento de uso e cobertura da terra (LULC) tornou-se uma tarefa científica comum. No entanto, a identificação sistemática e contínua de qualquer uso ou cobertura terrestre, seja em escala global ou regional, exige grande capacidade de armazenamento e processamento. Esta tese apresenta dois fluxos de processamento de dados orbitais, gerenciados por computação em nuvem para avaliar: 1) a extensão anual dos manguezais brasileiros de 1985 a 2018, em conjunto com a criação e avaliação de um novo índice espectral, o Índice Modular de Reconhecimento de Manguezais (MMRI), que foi projetado especificamente para melhor discriminar as florestas de manguezal da vegetação circundante; e 2) a situação anual da aquicultura e da salicultura nas planícies costeiras do Brasil, de 1985 a 2019. No que se refere ao item 1, a cobertura do manguezal apresentou dois períodos de ocupação distintas, 1985-1998 e 1999-2018. O primeiro período mostra uma tendência ascendente, que parece estar mais relacionada à distribuição temporalmente desigual dos dados Landsat do que à regeneração dos manguezais brasileiros. No segundo período, foi registrada uma tendência de perda de área de manguezal, atingindo até 2% das florestas de manguezal. Em uma escala regional, ~ 80% da cobertura de manguezais do Brasil está localizada na Amazônia, nos estados do Maranhão, Pará e Amapá. Em termos de persistência, ~ 75% dos manguezais brasileiros permaneceram inalterados por duas décadas ou mais, em especial na Amazônia. Já no que tange o item 2, faz-se importante lembrar que a aquicultura e a produção de sal, são dois dos mais clássicos usos da terra costeiros em todo o mundo. No Brasil não é diferente, ambos os usos compõem atividade econômica relevante na Zona Costeira Brasileira (BCZ). No entanto, a discriminação automática de tais atividades, dissociando-as de coberturas ou usos outros, igualmente relacionados a presença de água em superfície, não é uma tarefa fácil. Espectralmente falando, água é água e, a menos que apresente uma alta concentração de compostos opticamente ativos, pouco se consegue fazer para dissociar uma variedade de alvos aquosos. Nesse sentido, Redes Neurais Convolucionais (CNN) têm a vantagem de prever o rótulo de determinado pixel, fornecendo como entrada uma região/local (patches ou chips) no entorno desse pixel. Juntas, a natureza convolucional das CNN, bem como a utilização de mecanismos de segmentação semântica, fornecem ao classificador U-Net, um tipo de CNN, a capacidade de acessar o “domínio do contexto” ao em vez de apenas valores de pixel isolados. Apoiados no domínio do contexto, em detrimento ao domínio puramente espectral, os resultados obtidos nesta tese mostram que as aquiculturas/salinas ocupavam ~356 km² em 1985 e ~544 km² em 2019, refletindo uma expansão de 52% (~188 km²), um aumento de 1,5x em 35 anos de ocupação da BCZ. De 1997 a 2015, a área aquícola cresceu por um fator

de ~1.7x, saltando de 349 km² para 583 km², 67% de expansão. Regionalmente, em 2019, o setor Nordeste concentra 93% das superfícies aquícolas/salineiras da BCZ, 6% situa-se no Sudeste e 1% no Sul. Curiosamente, apesar de apresentar extensas zonas costeiras e condições adequadas para o desenvolvimento de diferentes produtos aquícolas, a Amazônia não apresenta sinais relevantes de infraestrutura aquícola/salineira ao longo das 3 décadas analisadas.

Palavras-chave: Manguezais. Aquicultura. Salicultura. Zona Costeira. Aprendizado de Máquina. Random Forest. U-Net. Landsat. Séries Temporais.

ABSTRACT

Since the 1980s, land-use and land-cover (LULC) mapping has become a common scientific task. However, the systematic and continuous identification of any terrestrial use or cover, whether on a global or regional scale, demands large storage and processing capacities. This thesis presents two cloud computing pipelines to analyze: 1) the annual status of Brazilian mangroves from 1985 to 2018, along with a new spectral index, the Modular Mangrove Recognition Index (MMRI), which has been specifically designed to better discriminate mangrove forests from the surrounding vegetation, and 2) the annual status of the aquaculture and salt-culture over the Brazilian coastal plains. The mangrove cover showed two distinct occupation periods, 1985-1998 and 1999-2018. The first period shows an upward trend, which seems to be related more to the uneven distribution of Landsat data than to the regeneration of Brazilian mangroves. In the second period, a mangrove loss trend was registered, reaching up to 2% of the mangrove forest. On a regional scale, ~80% of Brazil's mangrove cover is located in the Amazon, Maranhao, Para, Amapa states. In terms of persistence, ~75% of the Brazilian mangroves remained unchanged for two decades or more, especially in the Brazilian Amazon. As for item 2, aquaculture and salt-culture are two of the most classical coastal land-uses worldwide. It isn't different in Brazil, where both land-uses are related to relevant economic activities in the Brazilian Coastal Zone (BCZ). However, to automatically discriminate such activities from other water-related covers/uses is not an easy task. Spectrally speaking, water is water and, unless it presents a high concentration of optically active compounds, not much can be done to dissociate a variety of water-related targets. In this sense, convolutional neural networks (CNN) have the advantage of predicting a given pixel's label by providing as input a local region (named patches or chips) around that pixel. Both the convolutional nature and the semantic segmentation capability allow the U-Net classifier, a type of CNN, to access the "context domain" instead of solely isolated pixel values. Backed by the context domain, the results obtained show that the BCZ aquaculture/saline ponds occupied ~356 km² in 1985 and ~544 km² in 2019, reflecting an area expansion of 52% (~185 km²), a rise of 1.5x in 35 years. From 1997 to 2015, the saline/aquaculture area grew by a factor of ~ 1.7, jumping from 349 km² to 583 km², a 67% increase. In 2019, the northeast sector concentrated 93% of the coastal aquaculture/salt-culture surface, 6% in Southeast and 1% in South. Interestingly, despite presenting extensive coastal zones and suitable conditions for developing different aquaculture products, the Amazon coast shows no relevant aquaculture infrastructure sign.

Keywords: Mangroves. Aquaculture. Salt-culture. Coastal Zone. Machine Learning. Random Forest. U-Net. Landsat. Time Series.

LISTA DE ILUSTRAÇÕES

- Figura 1 - Em verde, a definição legal da faixa terrestre da Zona Costeira Brasileira..... 8
- Figura 2 – Data de lançamento de cada um dos satélites da família Landsat. Fonte: USGS. .. 10
- Figura 3 - Características espectrais e espaciais dos sensores da família Landsat. O Multispectral Scanner System (MSS) a bordo do Landsats 1–5 tinha quatro bandas. O Thematic Mapper (TM) a bordo do Landsats 4 e 5 tinha sete bandas. O Enhanced Thematic Mapper Plus (ETM+) do Landsat 7 tem 8 bandas e o Landsats 8 e 9 com 11 bandas. Fonte: USGS..... 10
- Figura 4 – Os três principais domínios de informação do sensoriamento remoto. Os domínios do espectro, do tempo e do espaço. 12
- Figura 5 – Representação da Plataforma SatVeg (EMBRAPA-CNPTIA). Variação temporal de séries NDVI-MODIS sobre pastagem cultivada e agricultura anual de dois ciclos. 13
- Figura 6 – Características espectrais e temporais de corpos d’água. Em A, variação do MNDWI, em três alvos distintos. Em B, percentual de ocorrência de água superficial ao longo de 34 anos. Ainda em B, a tonalidade azul escuro representa ~100% de ocorrência, azul claro/ciano menos de 10%. Os marcadores em verde apontam para amostras oceânicas, vermelhos para estuários e azuis para aquicultura/salicultura. 15
- Figura 7 - O AM pode ser categorizado em supervisionado e não-supervisionado. Quando o classificador estima padrões gerais usando de instâncias preexistentes (amostras), temos o aprendizado supervisionado. Quando não há um resultado comparativo ou rótulo pré-existente, anterior a extração de atributos, considera-se o aprendizado como não-supervisionado. 17
- Figura 8 – O algoritmo K-means separa n pontos de dados em k clusters. De tal forma que, cada elemento pixel do dado de entrada seja classificado em um cluster que tenha a menor diferença entre o valor médio do cluster e o valor do pixel em questão. 19
- Figura 9 - O RF é constituído de árvores de decisão diversas (daí deriva o nome Forest), que usa fundamentalmente de dois conceitos estatísticos para adicionar maior flexibilidade e robustez as decisões de suas árvores; a aleatoriedade e a majoritariedade. 20

- Figura 10 - NN, se apresentam como uma maneira genérica de encontrar uma solução viável, de melhor ajuste possível (fit), a qualquer problema de previsibilidade numérica ou classificação categórica, independentemente do padrão de distribuição numérico ou de sua complexidade estatística. 22
- Figura 11 - A rede tem a forma de “U”. Essa arquitetura é simétrica e consiste em duas partes principais – o ramo a esquerda é chamado de contração (downsampling), sendo constituído pelo processo de convolução tradicional; a parte a direita é o ramo expansivo da rede, que é constituído de convoluções transpostas (upsampling). .. 23
- Figura 12 - convolução é uma técnica comum de processamento digital de imagens, em que se altera o valor de um pixel central de acordo com os valores de seus pixels circundantes. São aplicadas por janelas moveis (kernels), de tamanhos variados (3x3, 5x5, etc). As convulsões, portanto, podem contribuir para adicionar pistas texturais e contextuais a determinados alvos. 24
- Figura 13 – Redução da dimensionalidade x, y e aumento da profundidade em z. A matriz resultante tem mais camadas de informação texturais e contextuais, porém tamanho menor. 25
- Figura 14 – Fase de contração/downsampling. As convoluções permitem a inserção de informação textural e contextual a imagem. Ao passo que o objetivo da operação de maxpooling é reduzir a dimensionalidade de entrada (imagem, input, output, hidden layers, etc.) e permitir inferências sobre feições contidas nas sub-regiões agrupadas. 26
- Figura 15 - Convoluída várias vezes, ao final da etapa de downsampling a U-net perdeu a capacidade de se localizar. Assim, duas outras operações estático-matriciais entram em ação; as convoluções transpostas e as operações de padding. 27
- Figura 16 – Etapa final da U-Net. Já com camadas de informação textural e contextuais adicionadas, tamanho e capacidade de localização restauradas, é concluída a etapa de segmentação semântica, satisfazendo assim o objetivo inicial de previsibilidade/classificação matricial. 27
- Figura 17 - Todo o processamento de dados ocorre dentro da plataforma do Google Earth Engine - GEE, exceto pela classificação do padrão aquícola/salineiro, dependente da biblioteca TensorFlow. Em verde, etapas relacionadas ao desenho amostral. Em amarelo, etapas relacionadas à classificação. Em vermelho, fase de avaliação de precisão cartográfica. Finalmente, em salmão, fase de publicação dos dados em um “Web Portal” 33

- Figura 18 - O litoral brasileiro foi dividido em seis (7) setores distintos. Site 1 - Amapá; Site 2 - Ilha do Marajó; Site 3 - Pará / Maranhão; Site 4 - Piauí / Bahia; Site 5 - Espírito Santo / São Paulo; Site 6 - Paraná / Laguna (SC), Site 7 – Laguna/Rio Grande do Sul. As siglas a seguir representam os estados costeiros brasileiros: AL (Alagoas), AP (Amapá), BA (Bahia), CE (Ceará), ES (Espírito Santo), MA (Maranhão), PA (Pará), PB (Paraíba), PE (Pernambuco), PI (Piauí), PR (Paraná), RJ (Rio de Janeiro), RN (Rio Grande do Norte), SC (Santa Catarina), SE (Sergipe) e SP (São Paulo). 35
- Figura 19 - No filtro de preenchimento (gap-fill), os valores de no-data (“buracos”), não são permitidos e são substituídos pela classificação válida mais próxima no tempo. Se nenhuma posição válida “futura” estiver disponível, o valor de no-data é substituído pela posição anteriormente válida. Até três anos anteriores podem ser usados para preencher posições persistentemente sem dados. 38
- Figura 20 - O filtro espacial remove pixels que não compartilham vizinhos de valor idêntico. O valor mínimo de conexão é de 10 pixels. 39
- Figura 21 - Vermelho, amarelo e verde representam pixels de manguezal com frequências de ocorrência alta (23 ou mais anos, $y \geq 23$), média (entre 11 e 22 anos, $11 \leq y \leq 22$) e baixa (dez anos ou menos, $y < 11$), respectivamente. A imagem superior mostra os pixels de Mg antes de aplicar o filtro de frequência. A imagem inferior mostra pixels de Mg após a aplicação do filtro de frequência. As caixas pretas são centralizadas nas áreas que foram significativamente afetadas pelo filtro. Observe que todas as ocorrências de Mg com menos de 10% de persistência temporal (3 anos em 33 anos possíveis) foram filtradas. 40
- Figura 22 - Um total de 900 amostras independentes foram obtidas no nível do pixel Landsat, em 1985 e 2019, totalizando 1.800 amostras. A amostragem foi dividida em 3 estratos: 1) 450 amostras dentro de uma superfície de temporalmente unificada de AS. 2) 400 em uma região de buffer do estrato 1 e 3) fora dos limites dos estratos 1 e 2 foram adquiridas 50 amostras. 43

LISTA DE TABELAS

Tabela 1 - Resumo dos subtópicos do item de “Resultados e Discussões”, aqui apresentados na forma de artigos científicos em revistas indexadas.	6
Tabela 2 - Informações gerais sobre cada um dos satélites. Adaptado de (Chander <i>et al.</i> 2009, Masek <i>et al.</i> 2020, USGS 2015)	9
Tabela 3 - Características espaciais e espectrais dos sensores a bordo dos Landsat 5, Landsat 7 e Landsat 8.....	11
Tabela 4 – Disposição básica de uma Matriz de Confusão/Erro. Em A, nas linhas (rótulo em azul) estão dispostas as classificações e nas colunas (rótulo em amarelo) os dados de referência. Para cada classe, em linhas ou colunas, há um somatório contabilizado, A, B, C. O termo T, refere-se a soma agregada do total de todas as classes. A diagonal principal, em verde, representa os valores verdadeiros, concordantes, entre cada classe; VA, VB, VC. Em magenta, a cima da diagonal, os falsos positivos (FP). Em laranja/pastel, a baixo da diagonal, os falsos negativos (FN). Em B, denota-se que a inversão do posicionamento das referências e classificações, em linhas ou colunas, inverte a posição de FN e FP, bem como de todas as métricas de acurácia a eles relacionadas.....	29
Tabela 5 – Métricas de Acurácia. Classicamente, as métricas são relações estatísticas envolvendo as proporções entre erros e acertos, tanto os globais, envolvendo a matriz como um todo, quanto os individuais, de cada uma das categorias. As métricas mais comumente utilizadas na análise de dados espaciais são; Acurácia do Consumidor (ou do Usuário) - AC, Acurácia do Produtor - AP, Erros de Omissão - EO, Erros de Comissão - EC, Acurácia Global - AG, Erros de Quantidade - QD e Erros de Alocação – AD.....	31
Tabela 6 - Atributos dos classificadores Random Forest e U-Net. No total, dez atributos de classificação foram usados para treinar o classificador Random Forest. Seis atributos foram usados para treinar o modelo U-Net.....	37

Tabela 7 – O filtro temporal inspeciona a posição central de três anos consecutivos ("ternário"), e se as extremidades do ternário são idênticas, mas a posição central não é, então o pixel central é reclassificado para coincidir com sua classe temporal vizinha. T1, T2 e T3 significam posição um (1), posição (2) e posição (3). GR significa "regra genérica", enquanto que Mg, N-Mg, AS e N-AS representam as classes de Manguezal, Não-Manguezal, Aqua-Salicultura, Não-Aqua-Salicultura, respectivamente. 39

SUMARIO

DEDICATÓRIA	iv
AGRADECIMENTOS	v
RESUMO	vii
ABSTRACT	ix
LISTA DE ILUSTRAÇÕES	xi
LISTA DE TABELAS	xiv
1 INTRODUÇÃO	1
1.1 PROBLEMÁTICA E JUSTIFICATIVA	4
1.2 OBJETIVOS	5
1.3 ESTRUTURA DA TESE	5
2 FUNDAMENTAÇÃO TEÓRICA	7
2.1 ZONA COSTEIRA BRASILEIRA (ZCB) E SUA DEFINIÇÃO LEGAL	7
2.2 FAMÍLIA LANDSAT	9
2.3 DOMÍNIOS DA INFORMAÇÃO	12
2.4 ALGORITMOS DE APRENDIZADO DE MÁQUINA	16
2.4.1 Clusterização por k-means	18
2.4.2 Random Forest	19
2.4.3 A Rede Neural Convolucionária: U-Net	21
2.5 AVALIAÇÃO DE ACURÁCIA	28
3 MATERIAIS E MÉTODOS	33
3.1 PROCESSAMENTO DE DADOS	33
3.2 PÓS-CLASSIFICAÇÃO	38
3.3 DESIGN AMOSTRAL E ANÁLISE DE ACURÁCIA	40
3.3.1 Manguezal e não-manguezal	40
3.3.2 Aqua-salicultura e não-aqua-salicultura	42
3 RESULTADOS	43
4.1 ARTIGO I BRAZILIAN MANGROVE STATUS: THREE DECADES OF SATELLITE DATA ANALYSIS.	43
4.2 ARTIGO II A LARGE-SCALE DEEP-LEARNING APPROACH FOR MULTI- TEMPORAL AQUACULTURE MAPPING	64
4 CONCLUSÕES, CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES	83
REFERÊNCIAS	85

1. INTRODUÇÃO

A ocupação das zonas costeiras sempre atraiu a nós, humanos, em parte devido à sua abundância de recursos, em especial o fornecimento de recursos de subsistência, assim como por razões logísticas, uma vez que oferece pontos de acesso ao comércio e transporte marítimo. As regiões costeiras têm, de maneira intrínseca a sua natureza, uma ampla gama de riscos inerentes a sua ocupação, que incluem a elevação do nível do mar, frequente interação com variações de marés, o constante impacto de ondas, além de fortes ventos. Ainda assim, as áreas costeiras são mais densamente povoadas do que o interior continental e exibem taxas mais altas de crescimento populacional e urbanização (Neumann *et al.* 2015). As áreas costeiras compreendem 20% da superfície continental do planeta (Burke *et al.* 2000) e abrigam quase metade da população da Terra (Small & Nicholls 2003).

A zona costeira do Brasil não é diferente, estendendo-se por aproximadamente 9.200 km, essa paisagem dinâmica e de rápidas mudanças físicas e socioeconômicas abriga aproximadamente 18% da população do país, na medida em que 16 das 28 regiões metropolitanas brasileiras se encontram ao longo da costa (Nicolodi & Petermann 2010). A zona costeira brasileira (ZCB) apresenta um conjunto muito diversificado de ambientes costeiros que evoluíram durante o Quaternário, em resposta às mudanças no clima e no nível do mar, interagindo com o fornecimento variável de sedimentos e um patrimônio geológico que remonta da separação dos continentes Sul Americano e Africano, durante o mesozoico (Dominguez 2009). Os sistemas de manguezais estão entre esta diversidade de ambientes costeiros.

Globalmente, os manguezais estão distribuídos na região intertidal dos trópicos e subtropicais, entre aproximadamente 30 ° N e 30 ° S (Giri 2016). Em 2000, as florestas de manguezal representavam uma área total de 137.760 km², distribuídas em 118 países, representando cerca de 1% do total de florestas tropicais do mundo. (Giri *et al.* 2011). As florestas de manguezal são um tipo de vegetação sempre-verde, distribuídas do nível médio do mar até a linha de maré mais alta (Alongi 2009), crescendo em condições ambientais extremas, como alta frequência de inundação, alta salinidade, alta temperatura e solos anaeróbicos lamacentos.

Os manguezais desempenham um papel importante na sustentabilidade humana, fornecendo uma ampla gama de serviços ecossistêmicos, incluindo ciclagem de nutrientes, formação de solo, produção de madeira, área de desova de peixes e armazenamento de carbono (C) (Alongi 2002, Murdiyarso *et al.* 2015, Saenger *et al.* 1983), sendo um dos ecossistemas

mais produtivos e biologicamente complexos da Terra (Donato *et al.* 2011). Manguezais e áreas úmidas costeiras, sequestram carbono, por ano, a uma taxa de duas a quatro vezes maior que florestas tropicais e armazenam três a cinco vezes mais carbono por área equivalente (Murdiyarso *et al.* 2015, Sanderman *et al.* 2018).

Devido à grande importância desse ecossistema, desde a década de 80 são realizados trabalhos de mapeamento da cobertura de manguezais e detecção de mudanças em seus domínios (Giri *et al.* 2011, Saenger *et al.* 1983, Spalding *et al.* 2010). No entanto, não existem estudos na literatura, que permitam a identificação sistemática e contínua de manguezais e suas alterações, seja em escala global ou continental. Parte desta escassez é devido a: 1 - grande demanda por recursos computacionais, 2 - persistente cobertura de nuvens sobre áreas costeiras e 3 - escassez de mecanismos matemáticos projetados especificamente para mapeamento costeiro.

Independentemente da cobertura terrestre a ser identificada, qualquer identificação sistemática e contínua de padrões, sejam padrões de vegetação ou não, requer grande capacidade de armazenamento e processamento. Essas duas variáveis foram apenas recentemente contornadas, com o advento das plataformas de computação em nuvem, como o Google Earth Engine - GEE (Gorelick *et al.* 2017) e o Amazon Web Service - AWS (Chen *et al.* 2017), que combinam vários petabytes de dados orbitais e geoespaciais com recursos de análise estatística em escala planetária.

Além disso, essas plataformas permitem superar um dos principais problemas inerentes ao mapeamento do uso e cobertura da terra, a persistente cobertura de nuvens em algumas áreas do planeta. (Zhu *et al.* 2015). Zonas costeiras intertropicais não são exceção de tal característica. Pelo contrário, as zonas costeiras são severamente afetadas por nebulosidade atmosféricas, condição que é intensificada em razão de sua proximidade aos oceanos e de sua posição tropical. No entanto, a plataforma Google Earth Engine (GEE) foi criada de maneira a permitir a rápida filtragem e ordenamento de dados, facilitando a tarefa de selecionar dentre milhões de imagens e pixels, dados que atendam a critérios espaciais, temporais ou a outro conjunto de critérios preestabelecidos, como por exemplo a ausência de nuvens e sombras (Gorelick *et al.* 2017).

Ao mesmo tempo, são escassos os mecanismos espectrais especificamente projetados para apoiar a identificação de manguezais e sua diferenciação da vegetação circundante. Tradicionalmente, a detecção de manguezais utiliza-se, entre outros, de índices de vegetação clássicos como o NDVI (Tucker 1979), o NDWI (Gao 1996, McFeeters 1996), interpretação visual, classificação supervisionada, classificação não supervisionada e uso imagens micro-ondas (Alsaaidh *et al.* 2013, Fei *et al.* 2011, Giri *et al.* 2011, Magris & Barreto 2010, Nardin

et al. 2016, Nascimento Jr *et al.* 2013, Tong *et al.* 2004). Aqui, o chamado MMRI, Índice Modular de Reconhecimento de Manguezais, é apresentado (Diniz *et al.* 2019).

Ainda mais escassos que os estudos sistemáticos sobre a cobertura de manguezais nacional, são os estudos sobre a evolução dos parques aquícolas e da salicultura no Brasil. A aquicultura e a salicultura estão classicamente associadas a ocupação de ambientes costeiros, sendo a origem de mudanças de larga escala na cobertura e uso de solo costeiros (Ottinger, Clauss & Kuenzer 2016). Atualmente, metade dos peixes e moluscos usados para alimentar a população mundial vem da aquicultura (FAO 2018), e o apetite mundial por peixes e produtos pesqueiros não mostra sinais de diminuir. De 1990 a 2018, houve um aumento de ~ 14% na pesca de captura global, um aumento de 527% na produção aquícola e um aumento de 122% no consumo total de peixes para fins alimentícios (FAO 2020).

Compartilhando características semelhantes às da infraestrutura aquícola, como a exploração dos recursos hídricos, a necessidade de reservatórios artificiais de água e a dependência de ambientes costeiros, a salicultura é outra atividade econômica relevante na Zona Costeira Brasileira (BCZ). Nesse sentido, a produção brasileira de sal marinho está associada principalmente a três estados: Rio Grande do Norte (RN), Ceará (CE) e Rio de Janeiro (RJ) (ANM 2018, Diniz *et al.* 2017). Em 2017, foram produzidos 6 Mt de sal marinho no Brasil; o estado do RN foi responsável por 95%, o RJ por 3% e outros 1,5% vem do CE (ANM 2018). Portanto, o mapeamento das áreas de aquicultura/salicultura é um desafio igualmente fundamental para a compreensão da expansão dessas atividades e do manejo sustentável desses recursos costeiros.

A principal dificuldade em distinguir, automaticamente, entre a aquicultura/salicultura ou qualquer outro alvo que apresente uma lâmina d'água superficial como característica predominante (por exemplo, rios, lagos ou águas costeiras), é a semelhança espectral entre todos eles. Espectralmente falando, água é água e, a menos que apresente uma alta concentração de compostos opticamente ativos (por exemplo, sedimentos em suspensão, algas, matéria orgânica, etc.), pouco pode ser feito para discriminar estes alvos.

Nesse cenário, a rede convolucionária U-Net (Ronneberger *et al.* 2015) tem a vantagem de prever o rótulo da classe de cada pixel, fornecendo como entrada uma região local (chamada de "patches" ou "chips") ao entorno de determinado pixel. Essa característica, juntamente com sua natureza convolucional e a capacidade de segmentação semântica, fornecem à U-Net a capacidade de acessar o "domínio de contexto" em vez de apenas valores de pixel isolados. Em casos como este, o domínio do contexto, ao invés do espectral ou temporal, passa a ser a chave para distinguir entre rios, lagos, aquicultura e pixels de água costeira.

O objetivo deste documento é apresentar os dados sobre o status anual da cobertura de manguezal e da evolução do uso aquícola/salineiro no país, de 1985 a 2019 e, ainda, verificar a robustez do novo índice espectral, o Modular Mangrove Recognition Index - MMRI, especificamente projetado para melhor discriminar as florestas de manguezal de sua vegetação circundante.

1.1. PROBLEMÁTICA E JUSTIFICATIVA

A expressão “uso e cobertura da terra” (LULC) carrega consigo duas terminologias distintas que são frequentemente usadas de forma intercambiável (McConnell 2015). O uso da terra (LU) se refere à forma como a terra foi usada pelos humanos, geralmente com ênfase no papel funcional da superfície terrestre para o desenvolvimento de atividades econômicas. Enquanto que, a cobertura da terra (LC) se refere às características biofísicas da superfície terrestre, incluindo a distribuição da vegetação, água, solo e outras características físicas naturais (Phiri & Morgenroth 2017).

Neste sentido, o processo de urbanização, por exemplo, pode transformar uma grande quantidade de terras agrícolas e/ou florestais em áreas urbanas. De maneira similar, em regiões costeiras a expansão urbana, por sua vez, pode ocupar terrenos dunares, frentes de praia, manguezais e outros sistemas-costeiros nativos. O termo mudança de uso e cobertura da terra (LULCC) refere-se, portanto, a conversões de coberturas naturais em diferentes tipos de uso da terra, sendo este o resultado de interações complexas entre humanos e o ambiente (Gómez *et al.* 2016).

Em estudos anteriores, as mudanças ambientais globais, como as emissões de gases de efeito estufa, as mudanças climáticas globais, a perda de biodiversidade e a perda de recursos do solo foram intimamente ligadas às mudanças de LULC (Guzha *et al.* 2018, Salazar *et al.* 2015, Vuuren van *et al.* 2017). A LULCC é um grande impulsionador das mudanças globais e tem impacto significativo nos processos ecossistêmicos, nos ciclos biológicos e sobre a biodiversidade (Alencar *et al.* 2020, Oertel *et al.* 2016, Rosa *et al.* 2021, Sasmito *et al.* 2019, Souza *et al.* 2020).

O mapeamento da dinâmica costeira, suas coberturas e usos, é tarefa tradicionalmente executada, porém em escalas espaciais/temporais curtas e, quase sempre, relativo ao mapeamento exclusivo da cobertura de manguezais (Giri *et al.* 2011, Nascimento Jr *et al.* 2013, Saenger *et al.* 1983, Souza Filho 2005, Spalding *et al.* 2010). Nesse sentido, esta tese contribui com a espacialização das mudanças de uso e cobertura costeiras ao longo de toda a ZCB,

revelando o padrão espaço-temporal não apenas da cobertura de manguezais, mas também de um dos mais clássicos usos costeiros, os parques aquícolas/salineiros instalados na costa brasileira.

Esta pesquisa cobre, portanto, um hiato de informação espaço-temporal na ciência brasileira, disponibilizando dados primários sobre a evolução dos padrões de uso e cobertura da terra na ZCB de maneira anual e sistemática, desde 1985 até 2019.

1.2. OBJETIVOS

O objetivo desta pesquisa é detectar e mapear as mudanças de uso e cobertura da terra associados a vegetação de manguezal e a atividade aquícola/salineira na costa brasileira, no período de 1985 a 2019.

Destacam-se, a seguir, alguns objetivos específicos:

- i. Criar e verificar a robustez de um índice espectral, especificamente projetado para melhor discriminar as florestas de manguezal de sua vegetação circundante
- ii. Realizar a detecção de mudanças entre os anos especificados, analisando as variações de aumento e/ou redução das áreas de manguezal.
- iii. Realizar a detecção de mudanças entre os anos especificados, analisando as variações de aumento e/ou redução dos parques aquícola/salineiros na costa brasileira.

1.3. ESTRUTURA DA TESE

A presente tese utiliza do mecanismo de agregação de artigos científicos e está estruturada de acordo à seguinte configuração:

1 Introdução, composta pelos itens: 1.1 Problemática e justificativa, 1.2 Objetivos e 1.3 Estrutura da tese;

2 Fundamentação teórica, responsável por reunir informações importantes para a construção do arcabouço teórico deste projeto, abordando os seguintes tópicos: 2.1 Zona Costeira Brasileira, 2.2 Família Landsat, 2.3 Os Domínios da Informação, 2.4 Algoritmos de Machine Learning e 2.5 Avaliação de Acurácia;

3 Materiais e métodos, composto por: 3.1 Processamento de Dados Orbitais 3.2 Pós-Classificação e 3.3 Design Amostral e Análise de Acurácia.

4 Resultados e discussões, são apresentados por agregação de artigos científicos, um publicado e um submetido, sendo o autor da tese o primeiro autor dos dois trabalhos,

satisfazendo assim o regimento do Programa de Pós-Graduação em Geologia e Geoquímica (PPGG) da Universidade Federal do Pará (UFPA). O corpo principal deste capítulo é constituído conforme a Tabela 1, onde são apresentadas as abordagens conduzidas em cada um dos artigos e o status de acompanhamento.

Tabela 1 - Resumo dos subtópicos do item de “Resultados e Discussões”, aqui apresentados na forma de artigos científicos em revistas indexadas.

Número do item – Título do manuscrito	Principais resultados (Highlights)	Status de acompanhamento
4.1 Artigo I – <i>Brazilian Mangrove Status: Three Decades of Satellite Data Analysis.</i>	<ul style="list-style-type: none"> • Apresentou e avaliou o MMRI – Modular Mangrove Recognition Index. • Apresentou os dados da evolução anual das florestas de manguezal de 1985 a 2018 • Avaliou a persistência temporal dos Manguezais para o País e para seus Estados. • Consolidou um método escalável e replicável em outras regiões e classes de uso e cobertura. 	Publicado na <i>Remote Sensing</i> (Diniz <i>et al.</i> , 2019).
4.2 Artigo II - <i>A Large-Scale Deep-Learning Approach for Multi-Temporal Aquaculture Mapping.</i>	<ul style="list-style-type: none"> • Utilizou-se de uma Rede Neural Convulsional (CNN) para mapear as regiões aquícolas e as salinas da ZCB. • Apresentou os dados da evolução anual das salinas/aquicultura de 1985 a 2019. • Avaliou a persistência temporal desta atividade em nos diferentes estados do país. • Consolidou um método escalável e replicável em outras regiões e classes de uso e cobertura. 	Submetido na <i>Remote Sensing</i> (22 de fevereiro de 2021)

Fonte: Elaborado pelo autor.

Por fim, o capítulo **5 Conclusões** representa a integração do conteúdo exposto nos artigos vinculados a esta tese, de modo a apresentar as conclusões obtidas e futuras lacunas a serem preenchidas.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. ZONA COSTEIRA BRASILEIRA (ZCB) E SUA DEFINIÇÃO LEGAL

De acordo com o Decreto nº 5.300 / 2004, a Zona Costeira Brasileira (BCZ) corresponde ao espaço de interação geográfica entre ar, mar e terra, incluindo seus recursos renováveis ou não renováveis, cobrindo uma faixa marítima e uma faixa terrestre, com os seguintes limites:

- Faixa marítima; espaço que se estende por doze milhas náuticas, medidas a partir de sua linha de base¹, abrangendo assim todo o mar territorial.
- Faixa terrestre; porção da terra diretamente influenciada pelos fenômenos da zona costeira; isto se refere especificamente aos municípios ao longo do Oceano Atlântico, e outros municípios onde existam ecossistemas costeiros de alta relevância e onde impactos ambientais e/ou de infraestrutura, devido às atividades da zona costeira, tenham sido registrados.

No que diz respeito a este documento, a região de mapeamento do ZCB engloba diretamente a "Faixa Terrestre/Continental", definida pelo sistema costeiro-marinho (IBGE 2019a), Figura 1.

¹**Linha de base** – De acordo com a UNCLOS (Convenção das Nações Unidas sobre o Direito do Mar), a linha de base normal para medir a largura do mar territorial é a linha de baixa-mar ao longo da costa, tal como indicada nas cartas marítimas de grande escala, reconhecidas oficialmente pelo Estado costeiro.

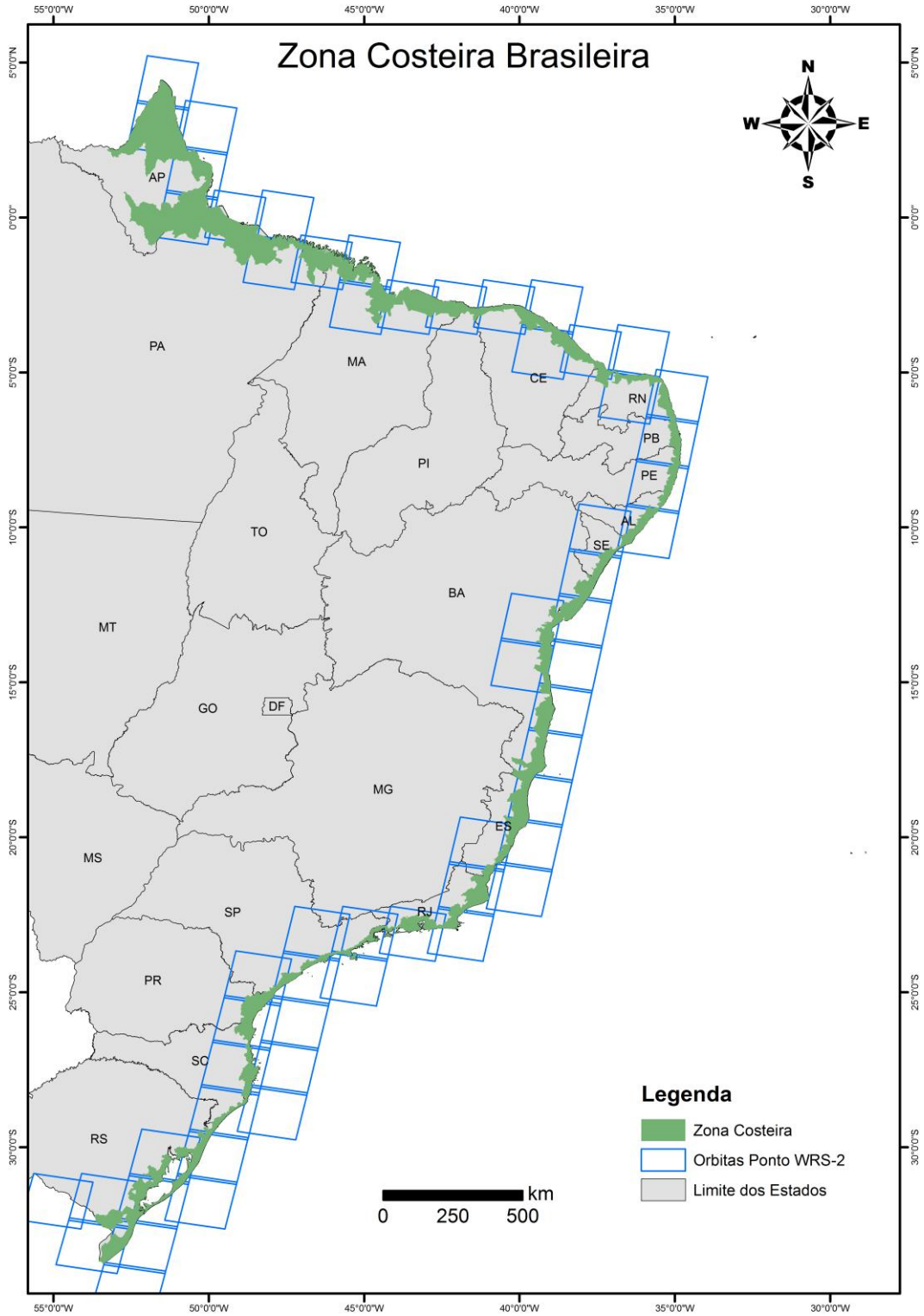


Figura 1 - Em verde, a definição legal da faixa terrestre da Zona Costeira Brasileira

2.2. FAMÍLIA LANDSAT

O programa Landsat iniciou uma revolução no sensoriamento remoto da Terra, ainda no começo da década de 1970 (Rocchio *et al.* 2018). Com sete missões bem sucedidas ao longo de mais de 40 anos, o programa Landsat documentou e continua a documentar a superfície terrestre e sua evolução. As missões e sensores Landsat evoluíram junto com a tecnologia, saltando de um projeto vinculado ao mundo analógico da interpretação visual, para uma missão plenamente digital, em escala planetária, com sucessivas melhorias ao longo do caminho em termos de desempenho espectral, espacial, radiométrico e geométrico, bem como na estratégia de aquisição e disponibilidade de dados e produtos (Emery & Camps 2017, Markham *et al.* 2018).

A série de satélites da família Landsat fornece o mais longo registro contínuo de observações terrestres. Como tal, o programa Landsat é um recurso inestimável para monitorar as mudanças globais, sendo uma fonte primária de observações da Terra, em média resolução, e portanto primordial ferramenta no auxílio a tomada de decisões (Chander *et al.* 2009, Fuller *et al.* 1994, Masek *et al.* 2008, Vogelmann *et al.* 2001).

O Programa Landsat começou no início de 1972 com o lançamento do primeiro satélite da série, inicialmente nomeado de Earth Resources Technology Satellite (ERTS-1), posteriormente renomeado para Landsat 1. À medida que as capacidades tecnológicas aumentaram, também aumentaram a quantidade e a qualidade dos dados capturados pelos vários sensores a bordo dos satélites. A Tabela 2 e Figura 2, apresentam informações gerais sobre cada um dos satélites Landsat.

Tabela 2 - Informações gerais sobre cada um dos satélites.

Satélite	Sensores	Lançamento	Desativação	Altitude (km)	Inclinação (graus)	Período (min)	Revisita (dias)	Passagem
Landsat 1	MSS e RBV	23/07/1972	07/01/1978	920	99.20	103.34	18	9:30
Landsat 2	MSS e RBV	22/01/1975	25/02/1982	920	99.20	103.34	18	9:30
Landsat 3	MSS e RBV	05/03/1978	31/03/1983	920	99.20	103.34	18	9:30
Landsat 4	MSS e TM	16/07/1982	30/06/2001	705	98.20	98.20	16	9:45
Landsat 5	MSS e TM	01/03/1984	05/06/2013	705	98.20	98.20	16	9:45
Landsat 6	ETM	05/10/1993	Não alcançou a órbita	---	---	---	---	---
Landsat 7	ETM+	15/04/1999	05/04/1999	705	98.20	98.20	16	10:00
Landsat 8	OLI e TIRS	11/02/2013	---	705	98.20	98.20	16	10:00
Landsat 9	OLI2 e TIRS2	09/2021 Previsão	---	705	98.20	98.20	16	10:00

Fonte: Adaptado de (Chander *et al.* 2009, Masek *et al.* 2020; USGS 2015)

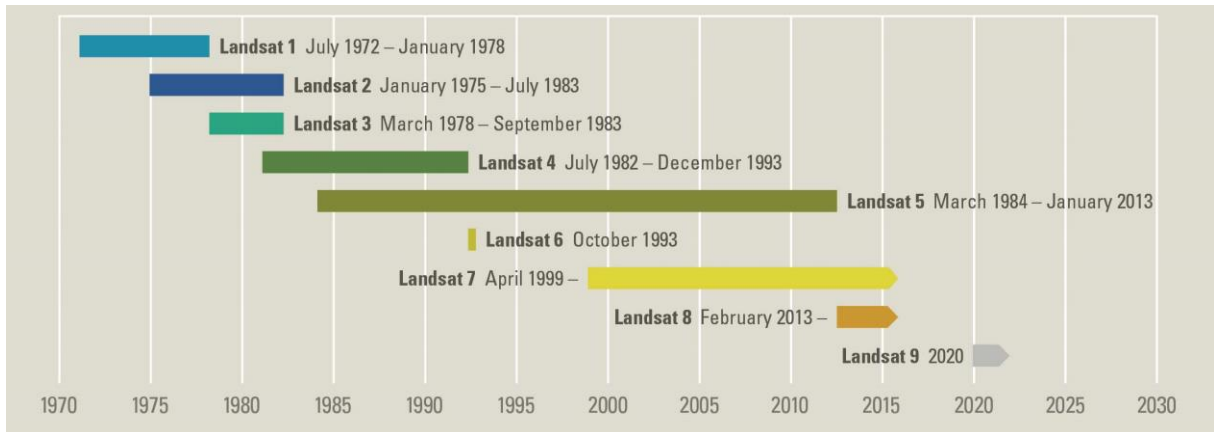


Figura 2 – Data de lançamento de cada um dos satélites da família Landsat. Fonte: USGS.

Desde o início do programa, ao menos um satélite Landsat esteve operacional, garantindo mais de 40 anos de imageamentos ininterruptos da superfície terrestre, Tabela 2 e Figura 2. Dos nove sistemas satelitários desenvolvidos, apenas um, o Landsat 6, falhou em alcançar sua órbita. Seu predecessor, Landsat 5, carrega o recorde de longevidade dentre os satélites de imageamento terrestre, permanecendo em atividade por 29 anos.

Do ponto de vista físico, as características espectrais dos sensores da família Landsat variam ao longo do tempo (Figura 3). Espectralmente, o Multispectral Scanner System (MSS) a bordo dos Landsats 1–5 possuía quatro bandas fisicamente distintas. O Thematic Mapper (TM) a bordo dos Landsats 4 e 5, sete bandas. O Enhanced Thematic Mapper Plus (ETM+) do Landsat 7 tem 8 bandas e os Landsats 8 e 9 com 11 bandas.

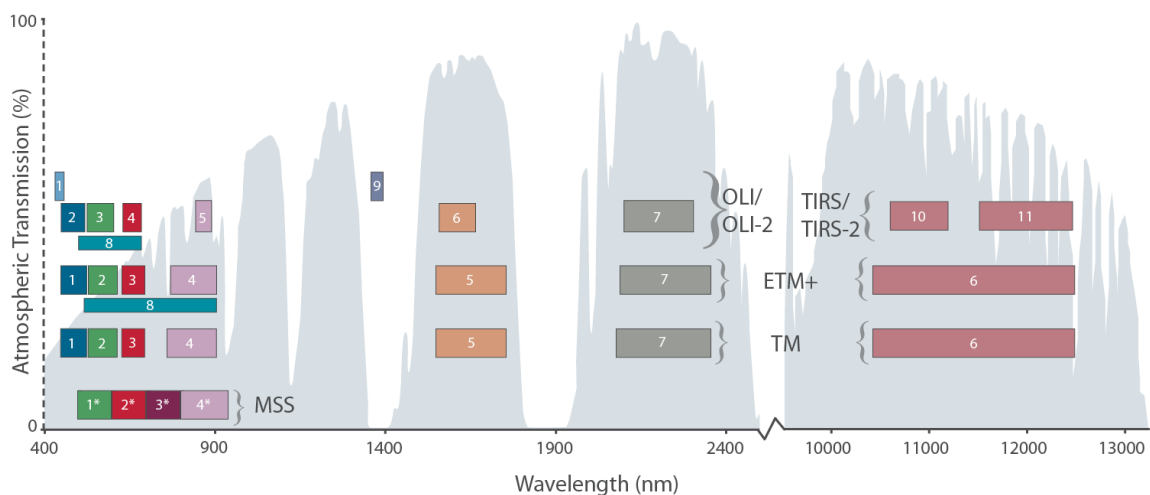


Figura 3 - Características espectrais e espaciais dos sensores da família Landsat. O Multispectral Scanner System (MSS) a bordo dos Landsats 1–5 tinha quatro bandas. O Thematic Mapper (TM) a bordo dos Landsats 4 e 5 tinha sete bandas. O Enhanced Thematic Mapper Plus (ETM+) do Landsat 7 tem 8 bandas e os Landsats 8 e 9 com 11 bandas. Fonte: USGS

Juntas, as bandas espectrais dos diferentes Landsats cobrem diferentes regiões do espectro eletromagnético. A região do visível, é representada pelas bandas azul (B), verde (G) e vermelho (R). A região do infravermelho, com as bandas do infravermelho próximo (NIR), do infravermelho de ondas curtas 1 e 2 (SWIR 1 e 2), e do infravermelho termal, com duas bandas no Landsat 8 e 9 (TIRS e TIRS-2) e apenas uma nos Landsats 5 e 7.

As bandas costeira/aerossol e a banda cirrus são inovações dos Landsat 8 e 9. A banda costeira, foi projetada para facilitar a detecção de clorofila, importante atributo da colorimetria de águas superficiais, em especial nas zonas costeiras. Já a banda cirrus, foi projetada para facilitar a detecção deste tipo de nuvem, juntamente com suas sombras. A banda cirrus está posicionada no intervalo do espectro eletromagnético onde as nuvens são mais visíveis.

Especialmente, dependendo do sensor e da banda em questão, a resolução espacial pode variar entre 15, 30, 60, 90 e 120 metros nos vários sensores da família Landsat. A tabela 3 agrupa as características espaciais e espectrais dos sensores a bordo dos Landsat 5, Landsat 7 e Landsat 8.

Tabela 3 - Características espaciais e espectrais dos sensores a bordo dos Landsat 5, Landsat 7 e Landsat 8

Bandas Landsat-5 TM (μm)			Bandas Landsat-7 ETM+ (μm)			Bandas Landsat-8 OLI e TIRS (μm)		
---			---			30 m Coastal/Aerosol	0.43 - 0.45	Band 1
Band 1	30 m Blue	0.45 - 0.52	Band 1	30 m Blue	0.44 - 0.51	30 m Blue	0.45 - 0.51	Band 2
Band 2	30 m Green	0.52 - 0.60	Band 2	30 m Green	0.52 - 0.60	30 m Green	0.53 - 0.59	Band 3
Band 3	30 m Red	0.63 - 0.69	Band 3	30 m Red	0.63 - 0.69	30 m Red	0.63 - 0.67	Band 4
Band 4	30 m NIR	0.76 - 0.90	Band 4	30 m NIR	0.77 - 0.90	30 m NIR	0.85 - 0.87	Band 5
Band 5	30 m SWIR-I	1.55 - 1.75	Band 5	30 m SWIR-I	1.54 - 1.75	30 m SWIR-I	1.56 - 1.65	Band 6
Band 6	120 m TIR	10.4 - 12.5	Band 6	60 m TIR	10.3 - 12.3	100 m TIR-I	10.60- 11.19	Band 10
						100 m TIR-2	11.50- 12.51	Band 11
Band 7	30 m SWIR-2	2.08 - 2.35	Band 7	30 m SWIR-2	2.06 - 2.34	30 m SWIR-2	2.10 - 2.29	Band 7
---			Band 8	15 m Pan	0.51 - 0.89	15 m Pan	0.50 - 0.67	Band 8
---			---			30 m Cirrus	1.36 - 1.38	Band 9

2.3. DOMÍNIOS DA INFORMAÇÃO

O sensoriamento remoto adquire informações sobre a superfície da Terra provenientes de três domínios essenciais (Jensen 2014, Novo 2010): o domínio do espectro (comprimento de onda ou frequência, propriedades reflexivas ou emissivas); o domínio do espaço (forma e tamanho do objeto, posição, local, distribuição, textura e vizinhança); e o domínio do tempo (mudanças do objeto/superfície ao longo do tempo), Figura 4.

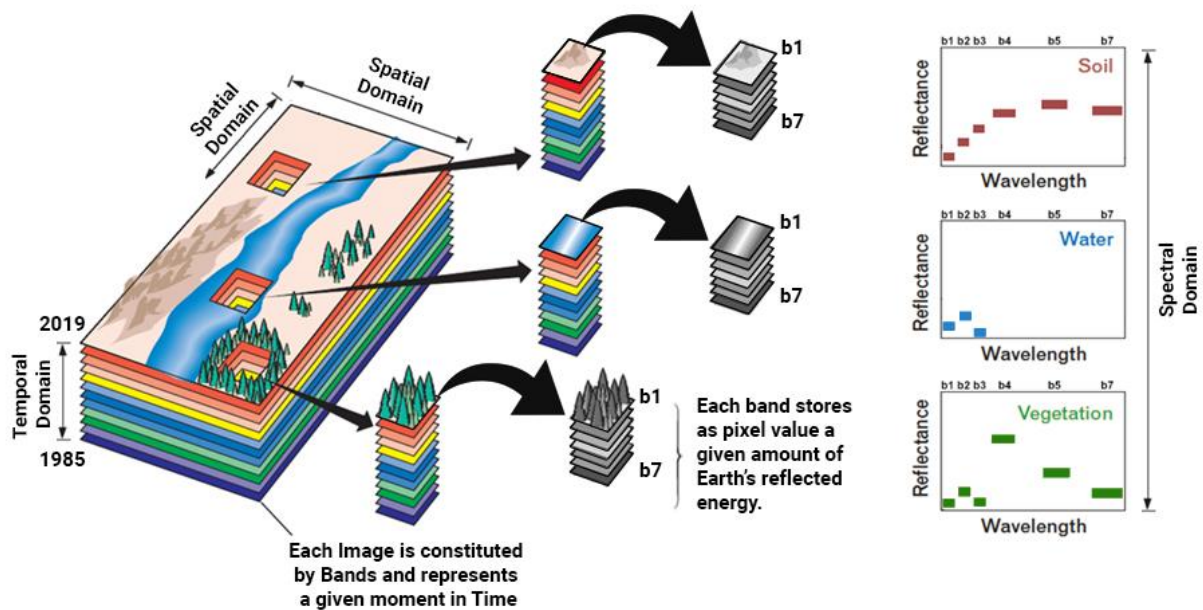


Figura 4 – Os três principais domínios de informação do sensoriamento remoto. Os domínios do espectro, do tempo e do espaço.

Maior parte dos estudos inerentes ao sensoriamento remoto se baseiam no desenvolvimento de uma relação determinística entre a quantidade de energia eletromagnética refletida, emitida ou espalhada e as características químicas, biológicas e físicas dos fenômenos, superfícies ou objetos que se investigam (Jensen 2014, Novo 2010). Nesse sentido, é o domínio espectral que fornece a informação responsável pela separação entre os alvos “A” ou “B” na superfície terrestre. Onde, a resolução espectral é definida pelo número e dimensão (tamanho) dos intervalos de comprimento de onda (bandas) para o qual um sensor é sensível (Jensen 2014, Novo 2010).

Todavia, as características puramente espectrais, entre dois ou mais alvos, podem não ser suficientes para que estes sejam distinguidos entre si. Em casos como este, de evidente similaridade espectral, é necessário lançar mão de atributos de domínios outros, como tempo ou espaço, para que se consiga atingir a individualização entre alvos.

É o que acontece, por exemplo, na separação entre pastagem e certos tipos de cultivos agrícolas (Adami *et al.* 2012, Rudorff *et al.* 2010). Estes dois usos da terra, podem apresentar sinal espectral característico de uma diversidade de momentos fenológicos. Pastagem e cultivos de cana-de-açúcar, por exemplo, experimentam, ainda que sejam usos da terra distintos, fases fenológicas similares e, portanto, diferentes padrões espectrais ao longo do desenvolvimento vegetal. Entretanto, ambos os usos, podem apresentar simultaneidade temporal entre uma ou mais fases de seu desenvolvimento (preparo do solo, semeadura, maturação vegetal e colheita).

Dessa maneira, sem o suporte das séries temporais, sempre existiria para estes dois usos a probabilidade de estarmos observando, em um momento instantâneo t , a mesma fase de desenvolvimento fenológico, sejam estas pastagens ou agricultura. O domínio do tempo, é então, absolutamente relevante para a separação de padrões agrícolas e pastoris (Figura 5).

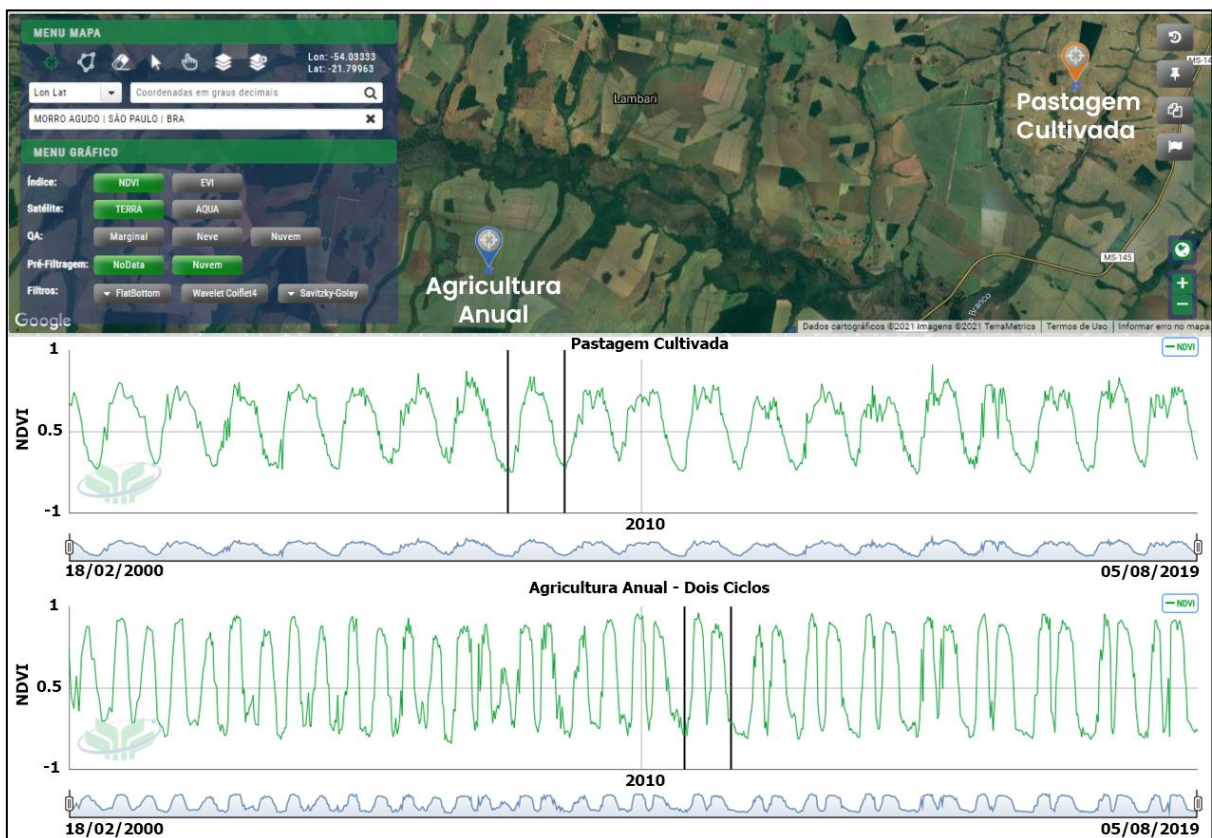


Figura 5 – Representação da Plataforma SatVeg (EMBRAPA-CNPTIA). Variação temporal de séries NDVI-MODIS sobre pastagem cultivada e agricultura anual de dois ciclos.

Por vezes, porém, nem propriedades temporais e tão pouco as espectrais, são suficientes para discriminar alvos "super semelhantes", que acabam por se comportar de maneira similar em ambos os domínios, o do espectro e do tempo. Esse é o caso da separação dos padrões

aquícolas e salineiros de alvos outros, que contenham água em superfície como característica predominante ao longo do tempo (Figura 6A e 6B).

No litoral brasileiro, a aquicultura e a salicultura compõem atividade econômica tradicional (Costa *et al.* 2013, IBGE 2019b, Lima 2017, Queiroz *et al.* 2013, Tenório *et al.* 2015). Assim, no que diz respeito a persistência temporal da atividade aquícola-salineira é natural que exista um conjunto diversificado de consolidação desta atividade, com infraestruturas mais antigas, mais recentes e padrões de abandono da atividade (Costa *et al.* 2013, Diniz *et al.* 2017, Lima 2017). Em decorrência disso, o domínio temporal pode renunciar à capacidade de distinguir entre padrões aquícolas/salineiros bem consolidados, de canais principais de rios, lagos e lagoas perenes e águas abertas (Figura 6B), uma vez que todas essas feições podem apresentar elevada persistência temporal ao longo dos anos (Pekel *et al.* 2016, Souza *et al.* 2019).

De maneira similar, no domínio do espectro, a menos que se apresente uma alta concentração de compostos opticamente ativos, por exemplo, sedimentos em suspensão, algas, matéria orgânica e outros (Marinho *et al.* 2021, Marinho *et al.* 2020, Mertes *et al.* 1993), pouco pode ser feito para discriminar espectralmente uma variedade de corpos d'água (Barbosa *et al.* 2019, Gao 1996, Xu 2006). Em casos como este, onde a cobertura natural entre os alvos a serem separados é a mesma e a persistência temporal é também similar, o domínio do espaço/contexto, ao invés do espectral ou temporal, pode ser a chave para distinguir entre rios, lagos e lagoas perenes e águas costeiras.

A análise da vizinhança, é essencial para acessar o domínio do espaço/contexto. Nesse cenário, os classificadores contextuais têm a vantagem de prever o rótulo de cada pixel, partindo de regiões de treinamento ao entorno de cada posição da matriz. A estratégia de agrupamento da vizinhança, de diferentes maneiras, fundamenta o funcionamento dos classificadores contextuais, sejam eles orientados a objeto - OBIA (Baatz & Schape., 2000); baseado em redes neurais convulsionárias – CNN (Jin *et al.*, 2019) ou no uso combinado dos dois – OCNN (Zhang *et al.*, 2018), fornecendo a todos estes a capacidade de acessar o "domínio do contexto" ao invés de se ater a valores de pixel isolados.

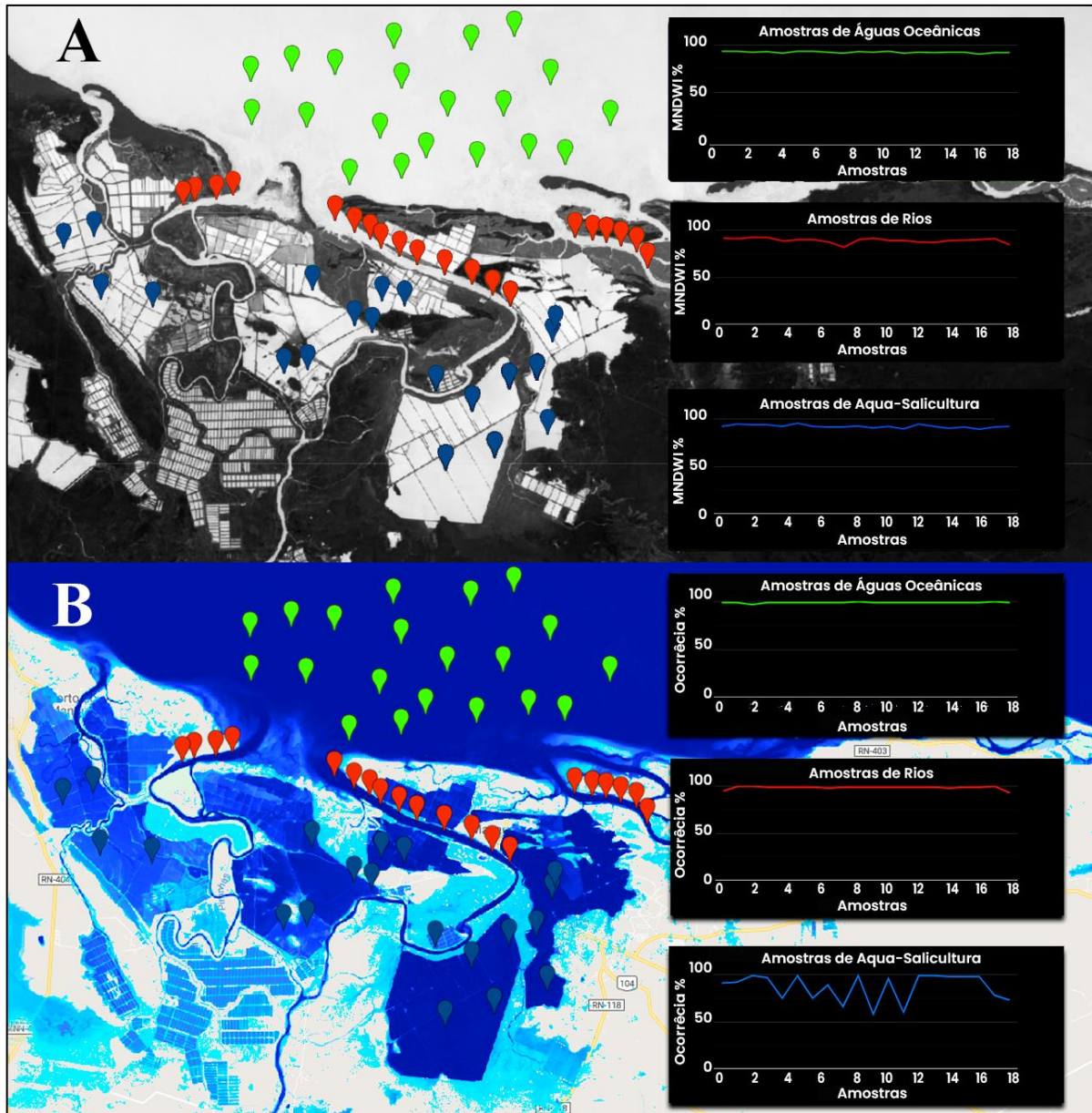


Figura 6 – Características espectrais e temporais de corpos d’água. Em A, variação do MNDWI, em três alvos distintos. Em B, percentual de ocorrência de água superficial ao longo de 34 anos. Ainda em B, a tonalidade azul escuro representa ~100% de ocorrência, azul claro/ciano menos de 10%. Os marcadores em verde apontam para amostras oceânicas, vermelhos para estuários e azuis para aquicultura/salicultura.

2.4. ALGORITMOS DE APRENDIZADO DE MÁQUINA

O aprendizado de máquina (AM), é uma subdivisão da inteligência artificial (IA). A abordagem do AM, lida com o design de algoritmos capazes de “aprender” por meio de dados que possam ser lidos por uma máquina (Lary *et al.* 2016). Em um sentido mais amplo do mundo das máquinas, a palavra "aprender", significa extrair atributos de um conjunto de dados, operar estatisticamente sobre eles, reconhecer a existência de padrões estatísticos, generalizar o padrão reconhecido e, finalmente, prever, detectar ou classificar padrões e cenários semelhantes ao “aprendido”.

O AM tornou-se um dos principais focos da literatura de suporte a classificação de imagens no sensoriamento remoto (Belgiu & Drăguț 2016, Gorelick *et al.* 2017, Hansen *et al.* 2000, Pekel *et al.* 2016, Souza *et al.* 2020). Os algoritmos de AM são, em geral, capazes de modelar assinaturas espectrais complexas, podendo aceitar uma variedade de dados de entrada e não fazem suposições sobre a distribuição de dados, ou seja, são não-paramétricos (Maxwell *et al.* 2018).

Vastamente documentado na literatura científica, esses métodos tendem a produzir maiores níveis de acurácia quando comparados a classificadores paramétricos tradicionais, especialmente para dados complexos, com um conjunto extenso de variáveis de entrada (Arruda *et al.* 2021, Belgiu & Drăguț 2016, Diniz *et al.* 2019, Ghimire *et al.* 2012, Hansen *et al.* 2000, Parente *et al.* 2019, Rosa *et al.* 2021, Souza *et al.* 2019, 2020).

As abordagens de aprendizado de máquina tornaram-se amplamente aceitas, como é evidenciado por sua utilização no mapeamento operacional do uso e cobertura da terra. Entretanto deve-se mencionar que classificadores baseados em AM continuam a obedecer, meticulosamente, a estatística da separabilidade de dados numéricos e a física do sensoriamento remoto. Assim, anterior a escolha de um classificador, é a necessidade do reconhecimento das características biofísicas de seu alvo de interesse, bem como do comportamento deste nos três domínios da informação; o espectro, o tempo e o espaço.

Do ponto de vista conceitual o AM pode ser categorizado de acordo com a estratégia de extração de atributos a qual obedece (Figura 7). Quando o classificador estima, prediz ou classifica padrões gerais usando de instâncias preexistentes (amostras), com o objetivo de prever ou rotular o resultado de instâncias futuras, chamamos de aprendizado supervisionado. Quando não há um resultado comparativo ou rótulo pré-existente, anterior a extração de atributos, considera-se o aprendizado como não-supervisionado.

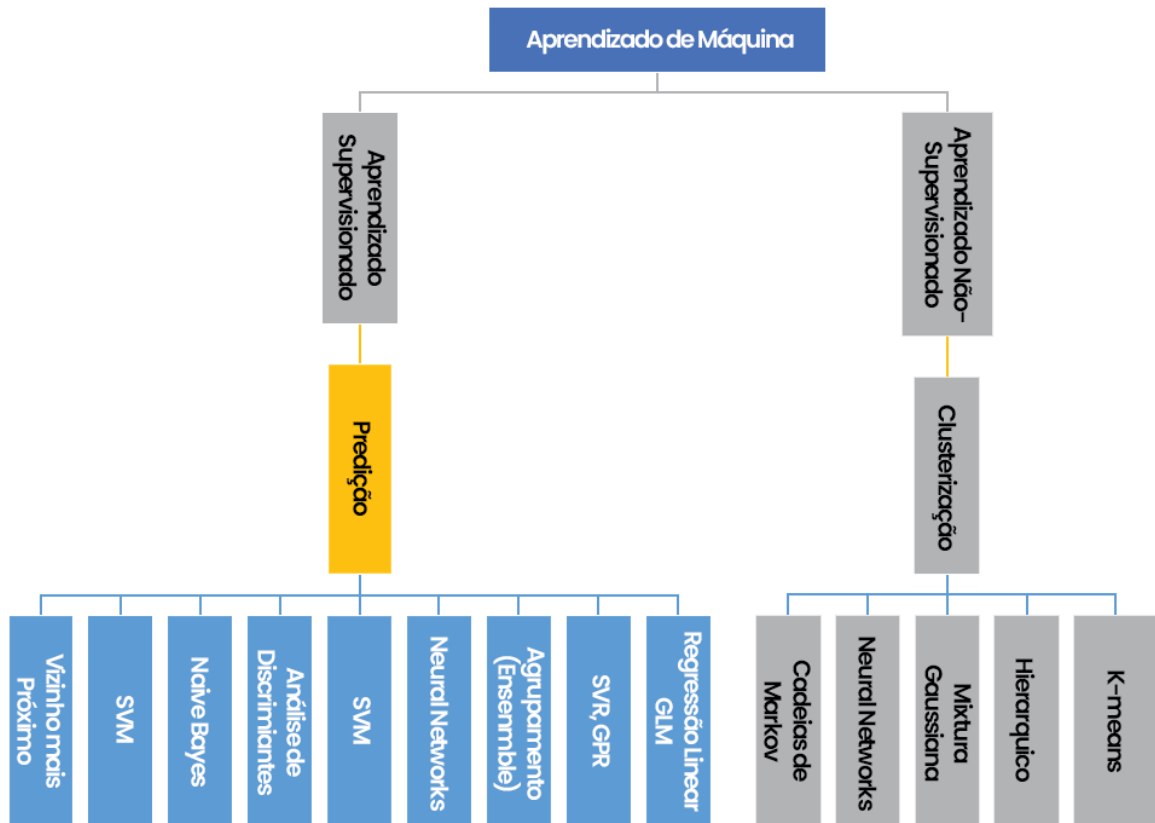


Figura 7 - O AM pode ser categorizado em supervisionado e não-supervisionado. Quando o classificador estima padrões gerais usando de instâncias preexistentes (amostras), temos o aprendizado supervisionado. Quando não há um resultado comparativo ou rótulo pré-existente, anterior a extração de atributos, considera-se o aprendizado como não-supervisionado.

Na família do aprendizado supervisionado, destaca-se o subgrupo dos algoritmos de predição, que englobam tanto os algoritmos de classificação quanto os de regressão. No ramo do aprendizado não-supervisionado, os clusterizadores são os representantes mais populares (Figura 7). Cabe ressaltar, que alguns algoritmos podem apresentar um comportamento dual, podendo o mesmo ser utilizado como um classificador ou como um regressor, dependendo de sua parametrização. Caso, por exemplo, do Random Forest (Breiman 2001), um algoritmo AM que parte de métodos de agrupamento (Ensemble Methods), podendo funcionar como regressor tanto quanto classificador.

Por diferentes razões e de diferentes modos, algoritmos supervisionados e não-supervisionados podem ser usados de maneira combinada. Mais adiante neste texto, utilizaremos uma combinação de K-means, U-Net e Random Forest, portanto um clusterizador não-supervisionado e dois regressores/classificadores supervisionados, como forma de; a) alcançar a redução da dimensionalidade do universo amostral inicial, b) alcançar e explorar o domínio do espaço/contexto como alternativa a separação de alvos super similares no domínio

do espectro e do tempo e c) explorar estratégias de separabilidade espectral de coberturas naturais relativamente similares, porém de características biofísicas distintas em algum aspecto.

Por brevidade e praticidade os tópicos que se seguem apresentam o funcionamento básico dos três algoritmos de aprendizado de máquina utilizados no desenvolvimento desta tese, o K-means, o Random Forest e a U-Net.

2.4.1. Clusterização por k-means

Do ponto de vista de sua implementação, o clusterizador K-means (Forgy 1965) é usado para separar n pontos de dados em k clusters. De tal forma que, cada elemento da matriz de entrada seja classificado no cluster que apresente a menor diferença entre o valor médio do cluster e o valor do pixel em questão.

Para um conjunto de dados de entrada $\mathbf{X} = \{x_j, i = 1, \dots, n\}$, os seus centros de cluster $\boldsymbol{\mu} = \{\mu_j, j = 1, \dots, k\}$, são inicialmente escolhidos de maneira aleatória, a partir da matriz de entrada. Então, o seguinte procedimento é repetido até que todos os pixels tenham sido avaliados e a convergência tenha sido alcançada. Para cada pixel x_j da matriz, calcula-se a distância euclidiana entre x_j e cada centro de cluster μ_j , atribuindo-se o pixel da matriz de entrada ao cluster que apresentar o centro mais próximo: $c_i = \underset{j}{\operatorname{argmin}} x_j - \mu_j^2$. Por fim, para cada

cluster, recalcula-se seu valor médio, afim de atualizar o valor central do cluster (Tahmasebi *et*

$$\text{al. 2020): } \mu_j = \frac{\sum_{i=1}^n c_i x_i}{\sum_{i=1}^n c_i}.$$

A clusterização por k-means tende a minimizar a diferença dentro dos clusters e maximizar a diferença entre os clusters. Na Figura 8, temos uma representação gráfica de seu funcionamento.

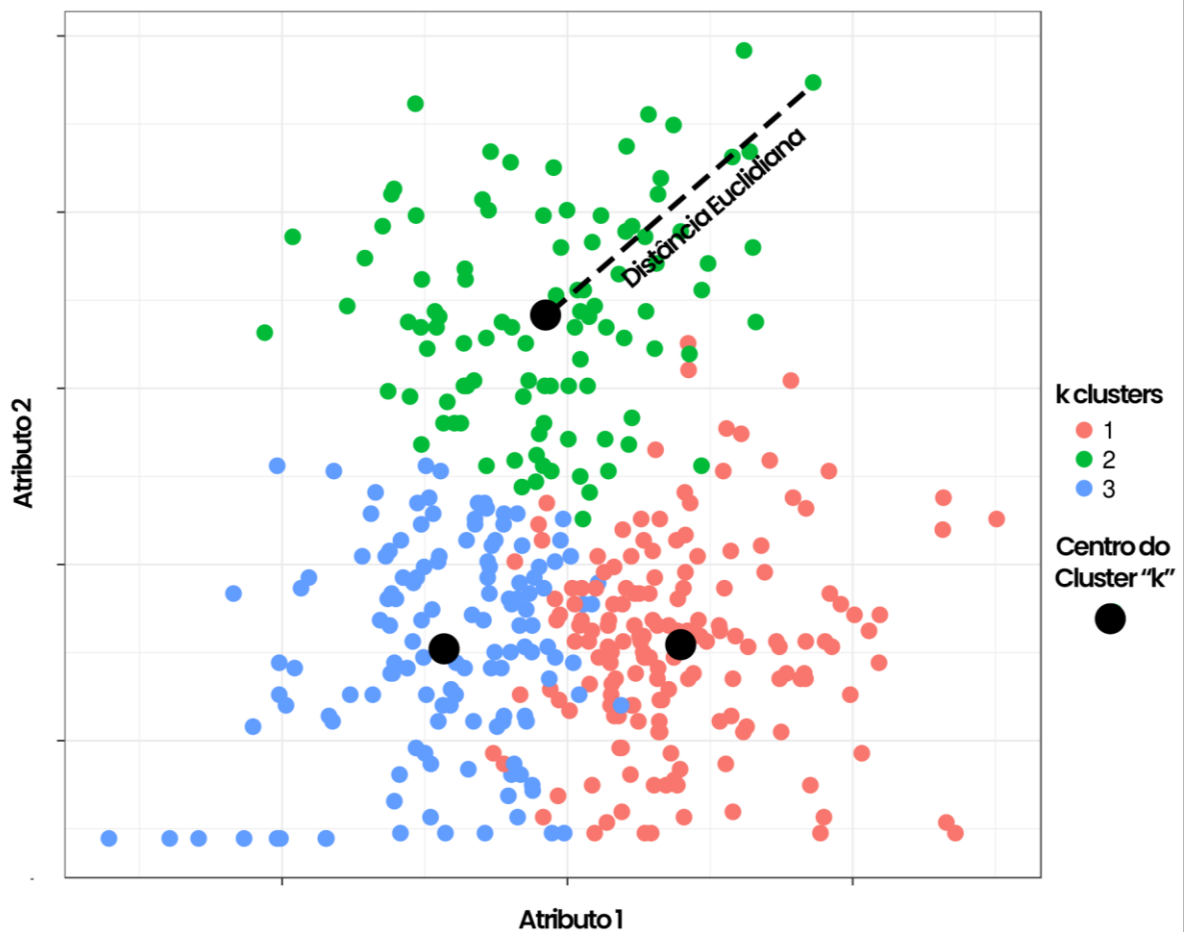


Figura 8 – O algoritmo K-means separa n pontos de dados em k clusters. De tal forma que, cada elemento pixel do dado de entrada seja classificado em um cluster que tenha a menor diferença entre o valor médio do cluster e o valor do pixel em questão.

2.4.2. Random Forest

O Random Forest (RF) é um algoritmo de aprendizado por agrupamento (ensemble), que pode ser usado tanto como um classificador, quanto como regressor. O RF é constituído de árvores de decisão diversas (daí deriva o nome Forest), que usa fundamentalmente de dois conceitos estatísticos para adicionar maior flexibilidade e robustez as decisões de suas árvores; a aleatoriedade e a majoritariedade (Breiman 2001), Figura 9.

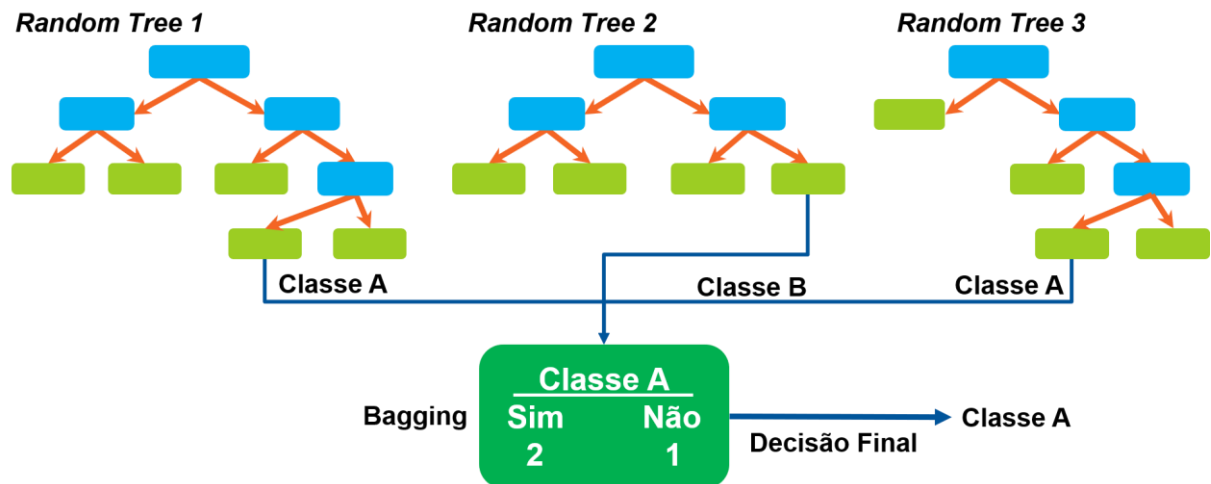


Figura 9 - O RF é constituído de árvores de decisão diversas (daí deriva o nome Forest), que usa fundamentalmente de dois conceitos estatísticos para adicionar maior flexibilidade e robustez as decisões de suas árvores; a aleatoriedade e a majoritariedade.

Estas duas características atuam em momentos diferentes da execução do algoritmo. A aleatoriedade do RF apresenta-se durante dois momentos distintos:

- Durante a construção do “Bootstrap Dataset” (BD), aplica-se a aleatoriedade para a escolha das amostras que integrarão o subconjunto de treinamento. Além da aleatoriedade, a repetição de amostras é característica deste conjunto. A replicação de amostras no “Bootstrap Dataset”, permite a criação do “Out-of-Bag Dataset”. Um conjunto de dados que será usado no cálculo da acurácia do RF, e é constituído das amostras não utilizadas no BD. No RF, cada árvore, tem seu próprio subconjunto aleatório de treinamento e de validação.
- A seleção dos atributos que integrarão cada nó da árvore, é igualmente aleatória. Nesse processo, compara-se ao menos um par de atributos entre si, sendo o atributo a preencher o nó da árvore, aquele que apresentar menor confusão/impureza entre seus valores. A impureza ou confusão de uma distribuição, é dada pelo cálculo da impureza de Gini ou pelo cálculo da Entropia, a depender da implementação de Random Forest que se tenha em mãos.

A majoritariedade é outra característica fundamental do RF. Apesar de baseado em árvore de decisão, o RF não é constituído de uma, porém de n árvores. Assim a decisão final por um rótulo A ou rotulo B, depende da decisão de várias árvores. Vem daí o conceito de Bagging.

- Parametrizado em geral com 100 ou mais árvores, a decisão final por uma classe A ou por um valor contínuo X, depende da decisão agregada de todas as árvores. Rearranjar aleatoriamente os dados (bootstrapping) e utilizar uma agregação (somatório de votos) para a tomada de decisão, é chamado de “Bagging”.
- O RF pode retornar uma classe/rótulo, uma probabilidade, multi-probabilidades, ou o resultado de uma regressão, como sendo a decisão final resultante da análise de todos os votos da rede.

2.4.3. A Rede Neural Convolucionária: U-Net

A U-Net é uma solução genérica de aprendizado profundo (Deep Learning), estruturada como uma rede neural convolucionária (CNN), que foi proposta para solucionar desafios de quantificação, de maneira mais rápida que as alternativas anteriores. (Ronneberger, Fischer e Brox, 2015). A U-Net foi inicialmente projetada para ser usada sobre imagens biomédicas, de maneira a auxiliar em tarefas como detecção, quantificação de células e medições de forma (Falk *et al.* 2019)

É importante termos em mente qual o objetivo central de uma rede neural (NN), qualquer que seja o tipo da rede em questão. Neste sentido as NN, se apresentam como uma maneira genérica de encontrar uma solução viável, de melhor ajuste possível (fit), a qualquer problema de previsibilidade numérica ou classificação categórica, com aplicações que vão do sensoriamento remoto à biomedicina, independentemente do padrão de distribuição numérico ou de sua complexidade estatística, Figura 10.

A U-Net, desenvolveu-se de uma rede neural convolucional tradicional. E, como uma CNN tradicional, o objetivo geral concentrava-se na tarefa de classificação de imagens, onde a entrada (input) é uma imagem e a saída é um rótulo (classe). Porém, em casos biomédicos, é necessário, não apenas confirmar ou negar a doença, mas também permitir localizar a área que apresenta a anormalidade clínica (Falk *et al.* 2019; Ronneberger *et al.* 2015).

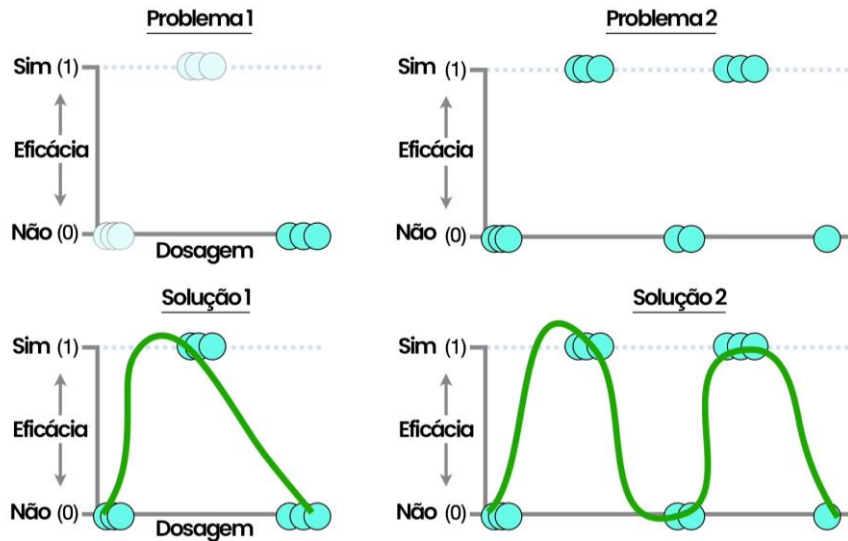


Figura 10 - NN, se apresentam como uma maneira genérica de encontrar uma solução viável, de melhor ajuste possível (fit), a qualquer problema de previsibilidade numérica ou classificação categórica, independentemente do padrão de distribuição numérico ou de sua complexidade estatística.

A U-Net, dedicou-se a resolver o problema de localização da anomalia detectada. A razão pela qual ela é capaz de localizar e distinguir anomalias, vem de dois motivos principais: 1) por fazer a classificação pixel a pixel, ainda que treinada por regiões (patches ou chips), sua entrada e saída precisam compartilhar o mesmo tamanho. 2) uma fase de expansão foi projetada, especificamente, para compensar a degradação da capacidade de se localizar, resultado das múltiplas convoluções da fase de contração.

Assim, uma imagem de entrada de tamanho 2×2 , terá como saída uma imagem de 2×2 . A compensação da fase de contração, com a existência da fase de expansão, dá a rede U-Net, o seu formato característico em U (Figura 11).

À primeira vista, a rede tem a forma de “U”. Essa arquitetura é simétrica e consiste em duas partes principais – o ramo a esquerda é chamado de contração (downsampling), sendo constituído pelo processo de convolução tradicional; a parte direita é o ramo expansivo da rede, que é constituído majoritariamente de convoluções transpostas (upsampling).

Antes de compreender em mais detalhes os momentos de downsampling e upsampling, é essencial que compreendamos o significado de convolução.

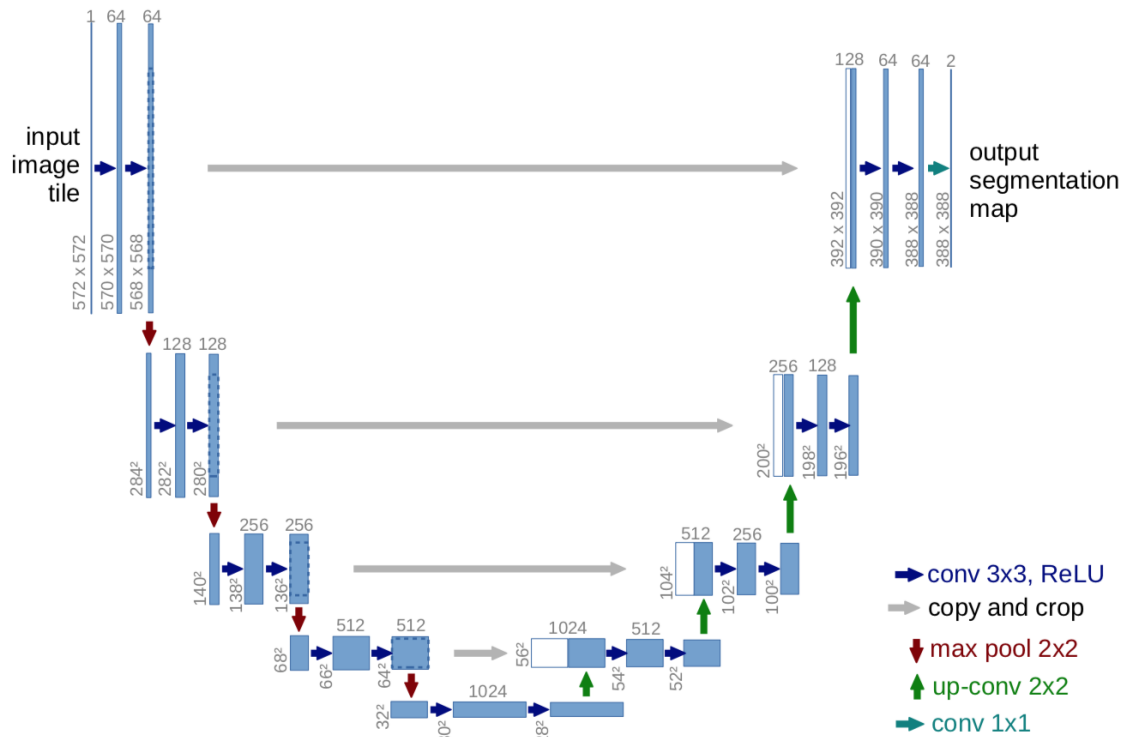


Figura 11 - A rede tem a forma de “U”. Essa arquitetura é simétrica e consiste em duas partes principais – o ramo a esquerda é chamado de contração (downsampling), sendo constituído pelo processo de convolução tradicional; a parte a direita é o ramo expansivo da rede, que é constituído de convoluções transpostas (upsampling).

Uma convolução é uma técnica comum de processamento digital de imagens, em que se altera o valor de um pixel central de acordo com os valores de seus pixels circundantes. As convoluções são aplicadas por janelas moveis (kernels), de tamanhos variados (3x3, 5x5, 7x7 etc), que deslizam sobre todos os pixels da imagem, de maneira ordenada e unidirecional. Muitos filtros matriciais, como o desfoque (blurring), a detecção de bordas, nitidez e relevo, derivam da convolução (Gonzalez & Woods, 2008; Jähne, 2012). As convulsões, portanto, podem contribuir para adicionar pistas texturais e contextuais a determinados alvos, porém, dependendo de sua utilização, podem alterar o tamanho do dado de saída (Figura 12).

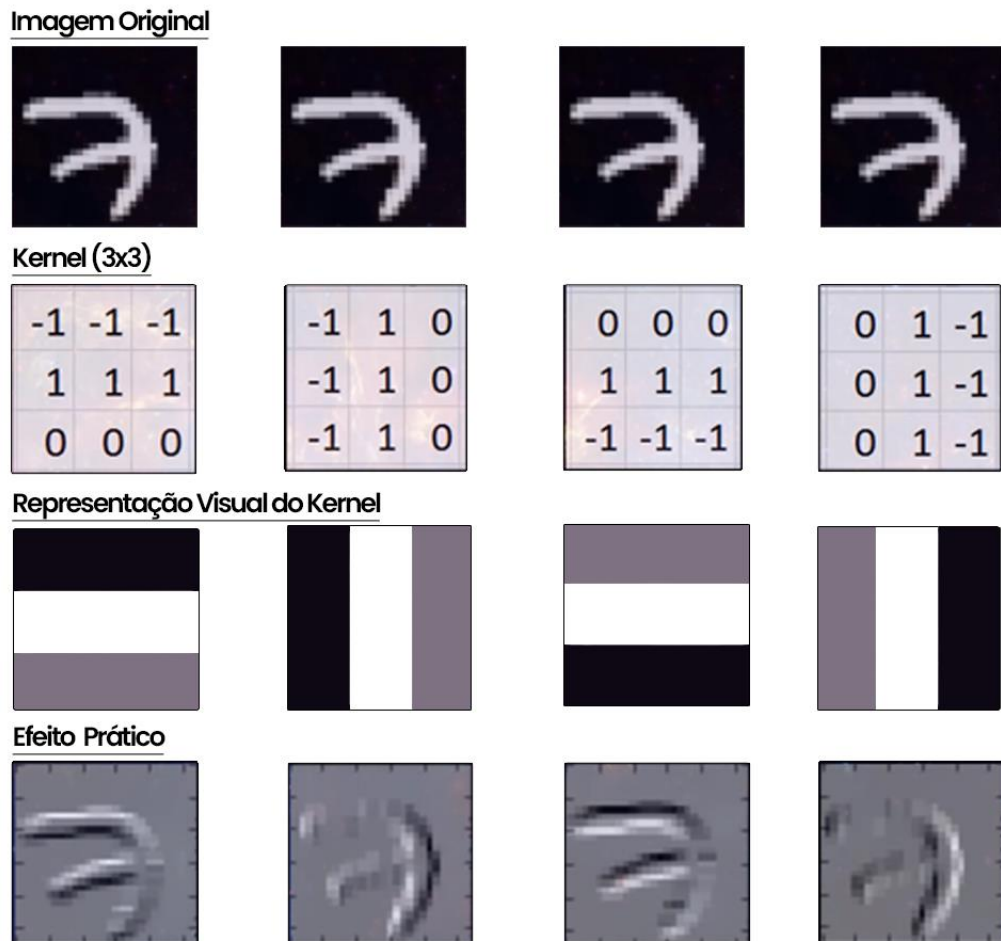


Figura 12 - convolução é uma técnica comum de processamento digital de imagens, em que se altera o valor de um pixel central de acordo com os valores de seus pixels circundantes. São aplicadas por janelas moveis (kernels), de tamanhos variados (3x3, 5x5, etc). As convulsões, portanto, podem contribuir para adicionar pistas texturais e contextuais a determinados alvos.

Agora que temos o conceito de convoluções tradicionais, podemos entender os dois ramos de uma U-Net, o de contração e o de expansão.

As convoluções permitem a inserção de informação textural e contextual a imagem, adicionando como camadas/bandas os resultados de suas filtragens (Figura 12). Entretanto, a cada passagem dos filtros de convolução, a matriz resultante tende a ser menor que a matriz original, ainda que tenha ganhado bandas de informação adicionais. Assim, as múltiplas filtragens/convoluções, aumentam a profundidade do eixo z, ao passo que a dimensão em x e y diminui. É exatamente o que acontece no trecho de downsampling (Figura 13).

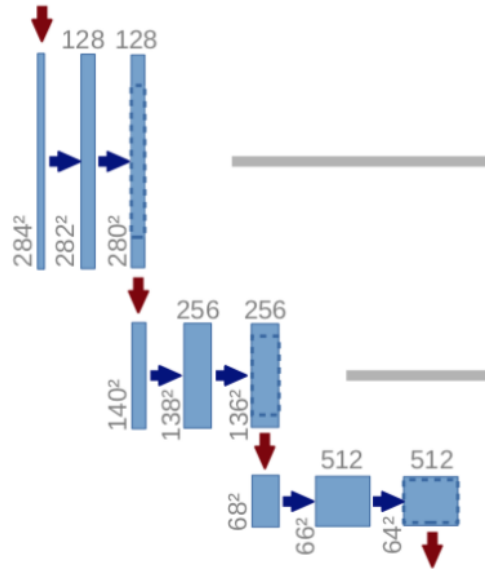
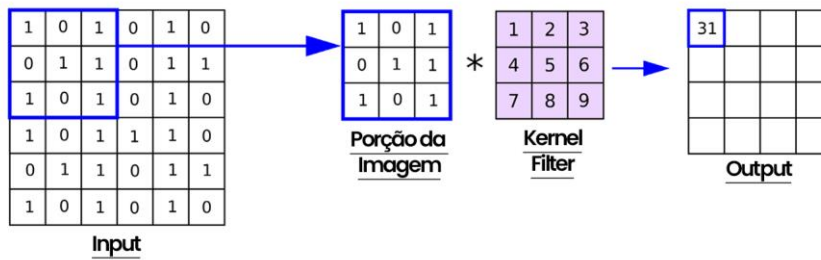


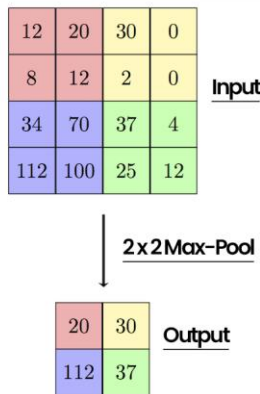
Figura 13 – Redução da dimensionalidade x, y e aumento da profundidade em z. A matriz resultante tem mais camadas de informação texturais e contextuais, porém tamanho menor.

No trecho descendente duas técnicas de convolução distintas são aplicadas. Nas setas horizontais em azul, convoluções tradicionais atuam gerando outputs pouco menores em x e y, resultado da perda da informação de suas bordas. Nas setas verticais, em vermelho, há a ação da convolução por maxpooling, que diminui a dimensão x e y da matriz, à metade de seu tamanho original, através da aplicação de um operador estatístico de máximo valores (max), Figura 14. O objetivo da operação de maxpooling é reduzir a dimensionalidade de entrada (imagem, input, output, hidden layers, etc.) e permitir inferências sobre recursos contidos nas sub-regiões agrupadas (Arruda *et al.* 2021, Falk *et al.* 2019, Parente *et al.* 2019, Saraiva *et al.* 2020).

Convolução Tradicional



MaxPool – Efeito Numérico



MaxPool – Efeito Prático

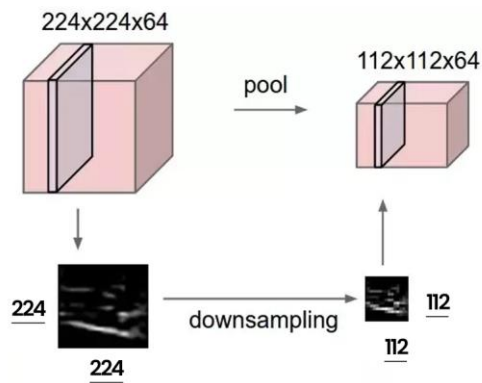


Figura 14 – Fase de contração/downsampling. As convoluções permitem a inserção de informação textural e contextual a imagem. Ao passo que o objetivo da operação de maxpooling é reduzir a dimensionalidade de entrada (imagem, input, output, hidden layers, etc.) e permitir inferências sobre feições contidas nas sub-regiões agrupadas.

Precisamos lembrar que a U-Net deve ser capaz de localizar e distinguir padrões em uma imagem. Convoluída várias vezes, ao final da etapa de downsampling a U-Net perdeu a capacidade de se localizar. Assim, duas outras operações estático-matriciais entram em ação. Na etapa de expansão, a matriz será devolvida ao seu tamanho original, através de uma série de convoluções transpostas e operações de padding, Figura 15.

Basicamente, na etapa de expansão da rede, são feitas operações de padding, seguidas por uma operação de convolução transposta (Gonzalez & Woods 2008, Jähne 2012). A convolução transposta é uma técnica de upsampling que expande o tamanho das imagens. Desta maneira, no começo do ramo ascendente da U-Net (Figura 11), a convolução transposta expande a imagem de entrada de $28 \times 28 \times 1024 \rightarrow 56 \times 56 \times 512$. Em seguida, esta imagem é concatenada com a imagem correspondente do caminho de contração e juntas formam uma imagem de tamanho $56 \times 56 \times 1024$. O motivo da concatenação aqui, é combinar as informações das camadas anteriores para obter uma previsão mais precisa.

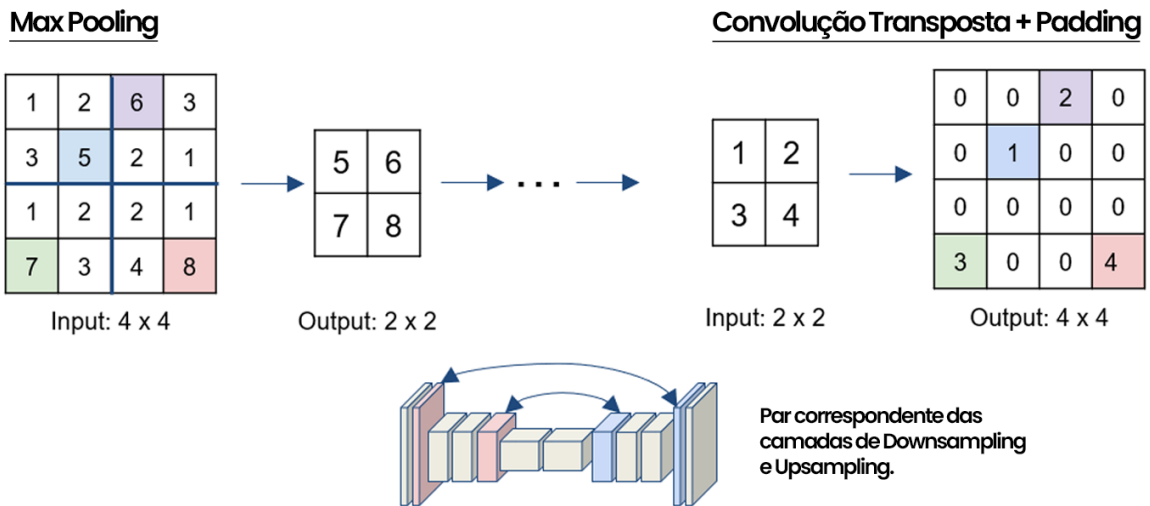


Figura 15 - Convoluída várias vezes, ao final da etapa de downsampling a U-Net perdeu a capacidade de se localizar. Assim, duas outras operações estático-matriciais entram em ação; as convoluções transpostas e as operações de padding.

Agora, alcançado o trecho final da arquitetura U-Net, com camadas de informação textural e contextuais adicionadas, graças a etapa de contração e a aplicação de consecutiva de filtros convolucionais, com o tamanho matricial e a capacidade de localização restauradas, graças a etapa de expansão, suas consecutivas convoluções transpostas e operações de padding, nos resta satisfazer a última etapa e concluir a segmentação semântica, alcançando o objetivo principal de uma CNN, a previsibilidade/classificação matricial, Figura 16.



Figura 16 – Etapa final da U-Net. Já com camadas de informação textural e contextuais adicionadas, tamanho e capacidade de localização restauradas, é concluída a etapa de segmentação semântica, satisfazendo assim o objetivo inicial de previsibilidade/classificação matricial.

2.5. AVALIAÇÃO DE ACURÁCIA

O princípio fundamental da análise de acurácia de um dado temático, passa pela comparação cruzada entre a classificação produzida (dado a ser avaliado), contraposta a uma referência espacial (considerada como verdade). Sendo a referência e o dado a ser avaliado, necessariamente representantes do mesmo espaço, do mesmo período de tempo e do mesmo universo categórico (classes ou temas). Para que então, através da distribuição de pontos amostrais, sorteados seguindo o rigor estatístico de designs amostrais, se consiga comparar categoricamente os dois dados, a classificação produzida e o dado de referência (Foody 2002, 2009, Olofsson *et al.* 2014, Stehman 2014)

Portanto, para existir análise de acurácia, há de existir dado de referência. Em geral, estes dados podem ser obtidos por meio pontos coletados in situ (verdade de campo), imagens de satélite ou fotografia aérea. A escolha por dados satelitários, se tornou cada vez mais comum, em razão do baixo custo de aquisição de imagens de orbitais frente ao custo de deslocamento in campo ou do aluguel de aeronaves (Finegold *et al.* 2016, Olofsson *et al.* 2014).

Neste ponto, os dados orbitais ou fotográficos devem privilegiar, quando possível, o uso de imagens com resolução espacial mais alta que o dado de suporte a classificação. Deve-se ponderar, porém, a história da fotografia aérea e das imagens orbitais, não sendo incomum a inexistência de dados de alta resolução para séries temporais pretéritas e muito longas. Neste caso, assume-se que a análise comparativa deve ser feita sobre a mesma imagem submetida ao processo de classificação.

Por se tratar essencialmente de uma comparação matricial, a tabulação cruzada ou tabela/matriz de contingência, se tornou a forma mais comum de alcançar a análise de acurácia entre dados espaciais. Vale ressaltar, que a tabulação cruzada é ferramenta genérica de análise de dados, e, portanto, seu uso pode estar atrelado a diferentes objetivos analíticos (Finegold *et al.* 2016, Foody 2002, 2009, Olofsson *et al.* 2014, Stehman 2014). Neste sentido, quando usada para calcular a proporção de erros ou acertos entre dois conjuntos de dados, leva o nome de “Matriz de Confusão” ou “Matriz de Erro”. Porém, quando se está avaliando a transição entre classes temáticas ao longo do tempo, ganha o nome de “Matriz de Transição”.

Como qualquer matriz, a de Confusão ou Erro, é disposta em linhas e colunas. As linhas em geral, correspondem a distribuição, classe a classe, dos dados inerentes a classificação, ao passo que as colunas, dispõem os dados de referência. Aqui, vale ressaltar, que a posição ocupada por referência e classificação, seja em linhas ou colunas, é absolutamente arbitrária e, portanto, dependente da vontade e gosto do usuário. No mundo da computação, é comum a disposição

de referências em linhas. Por outro lado, para o mundo do sensoriamento remoto, é mais comum que as referências sejam dispostas nas colunas (Tabela 4).

Tabela 4 – Disposição básica de uma Matriz de Confusão/Erro. Em A, nas linhas (rótulo em azul) estão dispostas as classificações e nas colunas (rótulo em amarelo) os dados de referência. Para cada classe, em linhas ou colunas, há um somatório contabilizado, $\sum A, \sum B, \sum C$. O termo $\sum T$, refere-se a soma agregada do total de todas as classes. A diagonal principal, em verde, representa os valores verdadeiros, concordantes, entre cada classe; VA, VB, VC. Em magenta, a cima da diagonal, os falsos positivos (FP). Em laranja/pastel, a baixo da diagonal, os falsos negativos (FN). Em B, denota-se que a inversão do posicionamento das referências e classificações, em linhas ou colunas, inverte a posição de FN e FP, bem como de todas as métricas de acurácia a eles relacionadas.

A		Dado de Referência			
		A	B	C	Total
Dado Classificado	A	VA	FP	FP	\sum_c^A
	B	FN	VB	FP	\sum_c^B
	C	FN	FN	VC	\sum_c^C
	Total	\sum_r^A	\sum_r^B	\sum_r^C	$\sum T$

B		Dado de Classificado			
		A	B	C	Total
Dado de Referência	A	VA	FN	FN	\sum_r^A
	B	FP	VB	FN	\sum_r^B
	C	FP	FP	VC	\sum_r^C
	Total	\sum_c^A	\sum_c^B	\sum_c^C	$\sum T$

A posição ocupada por determinado elemento em uma Matriz de Confusão, é sempre dependente de como estejam organizados os rótulos de referência e classificação, e carregam consigo o significado estatístico do elemento. Uma matriz de erro, tem essencialmente 4 posições principais: Os valores verdadeiros para cada classe (VA, VB, VC, etc.), sempre posicionados na diagonal principal da matriz. Os falsos positivos (FP), na organização vista na Figura 4A, posicionados acima da diagonal principal. E, por fim, os falsos negativos (FN), localizados sob a diagonal principal. A inversão do posicionamento das referências e classificações, em linhas ou colunas, inverte a posição de FN e FP (Figura 4B), bem como de todas as métricas de acurácia a eles relacionadas.

É absolutamente fundamental que compreendamos os conceitos práticos por trás dos elementos de uma matriz de erro. Assim, temos que o termo FP, refere-se a discordâncias de comissão, ou seja, por adição indevida do rótulo “A” a uma resposta referenciada como “não-A”. Por esta razão, são chamados de “Falsos-Positivos”. Por outro lado, os erros de falsos-negativos (FN), denotam omissão do classificador, portanto, uma classe classificada como “não-A”, é referenciada como pertencente a “A”. Por fim, as respostas

verdadeiras/concordantes são agrupadas na diagonal principal, constituindo acertos ou concordância entre referência e classificação.

Uma vez construída a estrutura fundamental de uma matriz de erros, podemos então extrair as métricas de acurácia. Classicamente, as métricas são relações estatísticas envolvendo as proporções entre erros e acertos, tanto globais, envolvendo a matriz como um todo, quanto os individuais, pertencentes a cada uma das categorias. As métricas mais comumente utilizadas na análise de dados espaciais são:

- Acurácia Global (AG):
- Erro Global (EG):
- Acurácia do Consumidor (AC):
- Acurácia do Produtor (AP):
- Erros de Omissão (EO):
- Erros de Comissão (EC):
- Erro de Quantidade (QD):
- Erro de Alocação (AD):

A acurácia global, é uma medida percentual da quantidade de amostras corretamente classificadas, em relação ao total de amostras disponíveis. O Erro Global (EG) por sua vez, é sua medida complementar e informa erro percentual global da matriz. A formulação de ambos é dada por:

$$AG = \frac{\sum V}{\sum T} * 100 \quad (1)$$

$$EG = 100 - AG \quad (2)$$

Onde,

V , representa os valores corretamente classificados,

T , representa o total de amostras disponíveis, e

AG , acurácia global.

A acurácia do consumidor (AC), ou do usuário, é a medida de acurácia do ponto de vista do usuário/consumidor do mapa. Aqui, a acurácia basicamente diz ao usuário com que frequência a classe no mapa estará realmente presente na no solo ou em sua referência. É também conhecida como medida de confiabilidade (reliability). O erro de comissão (EC) é complemento de AC e nos informa a quantidade de classificações associadas a falsos positivos ou falsos alarmes. A formulação de ambas é dada por:

$$AC = \frac{Vn}{\sum_c n} \quad (3)$$

$$EC = 100 - AC \quad (4)$$

Onde,

V , representa os valores corretamente classificados,

n , representa a classe a ser calculada. Devendo n , ser substituído por cada classe da matriz,

c , é o índice que representa a distribuição dos dados de classificação

EC , refere-se ao Erro de Comissão e

AC , para Acurácia do Consumidor.

A acurácia do produtor (AP) é a acurácia do ponto de vista do criador/produtor do mapa. Esta é a frequência com que as características reais no solo são mostradas corretamente no mapa classificado ou a probabilidade de uma determinada cobertura/uso da terra ser classificada como tal no mapa. A AP é complementada pelo erro de omissão (EO), que nos informa a quantidade percentual de erros de omissões, ou seja, classes que deveriam ter sido detectadas, mas que por alguma razão foram omitidas da classificação. A formulação de AP e EO são dadas por:

$$AP = \frac{Vn}{\sum_r n} \quad (5)$$

$$EO = 100 - AP \quad (6)$$

Onde,

V , representa os valores corretamente classificados

n , representa a classe a ser calculada. Devendo n , ser substituído por cada classe da matriz,

r , é o índice que representa a distribuição dos dados de referência,

EO , refere-se ao Erro de Omissão e

AP , para Acurácia do Produtor.

Mais recentemente as métricas de Erros de Quantidade (Quantity Disagreement) e Erros de Alocação (Allocation Disagreement) se tornaram mais populares. Ambas as métricas fornecem mais pistas para a identificação de eventuais falhas de uma classificação digital, separando a natureza dos discordâncias identificadas em, erros de quantidade e erros de alocação, característica ausente em seu predecessor, o coeficiente de Kappa (Pontius & Millones 2011, Pontius & Santacruz 2014).

Tabela 5 – Métricas de Acurácia. Classicamente, as métricas são relações estatísticas envolvendo as proporções entre erros e acertos, tanto os globais, envolvendo a matriz como um todo, quanto os individuais, de cada uma das

categorias. As métricas mais comumente utilizadas na análise de dados espaciais são; Acurácia do Consumidor (ou do Usuário) - AC, Acurácia do Produtor - AP, Erros de Omissão - EO, Erros de Comissão - EC, Acurácia Global - AG, Erros de Quantidade - QD e Erros de Alocação – AD.

A		Dado de Referência				Acurácia do Usuário		Erros de Comissão
		A	B	C	Total			
Dado Classificado	A	VA	FP	FP	27	$VA / \sum_c A$	78%	22%
	B	FN	VB	FP	37	$VB / \sum_c B$	84%	16%
	C	FN	FN	VC	31	$VC / \sum_c C$	70%	30%
	Total	33	39	23	95	Acurácia Global (AG)	$\frac{\sum V}{\sum T} * 100$	77.9%
Acurácia do Produtor		$VA / \sum_r A$	$VB / \sum_r B$	$VC / \sum_r C$		Erro Global (EG)	$100 - AG$	22.1%
		64%	80%	96%		Erros de Quantidade (QD)	$\frac{\sum Q}{2}$	8.4%
Erros de Omissão		36%	20%	4%		Erros de Alocação (AD)	$EG - QD$	13.7%

3. MATERIAIS E MÉTODOS

3.1. PROCESSAMENTO DE DADOS

Diferentes estratégias foram utilizadas para alcançar os objetivos estabelecidos nesta pesquisa, destacam-se: i) compilação dos mosaicos anuais livres de nuvens ii) definição do domínio de informação mais relevante para cada classe; iii) criação do banco de dados de treinamento de treinamento; iv) aplicação de dois métodos distintos de classificação de imagens orbitais, o Random Forest e U-Net para separação de vegetação de manguezais e parques aquícolas/salineiros; v) avaliação da acurácia do produto gerado. Estes itens serão detalhadamente descritos nos parágrafos subsequentes.

Todo o processamento de dados ocorre dentro da plataforma do Google Earth Engine - GEE, exceto pela classificação do padrão aquícola/salineiro, dependente da biblioteca TensorFlow (Figura 17). Todos os dados matriciais e seus subprodutos são derivados da família de dados “USGS Landsat TOA, Coleção 1, Tier 1”, que incluem somente dados com precisão de terreno nível 1 (L1TP), possuindo comportamento radiométrico bem caracterizado e intercalibração entre os diferentes sensores da série Landsat (Storey *et al.* 2014, Teillet *et al.* 2001, USGS 2015).

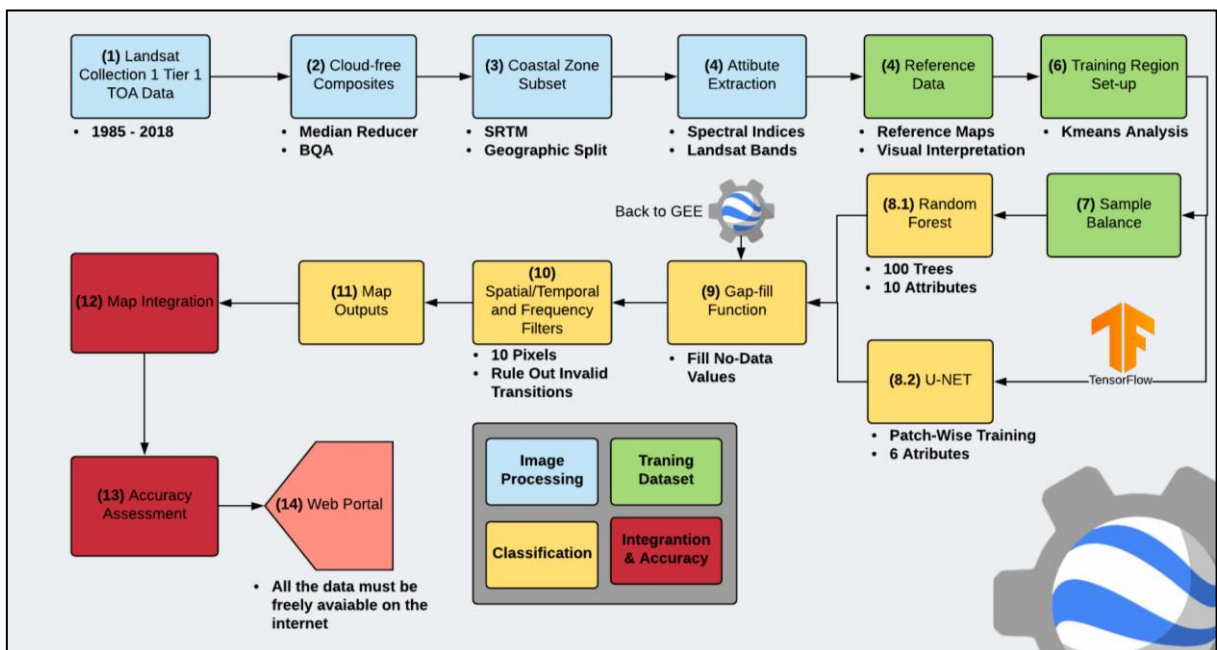


Figura 17 - Todo o processamento de dados ocorre dentro da plataforma do Google Earth Engine - GEE, exceto pela classificação do padrão aquícola/salineiro, dependente da biblioteca TensorFlow. Em verde, etapas relacionadas ao desenho amostral. Em amarelo, etapas relacionadas à classificação. Em vermelho, fase de avaliação de precisão cartográfica. Finalmente, em salmão, fase de publicação dos dados em um “Web Portal”.

Para cada ano, séries de dados Landsat são usados para gerar os “mosaicos anuais livres de nuvens”, variando de 1º de janeiro a 31 de dezembro de cada ano. O script de remoção de nuvens/sombra tira proveito da banda de qualidade “QA”, presente no conjunto de dados “Landsat Collection 1 Tier-1” e do redutor de mediana do GEE. Cada pixel da banda QA, contém valores que representam uma certa condição da superfície terrestre, da atmosfera ou do sensor. Quando apropriadamente utilizados, os valores expressos na banda QA podem melhorar a integridade dos dados, indicando quais pixels podem ser afetados por artefatos do instrumento ou estarem sujeitos à contaminação atmosférica (USGS 2017). Em conjunto, o GEE pode ser instruído a selecionar um pixel que represente os valores medianos de uma pilha de imagens. Ao fazer isso, o GEE rejeita valores muito altos (por exemplo, brilho excessivo de nuvens) ou muito baixos (por exemplo, baixo retorno de sinal, característico de sombras) e seleciona o valor mediano dos pixels em cada uma das bandas ao longo do tempo.

Subsequentemente, os mosaicos anuais são recortados, para incluir apenas áreas de maior probabilidade de ocorrência de manguezais e aquícolas/salineiros (por exemplo, áreas costeiras baixas e zonas intertidais) e excluir grandes áreas onde tipicamente este tipo de cobertura não exista (por exemplo, regiões de elevada topografia, áreas distantes da costa e mar aberto). Esse procedimento, auxilia na redução tanto do tempo de processamento quanto na diminuição de eventuais misturas espectrais, permitindo focar essencialmente na ZCB, diminuindo a diversidade de alvos que possam, eventualmente, comporta-se biofísicamente similar a vegetação de manguezal ou a superfícies aquosas, o que acaba contribuindo para melhorar a precisão global dos mapas finais.

Devido a sua extensão e características biogeográficas (Dominguez 2009, Schaeffer-Novelli *et al.* 1990), a ZCB foi dividida em 6 diferentes setores, conforme demonstrado na Figura 18.

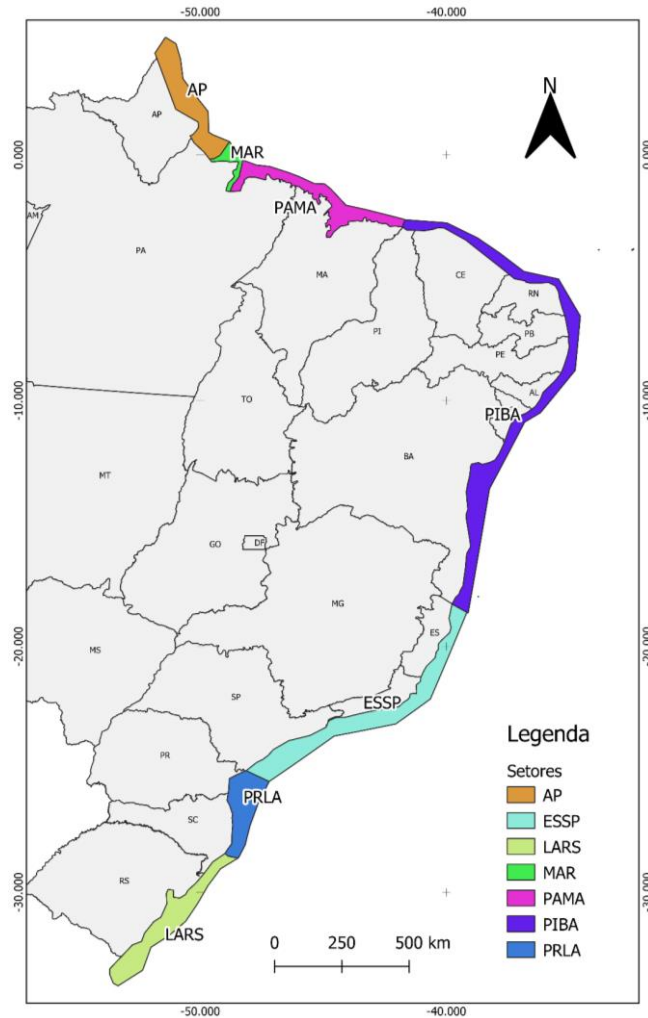


Figura 18 - O litoral brasileiro foi dividido em seis (7) setores distintos. Site 1 - Amapá; Site 2 - Ilha do Marajó; Site 3 - Pará / Maranhão; Site 4 - Piauí / Bahia; Site 5 - Espírito Santo / São Paulo; Site 6 - Paraná / Laguna (SC), Site 7 – Laguna/Rio Grande do Sul. As siglas a seguir representam os estados costeiros brasileiros: AL (Alagoas), AP (Amapá), BA (Bahia), CE (Ceará), ES (Espírito Santo), MA (Maranhão), PA (Pará), PB (Paraíba), PE (Pernambuco), PI (Piauí), PR (Paraná), RJ (Rio de Janeiro), RN (Rio Grande do Norte), SC (Santa Catarina), SE (Sergipe) e SP (São Paulo).

Em seguida, a extração de atributos é iniciada e os seguintes índices espectrais são calculados: o Normalized Vegetation Index (NDVI) (Tucker 1979), o Enhanced Vegetation Index (EVI) (Liu & Huete 1995), o Normalized Difference Water Index (NDWI), o Modified Normalized Difference Water Index (MNDWI) (Xu, 2006), o Normalized Difference Soil Index (NDSI) (Rogers & Kearney 2004) e o proposto MMRI. O MMRI foi inspirado no NDDI, Normalized Difference Drought Index (Gu *et al.* 2007), que é definido por:

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \quad (7)$$

Como os valores do NDVI e do NDWI estão entre -1 e 1, a aplicação direta do NDDI resultaria em um limite matemático indefinido, $-\infty$ a $+\infty$. Assim, o MMRI foi baseado em uma

formulação ligeiramente diferente que usa o MNDWI em vez do NDWI e o módulo de cada índice dentro de uma estrutura de diferença normalizada. Portanto, o MMRI é uma combinação de dois índices clássicos, um índice de vegetação e um índice de água, que realça o contraste da vegetação de manguezal. Sua equação é dada por:

$$MMRI = \frac{|MNDWI| - |NDVI|}{|MNDWI| + |NDVI|} \quad (8)$$

Para aquisição das amostras de treinamento, duas abordagens distintas são utilizadas, uma para coleta amostras de Manguezal (Mg) e Não-Manguezal (N-Mg) e outra para Aqua-Salicultura (AS) e Não-Aqua-Salicultura (N-AS).

De posse dos mosaicos anuais e dos atributos de entrada, o processo de aquisição de amostras é iniciado. O processo é projetado para ser o mais independente possível da aquisição de amostras visualmente interpretadas e, tão importante quanto, para ser um mecanismo escalável no tempo. Tal escalabilidade, deve permitir a aplicação do desenho amostral de treinamento a janelas temporais que possam diferir da característica temporal dos dados de referência e, ainda, permitir a aquisição de amostras de treinamento mesmo quando não houver dados de referência.

Para a aquisição das amostras de Mg e N-Mg, um buffer de ~50 km é aplicado sobre os dados globais de cobertura de manguezal (Giri *et al.* 2011). Então, ao longo de cada mosaico anual, mas apenas dentro da região de redefinida pelo buffer, uma análise de K-means é executada, resultando em uma área de amostragem refinada para as classes de Mg e N-Mg.

Para coleta de amostras de treinamento de AS e N-AS, em razão da inexistência de referências nacionais ou globais, uma máscara de água é inicialmente construída. A máscara é criada de maneira similar a abordagem anterior. Mas desta feita, sobre o Modified Normalized Water Index, é aplicado o K-means. Essa clusterização inicial ajuda a criar uma máscara binária de água, rotulando os pixels como Água ou Não-Água. Guiados pela existência de máscara de água/não-água, as estruturas aquícolas e salineiras são manualmente delineadas. É fundamental destacar que não é feita diferenciação entre aquicultura ou salicultura; ambos os elementos são agrupados como membros da classe de AS. Portanto, a partir deste ponto, toda vez que amostras ou a classe AS for mencionada, está se referindo a um padrão agrupado de aquicultura e salinas.

Em razão dos diferentes domínios de informação a serem acessados para que se consiga distinguir Manguezais e estruturas artificiais de Aqua-Salicultura, dois classificadores distintos são utilizados: 1) o Random Forest, para diferenciação espectral de Mg e N-MG e 2) a U-Net, para diferenciação espacial/contextual de AS e N-AS.

Para a classificação da cobertura de manguezais, sobre a região de treinamento previamente definida, ~1000 amostras aleatoriamente estratificadas foram distribuídas por classe, manguezal e não manguezal, por setor por ano. Depois de coletadas, as amostras foram estatisticamente filtradas (eliminando amostras que extrapolem o percentil de 80%) e visualmente inspecionadas para remover amostras de treinamento inadequadas. A amostragem estratificada e filtragem estatística são necessárias para resolver problemas de desbalanceamento de classes (Bogner *et al.* 2018), além de permitir a remoção outliers do conjunto de amostras de treinamento. A existência de desbalanceamento de classes na região costeira é um cenário provável, uma vez que as amostras N-Mg, especialmente sobre superfícies oceânicas, podem ultrapassar em muito a ocorrência de classes raras (como por exemplo, cobertura de manguezais).

Para classificar costeiras artificiais, foi usado um pipeline Earth Engine-TensorFlow com base no uso da rede neural convolucionária, U-Net. De posse do primeiro conjunto de amostras inicialmente delineadas, a primeira iteração da classificação U-net é executada. Interações diversas podem ocorrer, até que se atinja uma condição de classificação satisfatória. A classificação é então injetada de volta GEE, onde filtros espaço-temporais e a inspeção visual ocorrem. Esta fase é realizada para corrigir dados incorretamente classificados e avaliar a necessidade de adquirir (ou não) mais amostras de treinamento.

A Tabela 6 mostra os atributos dos classificadores Random Forest e U-Net, além dos parâmetros de classificação utilizados em ambos. No total, dez atributos de classificação foram usados para treinar o classificador Random Forest. Enquanto seis atributos foram usados para treinar o modelo U-Net.

Tabela 6 - Atributos dos classificadores Random Forest e U-Net. No total, dez atributos de classificação foram usados para treinar o classificador Random Forest. Seis atributos foram usados para treinar o modelo U-Net.

Parâmetros	Valores	
Classificador	Random Forest	U-Net
Árvores	100	---
Tile-Size	---	256 x 256 pixels
Amostras	~1000 por classe, por setor, por ano	8400 (geometries)
Atributos	10 (Verde, Red, NIR, SWIR1, SWIR2, NDVI, EVI, MNDWI, NDSI, MMRI)	6 (Red, SWIR1, NIR1, MNDWI, NDVI, and NDSI)
Classes	2 (Mg, N-Mg)	2 (AS, N-AS)

3.2. PÓS-CLASSIFICAÇÃO

Devido à natureza do método de classificação, baseado em pixels, e às séries temporais muito longas (mais de 30 anos), uma cadeia de filtros de pós-classificação foi aplicada. A cadeia começa preenchendo possíveis valores de no-data.

Em uma longa série temporal, de regiões severamente afetadas por nuvens, como a ZCB, espera-se que valores de no-data possam existir em algumas das posições resultantes das composições medianas. No filtro de preenchimento (gap-fill), os valores de no-data (“buracos”), não são permitidos e são substituídos pela classificação válida mais próxima no tempo. Se nenhuma posição válida “futura” estiver disponível, o valor de no-data é substituído pela posição anteriormente válida. Até três anos anteriores podem ser usados para preencher posições persistentemente sem dados.

Portanto, as lacunas de no-data só devem existir se um determinado pixel foi classificado permanentemente como no-data em todo o domínio temporal (30+ anos consecutivos). Para rastrear as origens temporais dos pixels, uma máscara de anos foi construída, conforme mostrado na Figura 19.

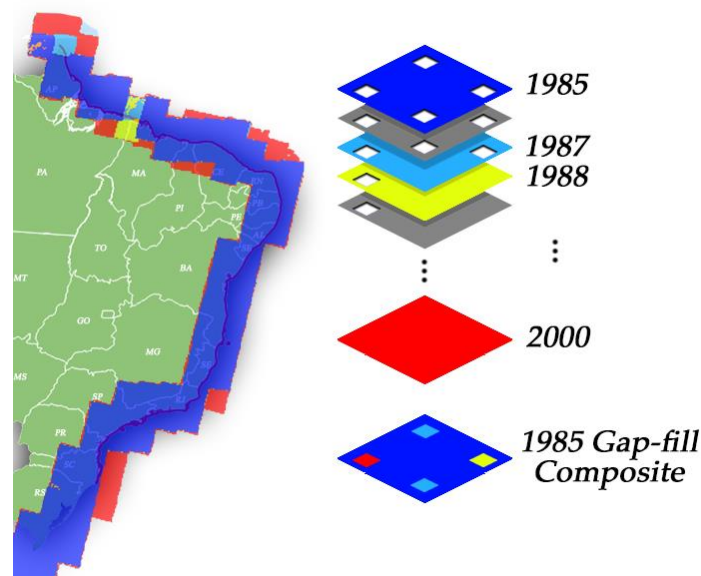


Figura 19 - No filtro de preenchimento (gap-fill), os valores de no-data (“buracos”), não são permitidos e são substituídos pela classificação válida mais próxima no tempo. Se nenhuma posição válida “futura” estiver disponível, o valor de no-data é substituído pela posição anteriormente válida. Até três anos anteriores podem ser usados para preencher posições persistentemente sem dados.

Após o gap-fill, um filtro temporal foi executado. O filtro temporal usa classificações sequenciais em uma janela móvel unidirecional de três anos para identificar transições temporalmente não permitidas. Com base em uma única regra genérica (GR), o filtro temporal inspeciona a posição central de três anos consecutivos (“ternário”), e se as extremidades do

ternário são idênticas, mas a posição central não é, então o pixel central é reclassificado para coincidir com sua classe vizinha temporal, conforme mostrado na Tabela 7.

Tabela 7 – O filtro temporal inspeciona a posição central de três anos consecutivos ("ternário"), e se as extremidades do ternário são idênticas, mas a posição central não é, então o pixel central é reclassificado para coincidir com sua classe temporal vizinha. T1, T2 e T3 significam posição um (1), posição (2) e posição (3). GR significa "regra genérica", enquanto que Mg, N-Mg, AS e N-AS representam as classes de Manguezal, Não-Manguezal, Aqua-Salicultura, Não-Aqua-Salicultura, respectivamente.

Rule	Input (Year)			Output		
	T1	T2	T3	T1	T2	T3
GR	Mg	N-Mg	Mg	Mg	Mg	Mg
GR	N-Mg	Mg	N-Mg	N-Mg	N-Mg	N-Mg
GR	AS	N-AS	AS	AS	AS	AS
GR	N-AS	AS	N-AS	N-AS	N-AS	N-AS

Em seguida, um filtro espacial foi aplicado. Para evitar modificações indesejadas nas bordas dos grupos de pixels (blobs), um filtro espacial foi construído com base na função "connectedPixelCount". Nativa da plataforma GEE, esta função localiza componentes conectados (vizinhos) que compartilham o mesmo valor de pixel. Assim, apenas os pixels que não compartilham conexões com um número predefinido de vizinhos idênticos são considerados isolados, conforme mostrado na Figura 20. Nesse filtro, são necessários pelo menos dez pixels conectados para atingir o valor mínimo de conexão. Conseqüentemente, a unidade mínima de mapeamento é diretamente afetada pelo filtro espacial aplicado e foi definida como 10 pixels (~1 ha).

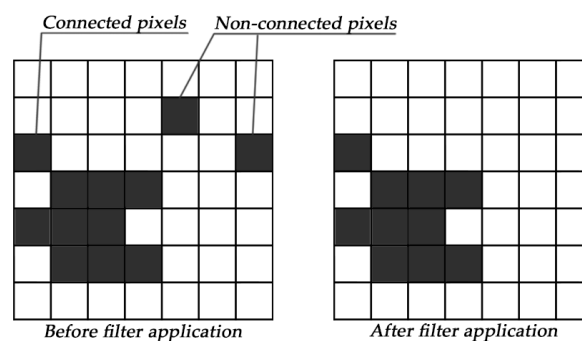


Figura 20 - O filtro espacial remove pixels que não compartilham vizinhos de valor idêntico. O valor mínimo de conexão é de 10 pixels.

A última etapa da cadeia de filtros é o filtro de frequência. Esse filtro é aplicado apenas sobre a classificação de manguezal, conforme mostrado na Figura 21. O filtro de frequência leva em ocorrência da cobertura de manguezais ao longo de toda a série temporal. Assim, todas as ocorrências de com menos de 10% de persistência temporal (3 anos ou menos, em 33) são

filtradas e incorporadas à classe de não-manguezal. Esse mecanismo contribui para reduzir a oscilação temporal do sinal do mangue, diminuindo o número de falsos positivos e preservando os pixels consolidados do mangue.

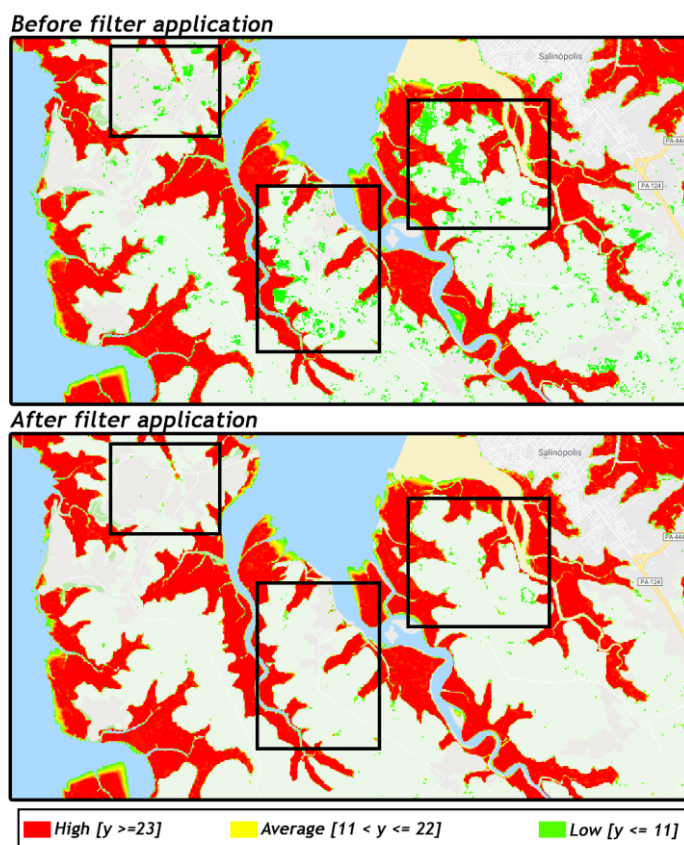


Figura 21 - Vermelho, amarelo e verde representam pixels de manguezal com frequências de ocorrência alta (23 ou mais anos, $y \geq 23$), média (entre 11 e 22 anos, $11 \leq y \leq 22$) e baixa (dez anos ou menos, $y < 11$), respectivamente. A imagem superior mostra os pixels de Mg antes de aplicar o filtro de frequência. A imagem inferior mostra pixels de Mg após a aplicação do filtro de frequência. As caixas pretas são centralizadas nas áreas que foram significativamente afetadas pelo filtro. Observe que todas as ocorrências de Mg com menos de 10% de persistência temporal (3 anos em 33 anos possíveis) foram filtradas.

3.3. DESIGN AMOSTRAL E ANÁLISE DE ACURÁCIA

3.3.1. Manguezal e não-manguezal

Tabelas de contingência (Olofsson *et al.* 2014, Pontius & Santacruz 2014) foram usadas para avaliar dois produtos distintos: 1) a separabilidade espectral do MMRI e 2) o mapa manguezais/não-manguezais da ZCB. A tabela de contingência faz uma tabulação cruzada das classes Mg e N-Mg do mapa de referência e dos produtos classificados; farto em referências nacionais e globais de mapeamentos, a avaliação da cobertura de manguezais adotou uma estratégia mapa-a-mapa (map-to-map).

A partir da tabela de contingência, foram calculadas as seguintes métricas: concordância geral (OA); a proporção de discordância positiva de manguezal (MgPDp), que se refere à proporção de pixels de manguezais exclusivos dos mapas classificados; e a proporção de discordância negativa do manguezal (MgNDp), que é a proporção de pixels de manguezal exclusivos dos mapas de referência.

Para examinar o desempenho da separabilidade de manguezais através do MMRI, foi realizada uma comparação de diferentes índices manguezais e índices de vegetação clássicos. Classificações individualizadas foram construídas, com base em um classificador de atributo único, no qual apenas um dos seguintes índices espectrais foi considerado a cada execução: CMRI (Alsaaidh *et al.* 2013, Gupta *et al.* 2018), NDVI (Tucker 1979), NDWI (McFeeters 1996) ou MMRI. O CMRI (Gupta *et al.* 2018) a diferença NDWI-NDVI (Alsaaidh *et al.* 2013) são tecnicamente iguais, pois ambos são diferenças aritméticas simples entre o NDVI e o NDWI, nas quais a ordem das variáveis muda de “NDVI-NDWI” para “NDWI-NDVI”. Esse rearranjo impõe uma reversão de brilho do histograma, mas não amplifica a separabilidade entre os estratos de manguezal e não-manguezal. Assim, por uma questão de simplicidade, apenas o NDVI (Tucker 1979), NDWI (McFeeters 1996), o CMRI (Gupta *et al.* 2018) e o MMRI foram comparados.

Cada classificação foi treinada com base no mesmo conjunto de amostras, com um total de mil (1000) amostras de treinamento coletadas, sendo quinhentas (500) por classe. A distribuição das amostras de treinamento obedeceu a uma estratégia estratificada, onde os dados de manguezal/não-manguezal do Centro de Dados e Aplicações Socioeconômicos (SEDAC) (Giri *et al.* 2013) serviram como referência para a estratificação. Para avaliar o desempenho de cada índice, em relação a capacidade de distinguir entre pixels de manguezal e não-manguezal e definir o quão distintos os estratos Mg/N-Mg são, uma matriz de confusão, o coeficiente Kappa (KC) e o coeficiente de Bhattacharya (BC) foram calculados.

O conjunto de dados de teste usado para avaliar o desempenho de cada índice é independente do conjunto de dados de treinamento. O conjunto de dados de teste foi projetado com base em uma nova rodada de amostragem aleatoriamente estratificada sobre a mesma referência (Giri *et al.* 2013). Esta análise ocorreu ao longo do manguezal de Caucaia, na região metropolitana de Fortaleza, no Nordeste do Brasil.

O nível de concordância dos mapas em escala nacional foi derivado de três conjuntos de dados distintos: do SEDAC (Giri *et al.* 2013), que serviu como uma referência para o mapeamento do ano 2000; do Global Mangrove Watch Project (GMW) (Bunting *et al.* 2018), fornecendo os dados de referência de 2010; e o ICMBio - Atlas Brasileiro de Manguezais

(ICMBio 2017), que foi utilizado para avaliar os níveis de concordância do mapeamento para o ano 2013. Como antes, todas as amostras usadas para avaliar as métricas de concordância eram totalmente independentes do conjunto de dados de treinamento. Foram utilizadas vinte mil (20.000) amostras por ano, com dez mil amostras em cada classe. Por fim, todos os resultados, desde os produtos auxiliares, aos mapas anuais finais, são disponibilizados ao público em formatos raster e vetoriais em uma plataforma Web.

3.3.2. Aqua-salicultura e não-aqua-salicultura

Sendo inexistentes as referências nacionais ou globais de estruturas aquícolas ou salineiras, a avaliação da acurácia lançou mão da interpretação visual para avaliar a acurácia do mapeamento. A análise de acurácia foi conduzida com base em 900 amostras independentes, tomadas no nível do pixel Landsat, por um par de anos, o de 1985 e de 2019. Um total de 1.800 amostras foram usadas, 900 por ano. Essas amostras foram geradas amostragem estratificada, onde a raridade da classe AS em comparação com a classe N-AS foi levada em consideração.

A distribuição das amostras de validação ocorreu em três diferentes estratos: 1) 450 dentro de uma superfície de AS temporalmente unificadas (as classificações anuais foram superpostas (merge), formando uma única superfície aqua-salicultura para toda série temporal). 2) o estrato previamente superposto foi ampliando através de uma zona de amortecimento (buffer), dentro da qual, 400 amostras foram coletadas e 3) fora dos estratos 1 e 2 foram adquiridas 50 amostras. O tamanho total do universo amostral foi calculado de acordo com a equação abaixo.

$$n = \frac{Nz^2\gamma p(1-p)}{(N-1)e^2 + z^2\gamma p(1-p)} \quad (9)$$

Onde;

n , número de amostras.

N , tamanho da população.

z , é o score da distribuição normal a um dado intervalo de confiança.

γ , é o grau de confiança

p , é a proporção da população a ser estimada.

e , é máxima margem de erro.

As amostras de validação foram distribuídas para permitir a possibilidade de avaliar a ocorrência de falsos positivos e falsos negativos (Figura 22). Um intervalo de confiança de 95%

e um erro padrão máximo de 5% foram assumidos para estabelecer o tamanho da amostra para cada estrato, seguindo a orientação boas práticas para a avaliação da acurácia de dados espaciais (Olofsson *et al.* 2014, Pontius & Santacruz 2014, Stehman 2014).

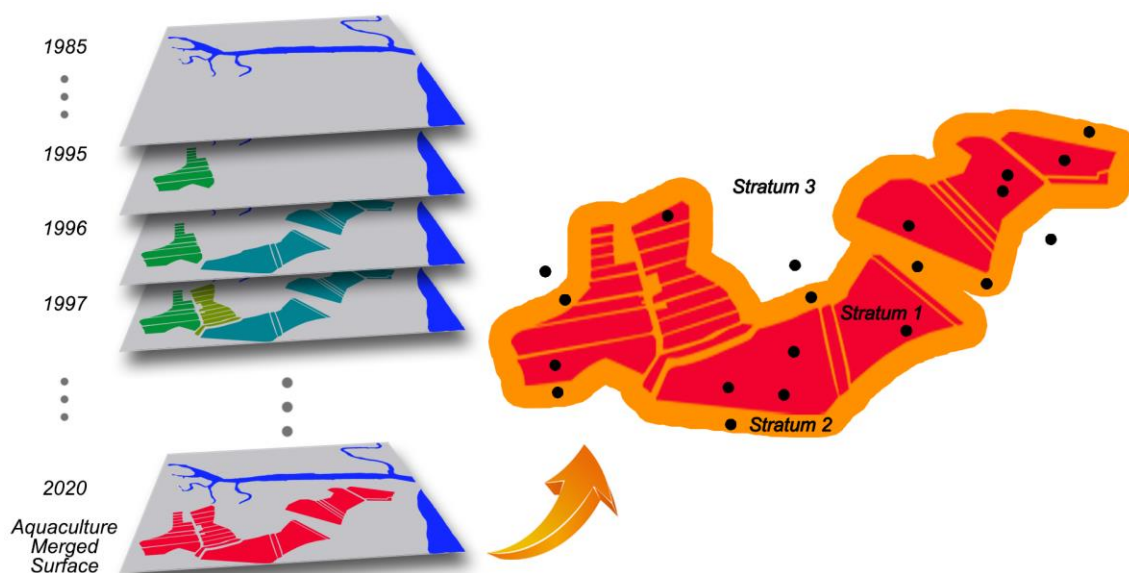


Figura 22 - Um total de 900 amostras independentes foram obtidas no nível do pixel Landsat, em 1985 e 2019, totalizando 1.800 amostras. A amostragem foi dividida em 3 estratos: 1) 450 amostras dentro de uma superfície de temporalmente unificada de AS. 2) 400 em uma região de buffer do estrato 1 e 3) fora dos limites dos estratos 1 e 2 foram adquiridas 50 amostras.

A avaliação da acurácia da classificação de 1985 e 2019 foi calculada com base em uma inspeção visual independente. Esta avaliação visual inspecionou os mesmos mosaicos anuais Landsat fornecidos como entrada para o classificador contextual, U-Net. Como dados de suporte, os intérpretes foram guiados também, por informações multitemporais Landsat, índices de água e imagens de alta resolução do Google Earth Pro.

Por fim, foi produzida uma matriz de erros para cada um dos anos inspecionados. As seguintes métricas de precisão foram calculadas, por classe, por ano: precisão geral (OA), os erros de quantidade e alocação (QD e AD), as acurácias do consumidor e do produtor (CA e PA), os erros de omissão e comissão (OE e CE), e o Coeficiente Kappa (KC) (Olofsson *et al.* 2014, Pontius & Santacruz 2014, Stehman 2014).




3. RESULTADOS

4.1. ARTIGO I - BRAZILIAN MANGROVE STATUS: THREE DECADES OF SATELLITE DATA ANALYSIS.

DOI: <https://doi.org/10.3390/rs11070808/>

Article

Brazilian Mangrove Status: Three Decades of Satellite Data Analysis

Cesar Diniz ^{1,2,*} , Luiz Cortinhas ¹, Gilberto Nerino ¹, Jhonatan Rodrigues ¹, Luís Sadeck ¹, Marcos Adami ^{2,3,4}  and Pedro Walfir M. Souza-Filho ^{2,5} 

¹ Solved—Solutions in Geoinformation, Belém 66075-750, Brazil; luiz.cortinhas@solved.eco.br (L.C.); gilberto.nerino@solved.eco.br (G.N.); jhonatan.rodrigues@solved.eco.br (J.R.); luis.sadeck@solved.eco.br (L.S.)

² Geoscience Institute, Federal University of Pará, Belém 66075-110, Brazil; marcos.adami@inpe.br (M.A.); pedro.martins.souza@itv.org (P.W.M.S.-F.)

³ Regional Center of the Amazon, National Institute for Space Research (INPE), São Paulo 12227-010, Brazil

⁴ Global Land Analysis and Discovery (GLAD) laboratory, Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA

⁵ Instituto Tecnológico Vale (ITV), Belém 66055-090, Brazil

* Correspondence: cesar.diniz@solved.eco.br; Tel.: +55-91-98239-1597

Received: 25 January 2019; Accepted: 22 March 2019; Published: 4 April 2019



Abstract: Since the 1980s, mangrove cover mapping has become a common scientific task. However, the systematic and continuous identification of vegetation cover, whether on a global or regional scale, demands large storage and processing capacities. This manuscript presents a Google Earth Engine (GEE)-managed pipeline to compute the annual status of Brazilian mangroves from 1985 to 2018, along with a new spectral index, the Modular Mangrove Recognition Index (MMRI), which has been specifically designed to better discriminate mangrove forests from the surrounding vegetation. If compared separately, the periods from 1985 to 1998 and 1999 to 2018 show distinct mangrove area trends. The first period, from 1985 to 1998, shows an upward trend, which seems to be related more to the uneven distribution of Landsat data than to a regeneration of Brazilian mangroves. In the second period, from 1999 to 2018, a trend of mangrove area loss was registered, reaching up to 2% of the mangrove forest. On a regional scale, ~85% of Brazil's mangrove cover is in the states of Maranhão, Pará, Amapá and Bahia. In terms of persistence, ~75% of the Brazilian mangroves remained unchanged for two decades or more.

Keywords: mangroves; machine learning; Google Earth Engine; spectral indices; Brazil; Landsat

1. Introduction

Following global change scenarios, coastal areas are exposed to a wide range of environmental hazards, including sea-level rise and its associated effects. At the same time, coastal areas are more densely populated than the hinterland and exhibit higher rates of population growth and urbanisation [1], hosting almost half of the planet's population [2]. Nevertheless, coastal areas comprise only 20% of the Earth's land area [3].

The Brazilian coastal zone (BCZ) shows the same pattern. Extending approximately 9200 km, this dynamic landscape of quick physical and socio-economic changes is home to ~18% of the country's population, along which 16 out of 28 metropolitan regions lie [4]. The Brazilian coastal zone presents a very diverse suite of coastal environments that evolved during the Quaternary in response to changes in climate and sea-level changes, showing an interaction between different sediment supplies and a geologic heritage that dates back to the breakup of South America and Africa [5]. The Brazilian mangrove systems are among this diverse suite of coastal environments.

Globally, mangrove forests are distributed in tropical and subtropical intertidal regions between approximately 30°N and 30°S [6]. In 2000, mangrove forests represented a total area of 137,760 km², distributed in 118 countries and making up ~1% of the tropical forests in the world [7]. Mangrove forests are an evergreen type of vegetation typically distributed from the mean sea level to the highest spring tide [8] and grow in extreme environmental conditions such as extreme tides, high salinity, high temperatures and muddy anaerobic soils [9].

Mangrove systems play an essential role in human sustainability, providing a wide range of ecosystem services, including nutrient cycling, soil formation and wood production. They also provide fish spawning grounds and carbon (C) storage [10–12], being one of the most productive and biologically complex ecosystems on earth [13]. Mangroves and coastal wetlands sequester carbon at an annual rate two to four times greater than that of mature tropical forests and store three to five times more carbon per equivalent area than do tropical forests [10]. Despite its importance, this environment is still highly threatened due to population growth and urbanisation processes.

Since the 1980s, mapping and change detection in mangrove areas at the global scale have been carried out [7,11,14–16]. However, there are few studies in the current literature that include the systematic and continuous identification of mangroves and associated changes, whether on the global or regional scale. In Brazil, the first national mangrove map was published in 1991 [17], based on airborne real aperture radar data collected from 1972 to 1975. At that time, the national mangrove area was ~13,800 km². In the same period, Schaeffer-Novelli et al. [18] described the variability in the mangrove ecosystems along the Brazilian coast. In the last two decades, several papers were published that focused on regional-scale mangrove mapping and short temporal windows, for example, [19–23] concentrated on the northern coast, [24,25] focused on the north-eastern coast and [26,27] focused on the south-eastern coast of Brazil.

In 2010, national-scale mangrove maps, based on 2009 Landsat-5 data, were again published and the Brazilian mangrove area reported was ~11,143 km² [28]. Global-scale mangrove maps published in the last two decades estimated Brazil's mangrove area to be ~9627 km² [7,16]. More recently, the Brazilian Environmental Ministry released the Brazilian Mangrove Atlas, based on 2013 Landsat-8 data, which estimated the national mangrove area to be near 13,989 km² [29]. However, the current literature lacks long time series analyses that allow an exhaustive and systematic understanding of Brazilian mangrove coverage dynamics.

Regardless of the terrestrial cover to be identified, any systematic and exhaustive identification of patterns, including vegetation patterns, requires large storage and processing capacities. These two requirements have only recently been circumvented with the advent of cloud computing platforms, such as Google Earth Engine (GEE) [30] and Amazon Web Services (AWS) [31], combining several petabytes of orbital and geospatial data with statistical analysis resources on the planetary scale.

Moreover, the integration of remotely sensed time series data in such platforms minimizes one of the major problems inherent to land use and land cover mapping: the persistent cloud cover over some areas of the planet [32]. Intertropical coastal zones are no exception to this characteristic. Coastal zones are severely affected by cloudy conditions due to their proximity to oceans and their position. The GEE platform provides fast filtering and sorting capabilities, which greatly facilitates searching through millions of individual images and pixels to select data that meet specific spatial, temporal, spectral or other criteria [30]. However, there are few spectral mechanisms specifically designed to support mangrove identification and to distinguish mangroves from surrounding vegetation. Traditionally, mangrove detection uses classical vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) [33] and the Normalized Difference Water Index (NDWI) [34,35], visual interpretation, supervised classification, unsupervised classification and microwave imagery [7,16,23,28,36–41].

This paper presents the annual Brazilian mangrove cover status as part of a continental-scale analysis from 1985 to 2018 based on Landsat time series data and GEE cloud computing capabilities. To facilitate the recognition, mapping and monitoring of mangrove forests from surrounding vegetation,

this study presents and verifies the robustness of a new spectral vegetation index called the MMRI, the Modular Mangrove Recognition Index.

2. Materials and Methods

Data processing and analysis occurred inside the GEE platform, as shown in Figure 1. All raster data and their sub-products were derived from the United States Geological Survey (USGS) Landsat Collection 1 Tier 1 Top of Atmosphere (TOA) data, which include Level-1 Precision Terrain (L1TP) processed data that have well-characterized radiometric values across the different Landsat sensors [42–44].

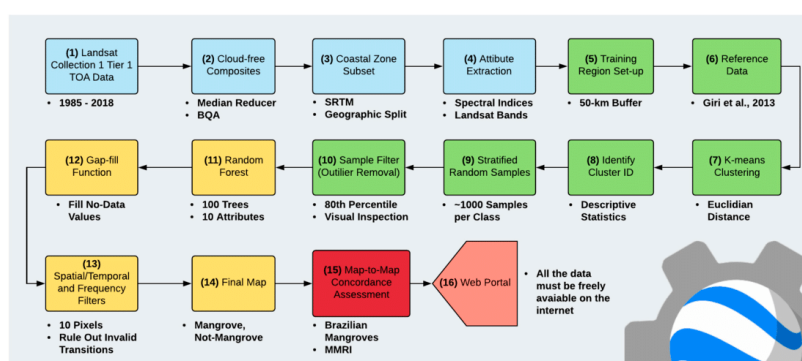


Figure 1. Data-flow diagram. All processing and analysis occur inside the Google Earth Engine (GEE) platform. Steps related to image processing are in blue. The steps in green are related to the sample design. Classification procedures are in yellow. The concordance assessment phase is in red and, finally, the data availability is in salmon. BQA and SRTM denotes Band Quality Assessment and Shuttle Radar Topography Mission, respectively.

For each year, Landsat TOA data were used to produce annual cloud-free composites, ranging from the 1 January to the 31 December. The cloud/shadow removal script takes advantage of the quality assessment (QA) band and the GEE median reducer. When used, QA values can improve data integrity by indicating which pixels might be affected by artefacts or subject to cloud contamination [45]. In conjunction, GEE can be instructed to pick the median pixel value in a stack of images. By doing so, the engine rejects values that are too bright (e.g., clouds) or too dark (e.g., shadows) and picks the median pixel value in each band over time.

Subsequently, the annual mosaics were sub-set to include only areas where mangrove forests are likely to occur (e.g., low-lying coastal areas and intertidal zones) and to exclude vast areas where mangrove forests are not expected to exist (e.g., highlands, areas distant from the shore and open ocean areas). Sub-setting allows the reduction of processing time and decreases the diversity of flooded vegetation types, because non-coastal areas were excluded, thereby improving the overall accuracy of the final maps.

Due to its length and biogeographical characteristics [5,18], the Brazilian coastal region was split into six (6) distinct sites, as shown in Figure 2.

The following spectral indices were calculated from extracted spectral attributes: the Normalized Difference Vegetation Index (NDVI), the Enhanced Vegetation Index (EVI) [46], the Normalized Difference Water Index (NDWI), the Modified Normalized Difference Water Index (MNDWI) [47], the Normalized Difference Soil Index (NDSI) [48] and the proposed MMRI. The MMRI was inspired by the NDDI, the Normalized Difference Drought Index [49], which is given by the following:

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \quad (1)$$

Since the values of both the NDVI and the NDWI are between -1 and 1 , the direct application of the NDDI would result in an undefined mathematical limit, $-\infty$ to $+\infty$. Thus, the MMRI was based on a slightly different formulation that uses the MNDWI instead of the NDWI and the modulus of each index within a normalized difference structure. Therefore, the MMRI is a combination of two classical indices, a vegetation and a water index, which enhances the mangrove cover contrast. Its equation is given by the following:

$$MMRI = \frac{|MNDWI| - |NDVI|}{|MNDWI| + |NDVI|} \quad (2)$$

In the sample acquisition process, the global mangrove cover data [7] was buffered (50 km) and set as the training boundary. Then, over each annual composite but solely inside the narrowed training region, a K-means cluster analysis was run, resulting in a refined sampling area for the mangrove and non-mangrove classes.

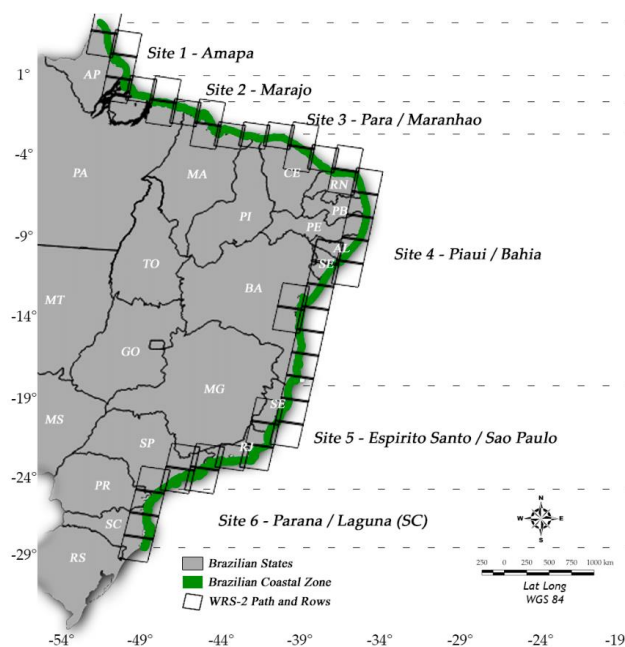


Figure 2. The Brazilian coastal region was split into six (6) distinct sectors. Site 1—Amapá; Site 2—Marajó Island; Site 3—Pará/Maranhão; Site 4—Piauí/Bahia; Site 5—Espírito Santo/São Paulo; and Site 6—Paraná/Laguna (SC). The north and south boundaries of the Brazilian mangroves are in green. The black grids represent the WRS-2 Path and Row footprints. The following acronyms represent the Brazilian coastal states: AL (Alagoas), AP (Amapá), BA (Bahia), CE (Ceará), ES (Espírito Santo), MA (Maranhão), PA (Pará), PB (Paraíba), PE (Pernambuco), PI (Piauí), PR (Paraná), RJ (Rio de Janeiro), RN (Rio Grande do Norte), SC (Santa Catarina), SE (Sergipe) and SP (São Paulo).

Having set the refined training region, ~1000 stratified random samples were distributed per class, mangrove and non-mangrove, per sector per year. Once collected, the samples were statistically filtered (with the 80th percentile function) and visually inspected to remove inadequate training samples.

Stratified sampling and statistical filtering are necessary to address imbalanced class problems [50], allowing the removal of outliers from the sample bag. The presence of imbalanced classes within the coastal region is a probable scenario, as water surface samples may, in general, greatly surpass other rare class occurrences (e.g., mangrove cover).

Among supervised and unsupervised methods, GEE presents more than 15 different classifiers. However, in the last 5 years, nearly 15,000 papers based on random forests (RFs) classifying a variety of land use/land cover classes were produced. Thus, due to its apparent robustness, the RF algorithm was selected here as the classification method to categorize the Brazilian coastal zone into two distinct classes: mangrove (Mg) and non-mangrove (N-Mg). This entire process was then repeated for each year from 1985 to 2018. Table 1 shows the classifier attributes and classification parameters. In total, ten classification attributes were used.

Table 1. Classifier attributes and classification parameters. In total, ten (10) distinct attributes were used. NIR, stands for near-infrared; SWIR1 and SWIR2, short-wave infrareds 1 and 2.

Parameters	Values
Classifier	Random Forest
Trees	100
Samples	~1000 per class per sector per year
Attributes	10 (Green, Red, NIR, SWIR1, SWIR2, NDVI, EVI, MNDWI, NDSI, MMRI)
Classes	2 (Mangrove and Non-Mangrove)

Due to the pixel-based nature of the classification method and the very long temporal series, a chain of post-classification filters was applied. The chain starts by filling in possible no-data values. In a long time series of severely cloud-affected regions, such as tropical coastal zones, it is expected that no-data values may populate some of the resultant median composite pixels. In this filter, no-data values (“gaps”) are theoretically not allowed and are replaced by the temporally nearest valid classification. In this procedure, if no “future” valid position is available, then the no-data value is replaced by its previous valid class. Up to three prior years can be used to fill in persistent no-data positions. Therefore, gaps should only exist if a given pixel has been permanently classified as no-data throughout the entire temporal domain. To keep track of pixel temporal origins, a mask of years was built, as shown in Figure 3.

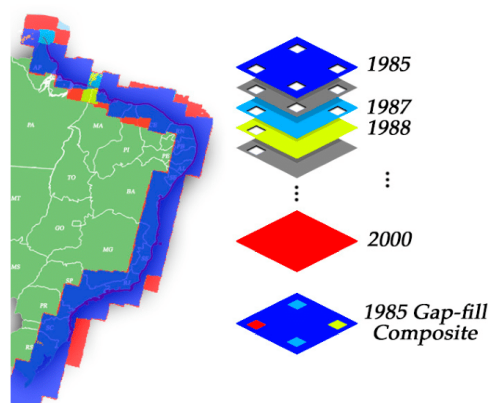


Figure 3. Gap-filling mechanism. The next valid classification replaces existing no-data values. If no “future” valid position is available, then the no-data value is replaced by its previous valid classification, based on up to a maximum of three (3) prior years. To keep track of pixel temporal origins, a mask of years was built.

After gap filling, a temporal filter was executed. The temporal filter uses sequential classifications in a three-year unidirectional moving window to identify temporally non-permitted transitions. Based on a single generic rule (GR), the temporal filter inspects the central position of three consecutive years (“ternary”), and if the extremities of the ternary are identical but the centre position is not, then the central pixel is reclassified to match its temporal neighbour class, as shown in Table 2.

Table 2. The temporal filter inspects the central position of three consecutive years, and in cases of identical extremities, the centre position is reclassified to match its neighbour. T1, T2 and T3 stand for positions one (1), two (2) and three (3), respectively. GR means “generic rule”, while Mg and N-Mg represent mangrove and non-mangrove pixels.

Rule	Input (Year)			Output		
	T1	T2	T3	T1	T2	T3
GR	Mg	N-Mg	Mg	Mg	Mg	Mg
GR	N-Mg	Mg	N-Mg	N-Mg	N-Mg	N-Mg

Next, a spatial filter was applied. To avoid unwanted modifications to the edges of the pixel groups (blobs), a spatial filter was built based on the “connectedPixelCount” function. Native to the GEE platform, this function locates connected components (neighbours) that share the same pixel value. Thus, only pixels that do not share connections to a predefined number of identical neighbours are considered isolated, as shown in Figure 4. In this filter, at least ten connected pixels are needed to reach the minimum connection value. Consequently, the minimum mapping unit is directly affected by the spatial filter applied, and it was defined as 10 pixels (~1 ha).

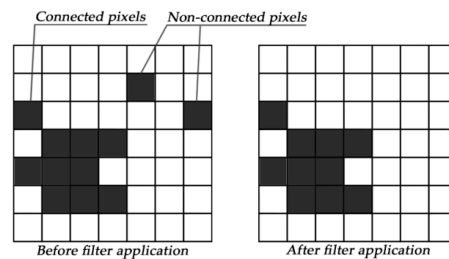


Figure 4. The spatial filter removes pixels that do not share neighbours of identical value. The minimum connection value was 10 pixels.

The last step of the filter chain is the frequency filter, as shown in Figure 5. This filter takes into consideration the mangrove occurrence frequency throughout the entire time series. Thus, all mangrove occurrences with less than 10% temporal persistence (3 years or fewer out of 33) are filtered out and incorporated into the non-mangrove class. This mechanism contributes to reducing the temporal oscillation of the mangrove signal, decreasing the number of false positives and preserving consolidated mangrove pixels.

Finally, contingency table analyses [51,52] were used to compare two distinct products: (1) the MMRI spectral separability and (2) the Brazilian mangrove/non-mangrove map. The contingency table cross-tabulates the Mg and N-Mg classes of the reference map and the classified products; thus, a map-to-map analysis was performed. From the contingency table, the following metrics were calculated: the overall agreement (OA); the mangrove positive disagreement proportion (MgPDp), which refers to the proportion of mangrove pixels that are exclusive to the classified maps; and the mangrove negative disagreement proportion (MgNDp), which is the proportion of mangrove pixels exclusive to the reference map.

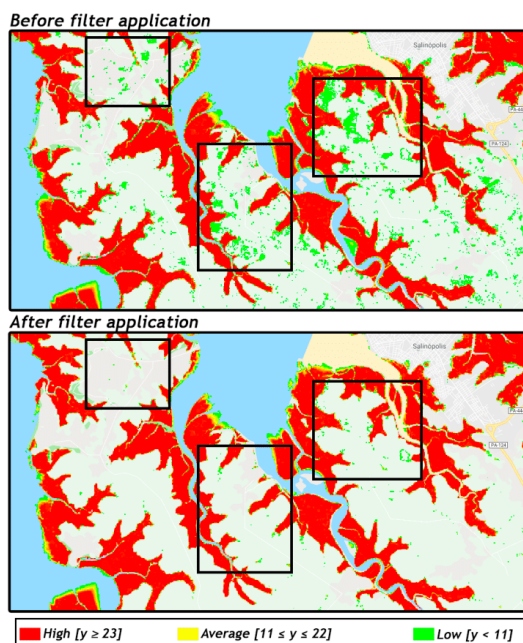


Figure 5. Red, yellow and green represent mangrove pixels with high (23 or more years, $y \geq 23$), average (between 11 and 22 years, $11 \leq y \leq 22$) and low (ten years or less, $y < 11$) occurrence frequencies, respectively. The top image shows mangrove pixels before applying the frequency filter. The bottom image shows mangrove pixels after applying the frequency filter. The black boxes are centred on areas that have been significantly affected by the filter. Note that all mangrove occurrences with less than 10% temporal persistence (3 years in 33 possible years) were filtered out.

To examine the performance of the MMRI mangrove separability, a comparison of distinct mangrove and classical vegetation indices was performed. Each map was constructed based on a single-attribute classifier, in which only one of the following spectral indices per execution were considered: Combined Mangrove Recognition Index (CMRI) [38,53], NDVI [33], NDWI [34] or MMRI. The CMRI [53] and the NDWI-NDVI difference [38] are technically the same, as both are simple arithmetical differences between the NDVI and the NDWI, in which the order of the variables changes from “ $NDVI - NDWI$ ” to “ $NDWI - NDVI$ ”. This rearrangement imposes a histogram brightness reversal but does not amplify the separability between the mangrove and non-mangrove strata. Thus, for the sake of simplicity, only the NDVI [33], the NDWI [34], the CMRI [53] and the MMRI were further compared.

Each classification was trained based on the same set of samples, with a total of one thousand (1000) training samples collected, five hundred (500) per class. The training sample distribution obeyed a stratified strategy, where the mangrove/non-mangrove data obtained from the Socioeconomic Data and Applications Center (SEDAC) [54] served as stratifying reference. To assess the performance of each index in distinguishing between mangrove and non-mangrove pixels and defining how distinct are the mangrove/non-mangrove strata, a contingency table analysis was performed and the kappa coefficient (KC) and the Bhattacharya coefficient (BC) were computed.

The testing dataset used to evaluate the performance of each index was independent of the training dataset. The testing dataset was designed based on a new round of stratified random sampling

over the same reference [54]. This analysis took place along the Caucaia mangrove patch in the metropolitan region of Fortaleza in the Northeast Region of Brazil.

The agreement level of the national-scale maps was derived from three distinct dataset sources: the SEDAC [54] dataset, serving as a reference for the year 2000 mapping; the Global Mangrove Watch (GMW) Project [16], providing the 2010 reference data; and the Chico Mendes Institute for Biodiversity Conservation (ICMbio)—Brazilian Atlas of Mangroves [29], which was used to evaluate the 2013 mapping agreement. As before, all samples used to assess the national agreement measures were entirely independent of the training dataset. A total of twenty thousand (20,000) samples per year were used, with ten thousand (10,000) samples in each class. Finally, all the results, from the auxiliary products to the final annual maps, were made publicly available in raster and vector formats on a Web platform: www.solved.eco.br/mangroveplatform.

3. Results

3.1. Modular Mangrove Recognition Index (MMRI)

Spectral indices may be described as the mathematical combinations of bands used to indicate the relative abundance of features of interest or to enhance the spectral separability between them. With this in mind, the MMRI was compared to already existing mangrove indices and classical vegetation indices, such as the CMRI [53], the NDVI [33] and the NDWI [34].

Based on a set of modular operations and to present a normalized difference structure (normalized ratio), the MMRI imposes an enhancement of the mangrove/non-mangrove contrast level while simultaneously redefining the resultant mathematical limits to fit the usual -1 to 1 interval. Figure 6 shows the visual aspect of each index, as well as a contingency table from which the overall agreement, the mangrove positive disagreement proportion, the mangrove negative disagreement proportion and the kappa coefficient were calculated, as well as the Bhattacharya coefficient, which denotes the spectral distance between mangrove and non-mangrove pixels.

Regardless of the aspect to be considered—whether visual characteristics, contingency table attributes or the spectral distinction between mangrove and non-mangrove classes—the classification based exclusively on the MMRI achieved more robust results than did those based solely on NDVI, NDWI or the difference between these two. All the metrics, the overall agreement, the mangrove positive disagreement proportion, the mangrove negative disagreement proportion and the kappa coefficient were better or equivalent in the MMRI classification (Figure 6).

Regarding the agreement metrics, the OA and the KC, the values reported by the MMRI mapping were ~ 1.5 and ~ 4 times higher, respectively, than the values reported by the remaining indices. Likewise, the disagreement metrics were ~ 4 times lower in the MMRI results, which indicates a more significant agreement between the MMRI classification and the reference used.

From the perspective of mangrove and non-mangrove spectral separability, the Bhattacharya coefficient applied to the MMRI reached the value of 0.22, which is the lowest among all the calculated coefficients (Figure 6). This suggests that the referred index provides greater distinction between the mangrove and non-mangrove strata compared to the other evaluated indices. The BC is a spectral distance metric that measures the overlap between two populations or samples, where a value of zero represents no overlap and a value of one represents perfect overlap [55].

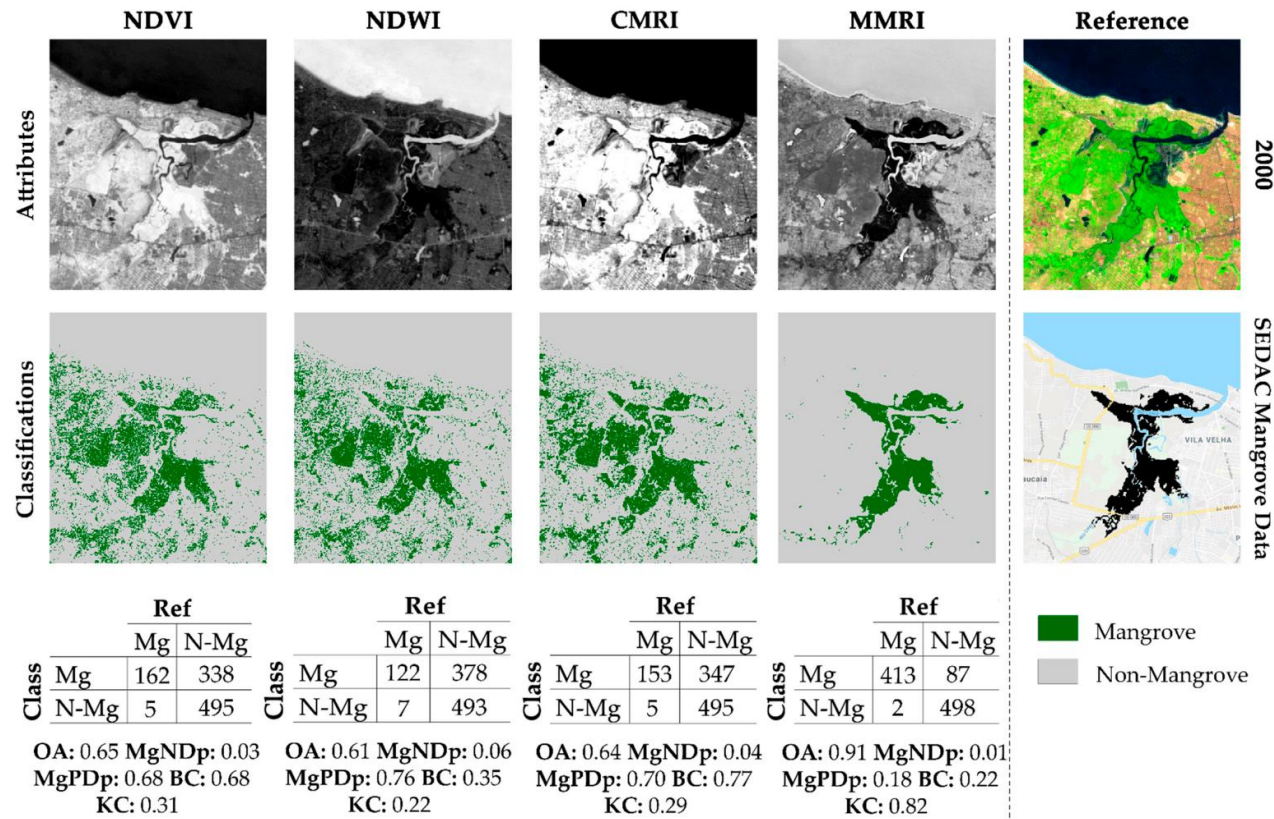


Figure 6. The first row shows the visual aspects of each index—NDVI, NDWI, CMRI and MMRI. The second row shows the classification results. The bottom row shows the contingency tables, where Mg stands for the mangrove class and N-Mg stands for the non-mangrove class. OA stands for overall agreement, MgPDp indicates mangrove positive disagreement proportion, MgNDp refers to mangrove negative disagreement proportion, KC means the kappa coefficient and BC denotes the Bhattacharya coefficient.

3.2. Brazilian Mangrove Status

For the first time in the scientific literature, Brazilian mangrove cover mapping was systematically and continuously carried out, producing maps and annual statistics ranging from 1985 to 2018, as shown in Figure 7. Data of this nature allow a better understanding of the dynamics of Brazilian mangroves, updating the status of this type of coverage throughout the country over a span of 33 years.

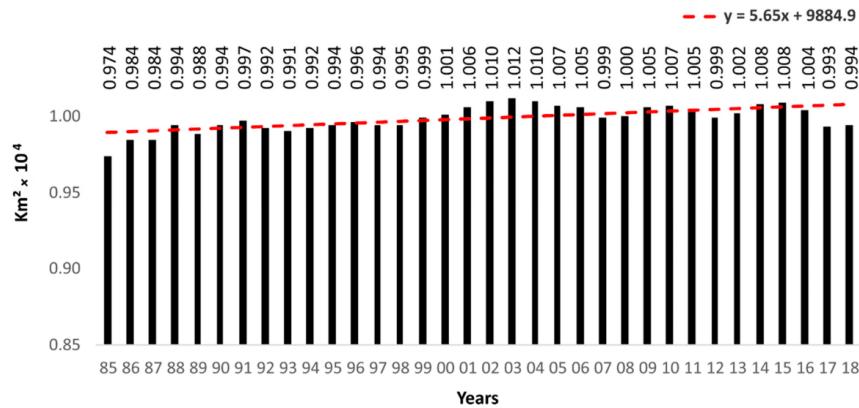


Figure 7. Brazilian mangrove areas from 1985 to 2018. The *x*-axis represents the years of the study. The *y*-axis represents total area of Brazilian mangroves in ten thousand kilometers squared. The numbers above each bar are the area values for that bar.

In Brazil, mangroves occupied areas of ~9700 km² in 1985 and ~9900 km² in 2018, representing an area increase of ~200 km². It is imperative, however, not to rush the interpretation that Brazilian mangroves have experienced a regeneration movement over the past three decades. The percentage difference between 1985 and 2018 is ~2%.

Compared to previous mappings [16,29,54], which refer to the years 2000, 2010 and 2013, respectively, the areas reported here are 4% more, 7% more and 28% less, respectively. In absolute terms, the difference for the year 2000 was ~400 km², with the value reported by the reference reaching ~9600 km² and the value calculated here was ~10,010 km². For the year 2010, the difference was 800 km², with the reference indicating 10,900 km², and this manuscript yielding 10,100 km². Finally, for 2013, the values differed by ~3900 km², with 13,989 km² according to the reference and ~10,020 km² calculated by us.

The overall pattern shows an upward trend (Figures 7 and 8). However, from 1985 to March 1998, only the Landsat-5 satellite remained operational. In this period, for the BCZ, the average number of images per year was ~500. In the last decade between 2008 and 2018, this figure tripled to ~1500 images per year, as shown in Figure 8.

On a regional scale, the states of Maranhão, Pará, Amapá and Bahia are the federal units with the most extensive mangrove cover in the country, reaching ~4350 km² (~46% of the national coverage), ~2100 km² (~22%), ~830 km² (~9%) and ~670 km² (~7%), respectively, in 2018. Together, the four states represent ~85% of Brazil's mangrove cover, as shown in Figure 9.

country, Espírito Santo, Rio de Janeiro, São Paulo and Santa Catarina, which all present levels of stability between 60% and 65%, as shown in Figure 10.

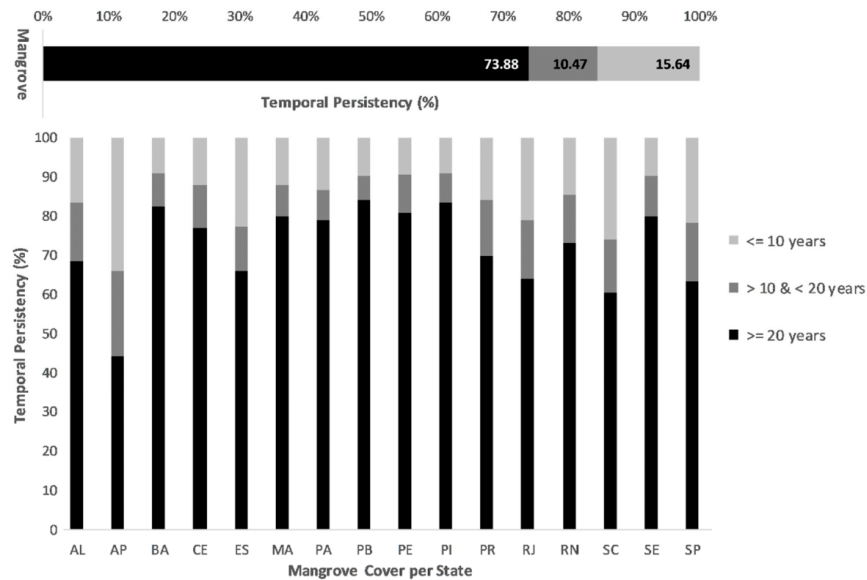


Figure 10. Brazilian mangrove cover persistence at the national and regional scale. The top bar shows the overall mangrove persistence. The bottom graph shows the mangrove persistence per state, abbreviated as in Figure 2. The x-axis represents the state distributions, whereas the y-axis represents the mangrove cover temporal persistence percentages (%). Black represents 20 years or more of stability, dark grey indicates stability between 10 and 20 years and light grey represents stability for less than 10 years.

Since the 1980s, mangrove cover mapping at diverse scales has been carried out, although in a non-systematic way. Here, three sets of data—two global datasets [16,54] from 2000 and 2010, respectively, and one national dataset [29] from 2013—were used to test the concordance levels between the mapping presented in this article and the references cited above. Figure 11 below shows the contingency tables for the years 2000, 2010 and 2013.

Compared to the year 2000 reference data [54], the mapping developed herein achieves an overall agreement of 87% (OA = 0.87) and a kappa coefficient of 74% (KC = 0.74), and presents a low proportion of mangrove negative disagreement, MgNDp = 2% (0.02). There are, however, a large number of positive disagreements, namely, 2474 pixels or ~25% (0.25).

In relation to the 2010 reference [16], the mapping developed herein also achieves a great level of overall agreement, with OA = 85% (0.85) followed by a kappa coefficient of 70%, KC = 0.70. The MgNDp reached ~1% (0.01). On the other hand, the proportion of positive disagreements was ~29% (0.29).

For the year 2013, based on the data produced by the Chico Mendes Institute for Biodiversity Conservation (ICMBio) [29], the smallest kappa coefficient and overall agreement were achieved, KC = 60% (0.60) and OA = 80% (0.80), and the highest negative disagreement proportion was reached, MgNDp = 40% (0.40). The proportion of positive disagreement was low, MgPDp = 1% (0.01).

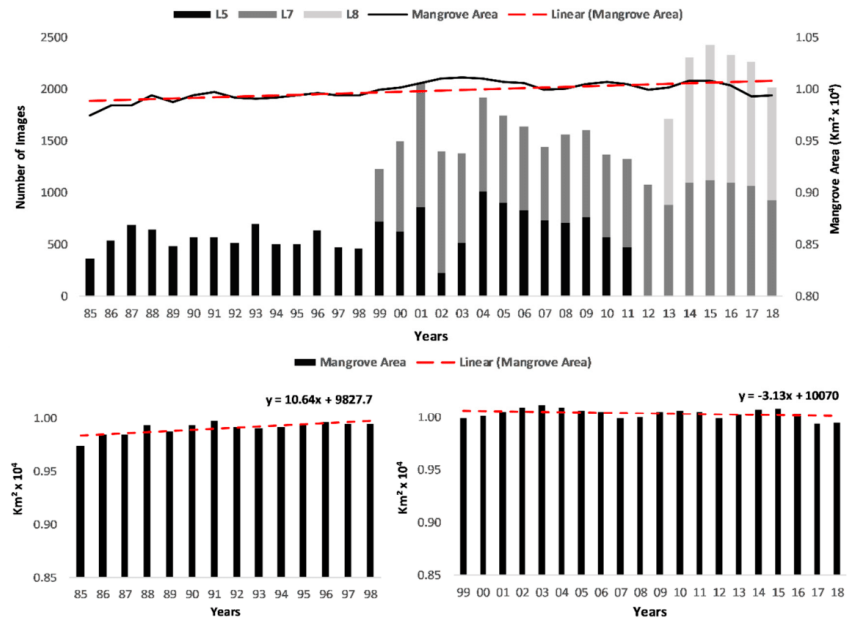


Figure 8. The top image shows the general pattern of fluctuation in the mangrove areas from 1985 to 2018. The solid line represents the variation in mangrove area over time. The dotted line shows the general trend. The bars show the distribution of Landsat images along the time series. L5 stands for Landsat-5, L7 refers to Landsat-7 and L8 stands for Landsat-8. The bottom left image shows the mangrove area variation and trend line from 1985 to 1998. The bottom right image shows the mangrove areas and trends from 1999 to 2018.

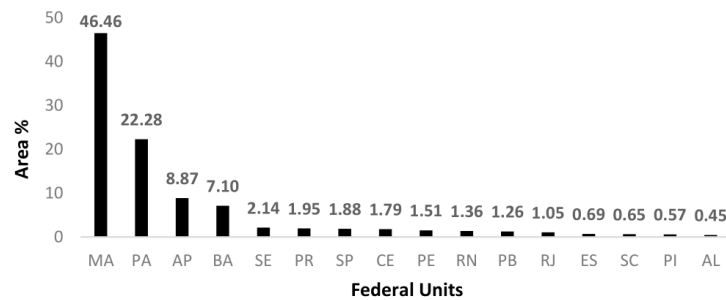


Figure 9. Brazilian mangrove area in percentage of total mangrove cover per state. The x-axis represents the coastal states’ two-letter abbreviations, as in Figure 2. The y-axis represents the mangrove area % of the year 2018.

From the perspective of mangrove cover persistence, ~75% of the Brazilian mangroves remained unchanged for two decades or more, ~10% remained stable between one and two decades and ~15% remained stable for ten or fewer years. In this scenario, the state of Amapá is the state that shows the lowest stability, at ~40%, followed by the states of the south-eastern and southern regions of the

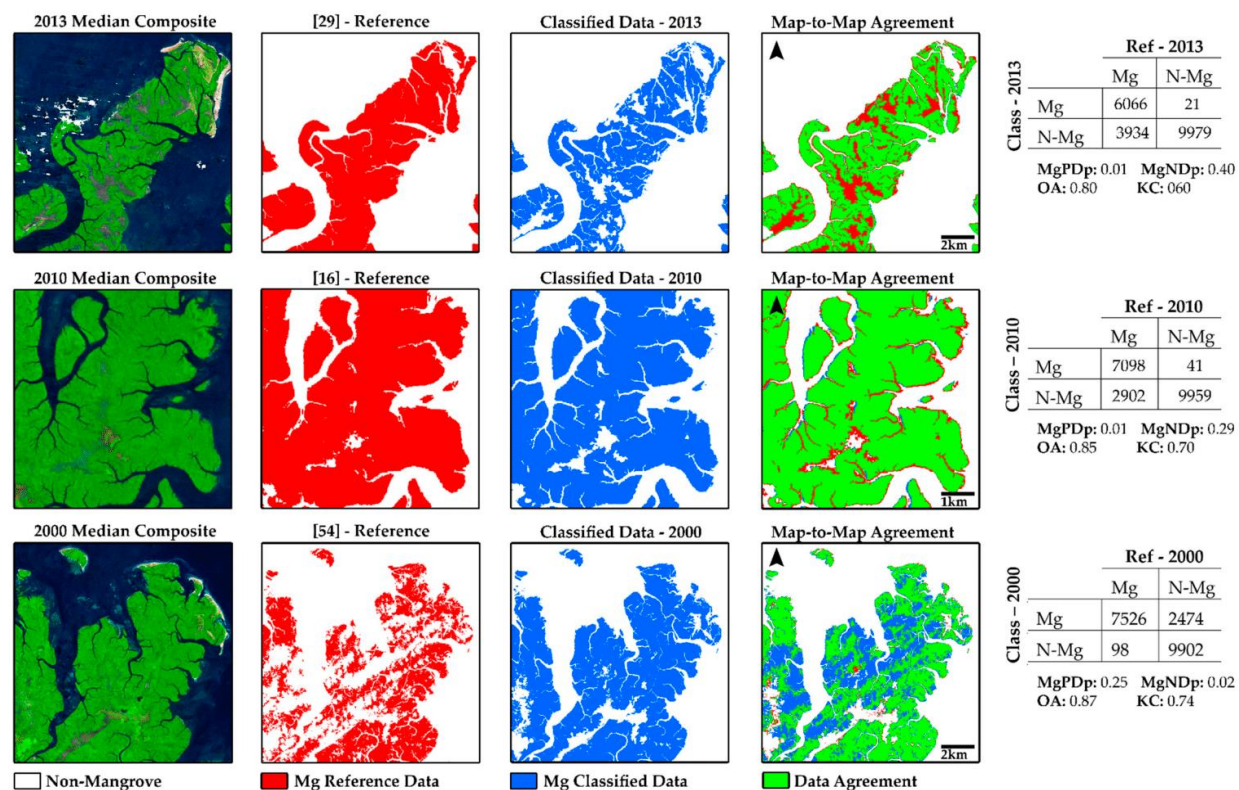


Figure 11. On the left are the 2000, 2010 and 2013 median composites. In the central portion, red represents the mangroves of the reference data, blue denotes the classified mangroves and green represents the agreement between blues and reds. The contingency tables (right) show the agreement levels between the reference and classified data. Values on the main diagonal are the numbers of concordant pixels. On the off-diagonal, those above are positive differences and those below are negative differences. OA stands for overall agreement, MgPDp means the mangrove positive disagreement proportion, MgNDp refers to the mangrove negative disagreement proportion and KC denotes the kappa coefficient.

3.3. Filter-Chain Influence

In total, four different filters may have influenced a resultant classification: gap-fill, temporal, spatial and frequency filters. The first one, gap-fill, may only add mangrove data into a given year; it is an exclusively positive filter (PF). The last two, the spatial and frequency filters, will only remove mangrove pixels from a given year and are therefore negative filters (NF). The temporal filter is the only filter that may add or remove mangrove classifications, depending on the pixel temporal trajectory. Below, Figure 12 and Table 3 categorized and aggregated each pixel, year by year, according to its filtered or no-change aspect, which varied as follows: positive filter (mangrove addition), negative filter (mangrove removal) or unmodified (no change).

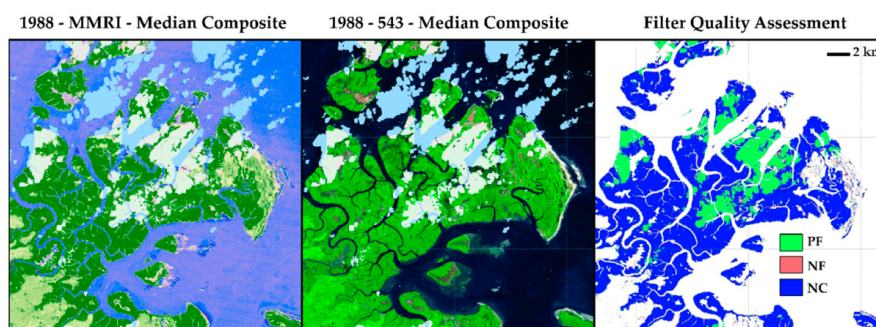


Figure 12. On the left is the 1988 MMRI median composite: dark green represents mangroves, light green is non-mangroves and blue is water. In the centre is the 1988 Landsat-5 band 543 median composite. On the right, PF denotes a positive filter (mangrove addition), NF stands for a negative filter (mangrove removal) and NC stands for no change.

Table 3. Annual filter-affected area, in percentages. PF% indicates a positive filter (mangrove addition), PN% stands for a negative filter (mangrove removal) and NC% denotes an unfiltered pixel (no-change).

Year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
PF%	21	17	18	17	28	24	38	40	47	47	51	47
NF%	23	18	15	14	18	22	13	13	17	14	09	11
NC%	79	83	82	83	72	76	62	60	53	53	49	53
Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
PF%	27	26	13	05	04	03	04	04	05	03	05	04
NF%	10	20	16	12	17	20	18	18	17	16	17	19
NC%	73	74	87	95	96	97	96	96	94	97	95	96
Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018		
PF%	06	04	05	10	09	12	07	07	07	10		
NF%	13	18	22	22	22	19	19	23	24	10		
NC%	94	96	95	91	91	88	93	93	93	90		

The mean percentage values for positive and negative filters were relatively low: 17% and 18%, respectively. On the other hand, the annual average values for the no-change pixels were quite high: 82%. The highest quantity of positive filtering happened in 1995, whereas the lowest appeared in 2006. The highest (24%) and lowest (9%) values of negative filtering appeared in 2017 and 1995, respectively. From 2000 until 2018, the NC percentages were always above the average.

4. Discussion

For the first time in the scientific literature, Brazilian mangrove cover has been systematically and exhaustively mapped. Although data of this nature allow a better understanding of the dynamics of mangrove coverage, it is necessary to understand the effects of the non-homogeneous distribution of Landsat data throughout the time series to avoid superficial interpretations regarding the recorded fluctuations in the country's mangrove areas.

In scenarios of low data frequency, such as the period from 1985 to 1998, it is reasonable to observe a degradation of the quality of "cloud-free composites" since these are based on the median behaviour of pixels composing the orbital data stack. Therefore, the smaller the number of pixels for a given position of the matrix, the higher the likelihood of including atmospheric and radiometric noise as part of the composite, and thus, the decision of the classifier is less accurate in distinguishing between the categories of mangrove and non-mangrove.

When compared separately, the periods from 1985 to 1998 and 1999 to 2018 show distinct trends, as shown in Figures 8 and 9. The first period, from 1985 to 1998, shows an upward trend with a positive slope, as shown by the equation $y = 10.64x + 9827.7$. Aside from environmental and human-related changes that are present in this period, most of this pattern of growth seems to be related more to the uneven distribution of Landsat data throughout the 33 years than to a regeneration movement of Brazilian mangroves. From 1989 to 1998, the PF% values were greater than 20%; for five of those years, the PF% values were greater than 40%.

The second period, from 1999 to 2018, when the PF% values are primarily smaller than 10% and the NC% values are above 85%, a slightly negative slope is shown, presenting a trend of mangrove area loss, which is evidenced by the equation $y = -3.13x + 10070$ and compatible with the global standard expected for mangrove coverage in recent decades [56].

In absolute terms, the direct difference between the years 2000 and 2018 shows an area loss of approximately 70 km²: a difference of ~0.6%. However, it is interesting to note that there is a wave-like pattern present in the variation in mangrove areas. Upon comparing the data from 2003, which has the largest reported area of mangroves, and the 2017 data, which has the smallest quantified area, the absolute difference is ~200 km², which represents an area loss of ~2%, a value almost three times greater than that previously reported by the comparison between 2000 and 2018.

This pattern may be associated with intra-annual oceanographic and climatic variables (rainy and dry seasons), which alter the prevailing tide and humidity conditions in the median composites [57–60]. On one-year intervals, especially over areas often covered by clouds such as the BCZ, the definition of additional image-selection parameters (e.g., the tide condition and the dominant climatic pattern) greatly reduces the number of images available, which makes spatial analysis over shorter time frames impractical.

Regarding the concordance levels, when compared to the year 2000 data [54], an overall agreement of 87% and a kappa coefficient of 74%, $KC = 0.74$, were reached. However, for this reference, there is a large quantity of positive disagreements, which is reflected by the MgPDp of ~25%. Nevertheless, most of these positive differences may be associated with unfiltered clouds and shadow residues in the reference, which were included as part of the non-mangrove class and thus generated positive disagreement.

In comparison to the 2010 data [16], there is significant overall agreement between the published data and our mapping effort, reaching an overall agreement of 85% and a kappa coefficient of 70%. However, the MgNDp is approximately 29%, which is likely to be associated with differences between the edges of the mangrove classes (border errors) and to the inclusion of the coastal rivers' dendritic patterns as part of the mangrove reference class.

For 2013, the year with the greatest mangrove area difference between the reference and our map, approximately 28% (~3900 km²), and the lowest overall agreement and kappa coefficient, 80% and 60%, respectively, the mangrove concept adopted by the reference was different. Unlike the previous concepts, [29] focused on ecosystem identification rather than mangrove forests [16,54]. Thus, the

omissions are quite high and the MgNDp reaches nearly 40%. Furthermore, the reference was created through visual interpretation and, therefore, had a reduced need for the application of spatial filters. As a result, the reported omissions are minorly associated with border differences and are mostly attributed to the inclusion of “apicuns” (salt marshes) as part of the mangrove reference class.

Moreover, it is necessary to highlight the difficulties of implementing continental-scale mapping in a systematic and exhaustive way. The uneven distribution of Landsat data, scarce prior to 1998 and much more frequent in the last two decades, imposes obstacles within the construction of annual cloud-free composites because the quantity of orbital data is greatly unbalanced over time.

Increasing the composites’ temporal windows to three or more years may be important to sharply select data that match specific climatic and oceanographic conditions. In addition, the use of multiple satellite families (Landsat and Sentinel, for example) should also be considered as an alternative to increase the frequency of observations.

Compared to previous mappings produced on an approximately annual timeframe, without considering the prevailing tide and climatic conditions and focusing on mangrove vegetation mapping, the results expressed herein are spatially coherent, reaching a kappa coefficient of 0.7 and an overall agreement above 0.85, as seen in Figure 11.

Along with a broader temporal window, the restriction of the climatic pattern (rainy or dry), as well as the selection of data in specific tide conditions, whether high or low, can reduce the wave-like pattern associated with the reported mangrove areas. Furthermore, the addition of more detailed land use/land cover classes seems to be essential to the ability to identify and separate natural coastal changes from anthropogenic coastal changes, as a more detailed pixel trajectory would allow a more accurate identification of human interference.

From the standpoint of MMRI robustness, compared to other indices, the MMRI seems to provide a greater distinction between the mangrove and non-mangrove strata, which facilitates the discrimination of such classes by automatic classifiers. The new index can be used in conjunction with multiple families of satellites, digital elevation models (DEM) and microwave data and, whenever possible, can incorporate information from the main domains and subdomains of remote sensing, space, time, spectra and context.

The next steps of this research include the creation of a mechanism capable of identifying tidal conditions from orbital data, the expansion of the analysis to the South American continent, and the insertion of new classes of coastal environments such as “apicuns”, beaches and dunes, aquaculture and urban areas, which would help separate human-related changes from natural coastal changes.

5. Conclusions

This manuscript detailed a GEE-managed pipeline to discriminate mangrove forests from surrounding vegetation. The developed pipeline is scalable and suitable for large-scale mangrove cover analyses, allowing systematic and continuous mapping of the Brazilian mangrove cover and producing maps and annual statistics ranging from 1985 to 2018. When compared to previous research at global and national scales, the data produced have high levels of spatial and statistical agreement. All the data produced are available for download through the website <http://www.solved.eco.br/mangroveplatform> and will be transferred to the MapBiomias project (www.mapbiomas.org) in MapBiomias Collection 4.0.

The developed spectral index, MMRI, was demonstrated to be robust and helped to increase the spectral separability between the mangrove and non-mangrove classes. The expansion of the temporal window for the construction of mosaics, the restriction of the climatic pattern (rainy or dry) and the ability to discern between the predominant tidal regimes, whether high or low, may help reduce the area fluctuations associated with the coverage of mangroves.

Author Contributions: Conceptualization, C.D. and L.C.; methodology, C.D. and L.C.; software, L.C., G.N. and J.R.; validation, C.D., L.S. and M.A.; formal analysis, C.D.; investigation, C.D.; resources, P.W.M.S.-F.; data curation, C.D., M.A. and P.W.M.S.-F.; writing—original draft preparation, C.D.; writing—review and editing, C.D., M.A. and P.W.M.S.-F.; visualization, C.D. and L.S.; supervision, P.W.M.S.-F.; project administration, P.W.M.S.-F.; funding acquisition, P.W.M.S.-F.

Funding: This research was funded by the Brazilian National Council for Scientific and Technological Development (CNPq): 870005/1997-9 and the MapBiomias Project. The processing charges were covered by the Dean of Research and Graduate Studies (PROPESP) from the Federal University of Para (UFPA).

Acknowledgments: The authors are grateful to Tasso Azevedo, general coordinator of the MapBiomias Project and the System for Estimation of Green House Gases Emissions (SEEG) initiative, Carlos Souza Jr., senior researcher at the Amazon Institute (IMAZON) and Washington Rocha, full professor at Feira de Santana State University (UEFS), for their useful comments, suggestions and encouragement. The authors would like to thank the anonymous reviewers and the editor for their helpful and constructive comments that greatly contributed to improving this paper. Lastly, the authors acknowledge the Norwegian Government for the financial support given to the Mapbiomas Project, via World Resources Institute (WRI), and Google for the development of the Google Earth Engine platform.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Neumann, B.; Vafeidis, A.T.; Zimmermann, J.; Nicholls, R.J. Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding—A Global Assessment. *PLoS ONE* **2015**, *10*, e0118571. [[CrossRef](#)]
2. Small, C.; Nicholls, R.J. A Global Analysis of Human Settlement in Coastal Zones. *J. Coast. Res.* **2003**, *19*, 584–599. [[CrossRef](#)]
3. Burke, L.; Kura, Y.; Kassem, K.; Revenga, C.; Spalding, M.; McAllister, D. *Pilot Analysis of Global Ecosystems: Coastal Ecosystems*; World Resources Institute: Washington, DC, USA, 2001.
4. Nicolodi, J.L.; Petermann, R.M. Potential vulnerability of the Brazilian coastal zone in its environmental, social, and technological aspects. *Panam. J. Aquat. Sci.* **2010**. [[CrossRef](#)]
5. Dominguez, J.M.L. The Coastal Zone of Brazil. In *Geology and Geomorphology of Holocene Coastal Barriers of Brazil*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 17–51. ISBN 978-3-540-44771-9.
6. Giri, C. Observation and Monitoring of Mangrove Forests Using Remote Sensing: Opportunities and Challenges. *Remote Sens.* **2016**, *8*, 783. [[CrossRef](#)]
7. Giri, C.; Ochieng, E.; Tieszen, L.L.; Zhu, Z.; Singh, A.; Loveland, T.; Masek, J.; Duke, N. Status and distribution of mangrove forests of the world using earth observation satellite data. *Glob. Ecol. Biogeogr.* **2011**, *20*, 154–159. [[CrossRef](#)]
8. Alongi, D.M. *The Energetics of Mangrove Forests*, 1st ed.; Springer: Dordrecht, The Netherlands, 2009; ISBN 978-1-4020-4270-6.
9. Alongi, D.M. Mangrove forests: Resilience, protection from tsunamis, and responses to global climate change. *Estuar. Coast. Shelf Sci.* **2008**, *76*, 1–13. [[CrossRef](#)]
10. Murdiyarso, D.; Purbopuspito, J.; Kauffman, J.B.; Warren, M.W.; Sasmito, S.D.; Donato, D.C.; Manuri, S.; Krisnawati, H.; Taberima, S.; Kurnianto, S. The potential of Indonesian mangrove forests for global climate change mitigation. *Nat. Clim. Chang.* **2015**, *5*, 1089–1092. [[CrossRef](#)]
11. Saenger, P.; Hegerl, E.J.; Davie, J.D.S. *Global Status of Mangrove Ecosystems*; International Union for Conservation of Nature and Natural Resources: Gland, Switzerland, 1983.
12. Alongi, D.M. Present state and future of the world's mangrove forests. *Environ. Conserv.* **2002**, *29*, 331–349. [[CrossRef](#)]
13. Donato, D.C.; Kauffman, J.B.; Murdiyarso, D.; Kurnianto, S.; Stidham, M.; Kanninen, M. Mangroves among the most carbon-rich forests in the tropics. *Nat. Geosci.* **2011**, *4*, 293–297. [[CrossRef](#)]
14. Spalding, M.; Kainuma, M.; Collins, L. *World Atlas of Mangroves*; Taylor & Francis Group: Abingdon-on-Thames, UK, 2010; ISBN 9781849776608.
15. Thomas, N.; Bunting, P.; Lucas, R.; Hardy, A.; Rosenqvist, A.; Fatoyinbo, T. Mapping Mangrove Extent and Change: A Globally Applicable Approach. *Remote Sens.* **2018**, *10*, 1466. [[CrossRef](#)]
16. Bunting, P.; Rosenqvist, A.; Lucas, M.R.; Rebelo, L.-M.; Hilarides, L.; Thomas, N.; Hardy, A.; Itoh, T.; Shimada, M.; Finlayson, M.C. The Global Mangrove Watch—A New 2010 Global Baseline of Mangrove Extent. *Remote Sens.* **2018**, *10*, 1669. [[CrossRef](#)]

17. Herz, R. *Manguezais do Brasil*; United States Pharmacopeia (USP): North Bethesda, MD, USA, 1991.
18. Schaeffer-Novelli, Y.; Cintrón-Molero, G.; Adaime, R.R.; de Camargo, T.M.; Cintron-Molero, G.; de Camargo, T.M. Variability of Mangrove Ecosystems along the Brazilian Coast. *Estuaries* **1990**, *13*, 204–218. [[CrossRef](#)]
19. Souza Filho, P.W.M. Costa de manguezais de macromaré da Amazônia: Cenários morfológicos, mapeamento e quantificação de áreas usando dados de sensores remotos. *Rev. Bras. Geofísica* **2005**, *23*, 427–435. [[CrossRef](#)]
20. Walfir Martins Souza Filho, P.; Renato Paradella, W. Recognition of the main geobotanical features along the Bragança mangrove coast (Brazilian Amazon Region) from Landsat TM and RADARSAT-1 data. *Wetl. Ecol. Manag.* **2002**, *10*, 121–130. [[CrossRef](#)]
21. Souza Filho, P.W.M.; Paradella, W.R. Use of RADARSAT-1 fine mode and Landsat-5 TM selective principal component analysis for geomorphological mapping in a macrotidal mangrove coast in the Amazon Region. *Can. J. Remote Sens.* **2005**, *31*, 214–224. [[CrossRef](#)]
22. Rodrigues, S.W.P.; Souza-Filho, P.W.M. Use of Multi-Sensor Data to Identify and Map Tropical Coastal Wetlands in the Amazon of Northern Brazil. *Wetlands* **2011**, *31*, 11–23. [[CrossRef](#)]
23. Nascimento Jr, W.R.; Souza-Filho, P.W.M.; Proisy, C.; Lucas, R.M.; Rosenqvist, A. Mapping changes in the largest continuous Amazonian mangrove belt using object-based classification of multisensor satellite imagery. *Estuar. Coast. Shelf Sci.* **2013**, *117*, 83–93. [[CrossRef](#)]
24. Santos, L.C.M.; Matos, H.R.; Schaeffer-Novelli, Y.; Cunha-Lignon, M.; Bitencourt, M.D.; Koedam, N.; Dahdouh-Guebas, F. Anthropogenic activities on mangrove areas (São Francisco River Estuary, Brazil Northeast): A GIS-based analysis of CBERS and SPOT images to aid in local management. *Ocean Coast. Manag.* **2014**, *89*, 39–50. [[CrossRef](#)]
25. Queiroz, L.; Rossi, S.; Meireles, J.; Coelho, C. Shrimp aquaculture in the federal state of Ceará, 1970–2012: Trends after mangrove forest privatization in Brazil. *Ocean Coast. Manag.* **2013**, *73*, 54–62. [[CrossRef](#)]
26. Rocha de Souza Pereira, F.; Kampel, M.; Cunha-Lignon, M. Mapping of mangrove forests on the southern coast of São Paulo, Brazil, using synthetic aperture radar data from ALOS/PALSAR. *Remote Sens. Lett.* **2012**, *3*, 567–576. [[CrossRef](#)]
27. Pereira, F.; Kampel, M.; Cunha-Lignon, M. Mangrove vegetation structure in Southeast Brazil from phased array L-band synthetic aperture radar data. *J. Appl. Remote Sens.* **2016**, *10*, 036021. [[CrossRef](#)]
28. Magris, R.A.; Barreto, R. Mapping and assessment of protection of mangrove habitats in Brazil. *Panam. J. Aquat. Sci.* **2010**, *5*, 546–556.
29. ICMBio. *Atlas dos Manguezais do Brasil*, 1st ed.; ICMBio: Brasília, Brazil, 2017; ISBN 978-85-61842-75-8.
30. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [[CrossRef](#)]
31. Chen, X.; Huang, X.; Jiao, C.; Flanner, M.G.; Raeker, T.; Palen, B. Running climate model on a commercial cloud computing environment: A case study using Community Earth System Model (CESM) on Amazon AWS. *Comput. Geosci.* **2017**, *98*, 21–25. [[CrossRef](#)]
32. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [[CrossRef](#)]
33. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
34. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
35. Gao, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
36. Tong, P.H.S.; Auda, Y.; Populus, J.; Aizpuru, M.; Al Habshi, A.; Blasco, F. Assessment from space of mangroves evolution in the Mekong Delta, in relation to extensive shrimp farming. *Int. J. Remote Sens.* **2004**, *25*, 4795–4812. [[CrossRef](#)]
37. Fei, S.X.; Shan, C.H.; Hua, G.Z. Remote Sensing of Mangrove Wetlands Identification. *Procedia Environ. Sci.* **2011**, *10*, 2287–2293. [[CrossRef](#)]
38. Alsaadeh, B.; Al-Hanbali, A.; Tateishi, R.; Kobayashi, T.; Hoan, N.T. Mangrove Forests Mapping in the Southern Part of Japan Using Landsat ETM+ with DEM. *J. Geogr. Inf. Syst.* **2013**, *5*, 9. [[CrossRef](#)]

39. Nardin, W.; Locatelli, S.; Pasquarella, V.; Rulli, M.C.; Woodcock, C.E.; Fagherazzi, S. Dynamics of a fringe mangrove forest detected by Landsat images in the Mekong River Delta, Vietnam. *Earth Surf. Process. Landforms* **2016**, *41*, 2024–2037. [[CrossRef](#)]
40. Pham, D.T.; Yokoya, N.; Bui, T.D.; Yoshino, K.; Friess, A.D. Remote Sensing Approaches for Monitoring Mangrove Species, Structure, and Biomass: Opportunities and Challenges. *Remote Sens.* **2019**, *11*, 230. [[CrossRef](#)]
41. Kuenzer, C.; Bluemel, A.; Gebhardt, S.; Quoc, T.V.; Dech, S. Remote Sensing of Mangrove Ecosystems: A Review. *Remote Sens.* **2011**, *3*, 878–928. [[CrossRef](#)]
42. USGS Landsat. *8 (L8) Data Users Handbook*; USGS Landsat: Sioux Falls, SD, USA, 2015.
43. Storey, J.; Choate, M.; Lee, K. Landsat 8 Operational Land Imager On-Orbit Geometric Calibration and Performance. *Remote Sens.* **2014**, *6*, 11127–11152. [[CrossRef](#)]
44. Teillet, P.M.; Barker, J.L.; Markham, B.L.; Irish, R.R.; Fedosejevs, G.; Storey, J.C. Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sens. Environ.* **2001**, *78*, 39–54. [[CrossRef](#)]
45. USGS Landsat. *USGS Landsat Collection 1 Level 1 Product Definition*; USGS Landsat: Sioux Falls, SD, USA, 2017; Volume 26.
46. Liu, H.Q.; Huete, A. Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 457–465. [[CrossRef](#)]
47. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [[CrossRef](#)]
48. Rogers, A.S.; Kearney, M.S. Reducing signature variability in unmixing coastal marsh Thematic Mapper scenes using spectral indices. *Int. J. Remote Sens.* **2004**. [[CrossRef](#)]
49. Gu, Y.; Brown, J.F.; Verdin, J.P.; Wardlow, B. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophys. Res. Lett.* **2007**, *34*. [[CrossRef](#)]
50. Bogner, C.; Seo, B.; Rohner, D.; Reineking, B. Classification of rare land cover types: Distinguishing annual and perennial crops in an agricultural catchment in South Korea. *PLoS ONE* **2018**, *13*, e0190476. [[CrossRef](#)] [[PubMed](#)]
51. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
52. Pontius, R.G.; Santacruz, A. Quantity, exchange, and shift components of difference in a square contingency table. *Int. J. Remote Sens.* **2014**, *35*, 7543–7554. [[CrossRef](#)]
53. Gupta, K.; Mukhopadhyay, A.; Giri, S.; Chanda, A.; Datta Majumdar, S.; Samanta, S.; Mitra, D.; Samal, R.N.; Pattnaik, A.K.; Hazra, S. An index for discrimination of mangroves from non-mangroves using LANDSAT 8 OLI imagery. *MethodsX* **2018**, *5*, 1129–1139. [[CrossRef](#)] [[PubMed](#)]
54. Giri, C.; Ochieng, E.; Tieszen, L.L.; Zhu, Z.; Singh, A.; Loveland, T.; Masek, J.; Duke, N. *Global Mangrove Forests Distribution, 2000*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2013.
55. Ben Ayed, I.; Punithakumar, K.; Li, S. Distribution Matching with the Bhattacharyya Similarity: A Bound Optimization Framework. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1777–1791. [[CrossRef](#)] [[PubMed](#)]
56. Thomas, N.; Lucas, R.; Bunting, P.; Hardy, A.; Rosenqvist, A.; Simard, M. Distribution and drivers of global mangrove forest change, 1996–2010. *PLoS ONE* **2017**, *12*, e0179302. [[CrossRef](#)]
57. Xia, Q.; Qin, C.-Z.; Li, H.; Huang, C.; Su, F.-Z. Mapping Mangrove Forests Based on Multi-Tidal High-Resolution Satellite Imagery. *Remote Sens.* **2018**, *10*, 1343. [[CrossRef](#)]
58. Rogers, K.; Lymburner, L.; Salum, R.; Brooke, B.P.; Woodroffe, C.D. Mapping of mangrove extent and zonation using high and low tide composites of Landsat data. *Hydrobiologia* **2017**, *803*, 49–68. [[CrossRef](#)]
59. Zhang, X.; Treitz, P.M.; Chen, D.; Quan, C.; Shi, L.; Li, X. Mapping mangrove forests using multi-tidal remotely-sensed data and a decision-tree-based procedure. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *62*, 201–214. [[CrossRef](#)]
60. Chen, B.; Xiao, X.; Li, X.; Pan, L.; Doughty, R.; Ma, J.; Dong, J.; Qin, Y.; Zhao, B.; Zhixiang, W.; et al. A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. *ISPRS J. Photogramm. Remote Sens.* **2017**, *131*, 104–120. [[CrossRef](#)]



4.2. ARTIGO II - A LARGE-SCALE DEEP-LEARNING APPROACH FOR MULTI-TEMPORAL AQUACULTURE MAPPING

SUBMITTED ON 25 FEB 2021

Proof of Submission:

[Remote Sensing] Manuscript ID: remotesensing-1141650 - Submission Received



Editorial Office

para Cesar, Luiz, Maria, Luis, Alexandre, Luis, Marcos, Pedro ▾

Dear Mr. Diniz,

Thank you very much for uploading the following manuscript to the MDPI submission system. One of our editors will be in touch with you soon.

Journal name: Remote Sensing

Manuscript ID: remotesensing-1141650

Type of manuscript: Article

Title: A Large-Scale Deep-Learning Approach for Multi-Temporal Aqua and Salt-Culture Mapping

Authors: Cesar Diniz *, Luiz Cortinhas, Maria Luize Pinheiro, Luis Sadeck, Alexandre S. F. Filho, Luis Baumann, Marcos Adami, Pedro Walfir M. Souza-Filho

Received: 25 February 2021

E-mails: cesar.diniz@solved.eco.br, luiz.cortinhas@solved.eco.br, m.luize.s.pinheiro@gmail.com, luis.sadeck@solved.eco.br, alexandresff25@gmail.com, fbaumann@ufg.br, marcos.adami@inpe.br, pedro.martins.souza@itv.org

State-of-the-Art Remote Sensing in South America

https://www.mdpi.com/journal/remotesensing/special_issues/South_America

You can follow progress of your manuscript at the following link (login required):

https://susy.mdpi.com/user/manuscripts/review_info/13b9a116d76716d876be825250c4b565

1 Article

2 A Large-Scale Deep-Learning Approach for Multi- 3 Temporal Aqua and Salt-Culture Mapping

4 Cesar Diniz^{1,2*}, Luiz Cortinhas^{1,3}, Maria Luize^{1,3}, Luis Sadeck¹, Alexandre Pinho^{1,4}, Luis R. F.
5 Baumann⁵, Marcos Adami^{2,6}, Pedro Walfir M. Souza-Filho^{2,7}

6 ¹ Solved – Solutions in Geoinformation7 ² UFPA – Federal University of Pará, Geoscience Institute8 ³ UFPA – Federal University of Pará, Technology Institute9 ⁴ UFRA – Federal Rural University of the Amazon, Institute of Agricultural Sciences10 ⁵ UFG – Federal University of Goiás, Institute of Mathematics and Statistics11 ⁶ INPE – National Institute for Space Research, Amazon Spatial Coordination12 ⁷ ITV – Instituto Tecnológico Vale

13

14 * Correspondence: cesar.diniz@solved.eco.br; Tel.: +55-91-98239-1597

15 **Abstract:** Aquaculture and salt-culture are relevant economic activities in the Brazilian Coastal Zone
16 (BCZ). However, automatic discrimination of such activities from other water-related covers/uses
17 is not an easy task. In this sense, convolutional neural networks (CNN) have the advantage of
18 predicting a given pixel's class label by providing as input a local region (named patches or chips)
19 around that pixel. Both the convolutional nature and the semantic segmentation capability provide
20 the U-Net classifier with the ability to access the "context domain" instead of solely isolated pixel
21 values. Backed by the context domain, the results obtained show that the BCZ aquaculture/saline
22 ponds occupied ~356 km² in 1985 and ~544 km² in 2019, reflecting an area expansion of ~51%, a rise
23 of 1.5x in 34 years. From 1997 to 2015, the aqua-salt-culture area grew by a factor of ~1.7, jumping
24 from 349 km² to 583 km², a 67% increase. In 2019, the Northeast sector concentrated 93% of the
25 coastal aquaculture/salt-culture surface, while, Southeast and South sectors contained 6% and 1%,
26 respectively. Interestingly, despite presenting extensive coastal zones and suitable conditions for
27 developing different aqua-salt-culture products, the North coast shows no relevant aqua or salt-
28 culture infrastructure sign.

29 **Keywords:** Aquaculture; Salt-culture; U-net; Tensor-Flow; Google Earth Engine; Landsat

30

31 1. Introduction

32 Aquaculture is the farming of aquatic organisms, including fish, mollusks, crustaceans, and
33 aquatic plants. In this sense, farming implies some form of intervention in the rearing process to
34 enhance production, such as regular stocking, feeding, and predators' protection. Farming also
35 presupposes individual or corporate ownership of the stock being cultivated [1].

36 Currently, half of the fish and mollusks used to feed the world's population comes from
37 aquaculture [2], and the world's appetite for fish and fish products shows no sign of slowing. From
38 1990 to 2018, there was a ~14% rise in global capture fisheries, a 527% rise in aquaculture production,
39 and a 122% rise in total food fish consumption [3].

40 Sharing similar characteristics with the aquaculture infrastructure, such as the exploitation of
41 water resources, the need for artificial water-reservoir, and the dependency on coastal environments,
42 the sea salt culture is another relevant economic activity in the Brazilian Coastal Zone (BCZ).
43 Therefore, mapping aquaculture and salt culture areas is a fundamental challenge for understanding

44 the expansion of such activities and the sustainable management of these global coastal resources.

45 Aquaculture and salines have been mainly associated with valuable fertile environments and
46 maybe the origin of large-scale land-use changes. Due to its water-dependent nature, coastal
47 aquaculture and saline activity require special attention, being directly associated with the loss of
48 natural coastal wetlands and pollution of waters and soils [4]. Other studies have relayed on remote
49 sensing images to understand the dynamics of classical coastal features such as mangroves, beaches,
50 estuaries, and shoreline analysis [5–13]. However, there are fewer large-scale studies exhaustively
51 identifying coastal aquaculture and salines, whether on the global or regional scale and yet, none rely
52 on deep-learning algorithms [14,15].

53 The primary difficulty in automatically distinguishing between aqua or salt-culture and any
54 other surface water target (e.g., rivers, lakes, or coastal waters) is the spectral similarity between them
55 all. Spectrally speaking, water is water and, unless it presents a high concentration of optically active
56 compounds (e.g., suspended sediments, algae, organic matter, and others.), not much can be done to
57 discriminate a variety of water-dependent targets spectrally. In cases like this, when the natural cover
58 among the targets to be set apart are the same, the context domain, rather than the spectral or
59 temporal one, may hold the key to distinguish between rivers, lakes, natural or artificial ponds, and
60 coastal water pixels.

61 In this scenario, the U-Net: Convolutional Networks [16,17] has the advantage of predicting the
62 class label of each pixel by providing as input a local region (named patches or chips) around that
63 given pixel. Such labeling characteristics, together with its convolutional nature and the semantic
64 segmentation capability, provide the U-Net with the ability to access the "context domain" instead of
65 solely isolated pixel values.

66 This research tests the application of a U-net classifier over the BCZ, a coastal extension of ~9000
67 km, to assess its robustness in the multitemporal discrimination of artificial coastal ponds of aqua
68 and salt-culture. The region presents a great diversity of coastal environments, such as mangroves,
69 salt-flats, beaches, dunes, and coastal waters (e.g., estuaries, deltas, and tidal-rivers) that influence
70 the organization and spatial distribution of coastal aquaculture [18–23]. All this variability and
71 complexity make the BCZ an appropriate region to evaluate the proposed method's robustness. This
72 paper aims to present, describe, and assess an Earth Engine-TensorFlow pipeline based on a U-net
73 large-scale deep-learning approach. The pipeline has been designed to classify artificial coastal ponds
74 to pave the understanding of the aqua-salt culture dynamics in the BCZ, based on 34 years of Landsat
75 time-series data, from 1985 to 2019.

76

77 2. Materials and Methods

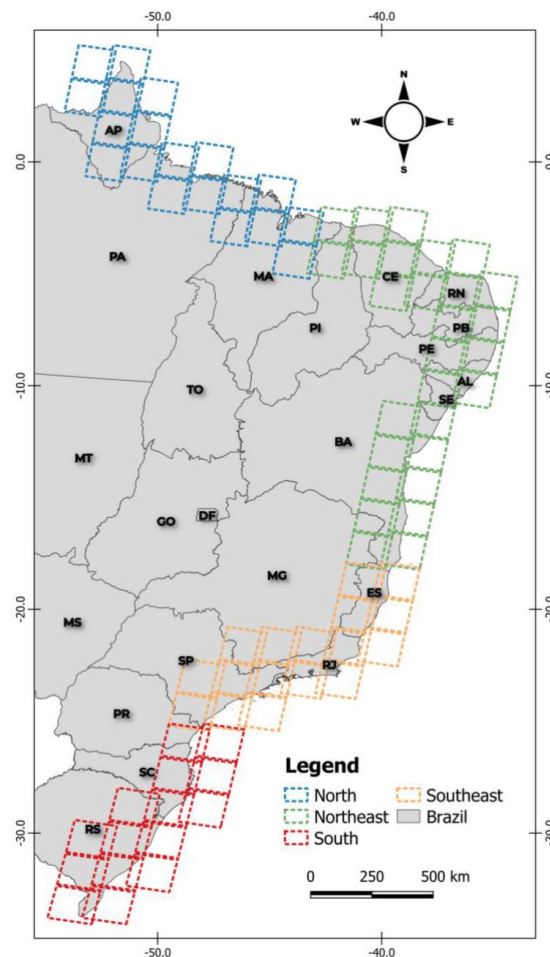
78 2.1. Study Site

79 In Brazil, between 2013 and 2018, aquaculture activity production increased by ~18%, jumping
80 from 476 tons to 576 tons [24]. The Northeast and South regions lead the national aquaculture
81 production; together, these regions are responsible for ~56% of the domestic market [24]. Compared
82 to the global scenario, the experiment rate of increase can be considered relatively low, a characteristic

83 that may be associated with the rapid expansion of white spot syndrome disease (WSSD), especially
 84 in the northeastern region of the country [25–27].

85 From the diversity perspective, Brazilian aquaculture is primarily dominated by fish farming,
 86 followed by shrimp farming, and finally by bivalve culture (oysters, scallops, and mussels). In 2018
 87 these three crops represented near 89%, 8%, and 2% of national production, respectively [24]. While
 88 fish farms are widely spread throughout the country, shrimp farming and bivalve culture are mostly
 89 concentrated in the Northeast and South regions, respectively [28–32].

90 Together with traditional aquaculture, sea salt culture is also common in the BCZ. Based on the
 91 evaporation of saline reservoirs', this activity is responsible for 81% of Brazil's salt production. Sea
 92 salt production is mainly associated with three states: Rio Grande do Norte (RN), Ceará (CE), and
 93 Rio de Janeiro (RJ) [4,5]. In 2017, 6 Mt of sea salt was produced in Brazil, 95% in RN, 3% from RJ, and
 94 another 1.5% from CE [5]. To better present the achieved results, the BCZ was divided into four
 95 regional sectors; North, Northeast, Southeast, and South (Figure 1).

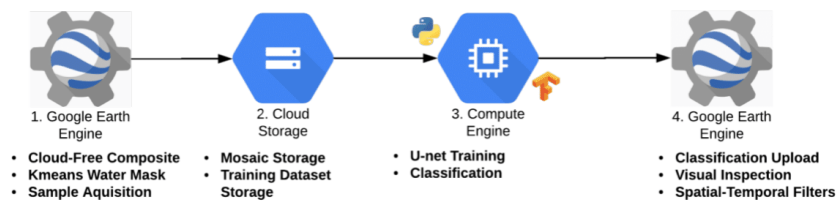


96
 97

Figure 1 –The Brazilian Coastal Zone (BCZ) was divided into four regional sectors; North, Northeast, Southeast, and South.

98 2.2. Data Processing

99 Data processing and analysis occurred inside a cloud computing environment, a setup
 100 combining the Google Earth Engine platform (GEE) and Tensor Flow library (Figure 2). All raster
 101 data and its sub-products are derived from the United States Geological Survey (USGS) Landsat
 102 Collection 1 Tier 1 Top of Atmosphere (TOA) data, which includes Level-1 Precision Terrain (L1TP)
 103 data [33–35].



104

105

106

107

108

109

Figure 2 - Aquaculture Earth Engine-TensorFlow pipeline. The process is structured in 4 steps. First (1), GEE is used to generate the cloud-free composites and create the initial training dataset. Second (2), the mosaics and training data are stored inside the Google Cloud Storage. Three (3), the patch-wise training and classification are initiated. In the fourth step (4), the classified product is spatial and temporally filtered. The filtered product is visual and statistically inspected. Multiple iterations may be used until a satisfactory degree of spatial and temporal quality is achieved.

110

111

112

113

114

115

116

For each year, Landsat Collection 1, Tier 1, TOA data was used to produce annual cloud-free composites, ranging from January 1st to December 31st. The cloud/shadow removal script takes advantage of the quality assessment (QA) band and the GEE median reducer. When used, QA values can improve data integrity by indicating which pixels might be affected by artifacts or subject to cloud contamination [36]. In conjunction, GEE can be instructed to pick the median pixel value in a stack of images. By doing so, the engine rejects values that are too bright (e.g., clouds) or too dark (e.g., shadows) and picks the median pixel value in each band over time.

117

118

119

120

Subsequently, the annual mosaics were sub-set to the area that comprises the BCZ to include areas where coastal aquaculture and saline ponds are more likely to occur (e.g., low-lying coastal and intertidal regions) and to exclude large areas where such characteristics are not expected to occur (e.g., highlands distant from the coastal plain and open ocean areas).

121

122

123

124

125

126

127

128

Next, the Modified Normalized Difference Water Index (MNDWI) [37] is computed to each of the annual mosaics, and over each, a K-means cluster analysis is executed. The K-means analysis helps create a binary water mask, labeling the pixels as water or non-water classes. Guided by the water/non-water mask's existence, the coastal ponds of aqua or salt-culture features are visually delineated. It is essential to highlight that no differentiation is made between coastal aquaculture or saline ponds; both elements are grouped as "aqua-salt culture" class members. Therefore, from this point on, every time aquaculture and salt-culture samples or classes are mentioned, it is referring to an aqua-salt culture pattern.

129

130

131

132

133

134

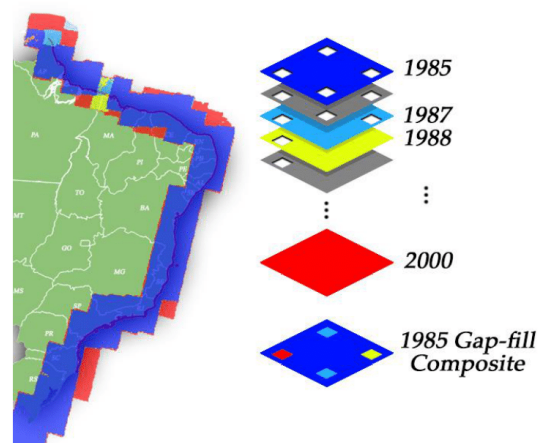
This first visual delineation generates aqua-salt culture and non-aqua-salt culture samples. Once the sample acquisition is finished, the U-net classification is run, resulting in the pre-filtered classification product. The classified data was injected back into GEE, where spatial-temporal filters and visual inspection occur. This phase was undertaken to correct misclassified data and evaluate the necessity of acquiring (or not) more training samples. Table 1 shows the classifier parameters.

135 **Table 1** - Classifier attributes and classification parameters. In total, six distinct attributes were used.

Parameters	Values
Classifier	U-Net
Tile-Size	256 x 256 pixels
Samples	8400 (geometries)
Attributes	Swir1, Nir1, Red, MNDWI, NDVI, and NDSI
Classes	2 (Aqua-Salt-culture and Non-Aqua-Salt-culture)

136

137 Due to the classification methods' pixel-based nature and the very long temporal series (35
 138 years), a chain of post-classification filters was implemented. The chain starts by filling in possible
 139 no-data values. In a long time series of severely cloud-affected regions, such as tropical coastal zones,
 140 it is expected that no-data values may populate some of the resultant median composite pixels. In
 141 this filter, no-data values ("gaps") are theoretically not allowed and are replaced by the temporally
 142 nearest valid classification. If no "future" valid position is available in this procedure, then the no-
 143 data value is replaced by its previous valid class. To fill persistent no-data positions, up to three
 144 previous years can be used. Therefore, gaps should only exist if a given pixel has been permanently
 145 classified as no-data throughout the entire temporal domain. As shown in Figure 3, a mask of years
 146 was designed to keep track of the pixels' temporal origin.



147

148 **Figure 3** – Gap-filling mechanism. The next valid classification replaces existing no-data values. If no "future"
 149 valid position is available, then the no-data value is replaced by its previous valid classification, based on up to
 150 a maximum of three (3) prior years. To keep track of pixel temporal origins, a mask of years was built.

151 After gap filling, a temporal filter was executed. The temporal filter uses sequential
 152 classifications in a 3-year unidirectional moving window to identify temporally non-permitted
 153 transitions. Based on a single generic rule (GR), the temporal filter inspects the central position of
 154 three consecutive years ("ternary"), and if the extremities of the ternary are identical but the center
 155 position is not, then the central pixel is reclassified to match its temporal neighbor class, as shown in
 156 Table 2.

157

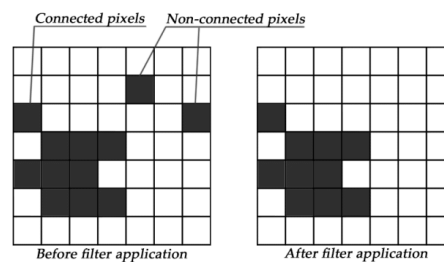
158 **Table 2** - The temporal filter inspects the central position of three consecutive years, and in cases of identical
 159 extremities, the center position is reclassified to match its neighbor. T1, T2, and T3 stand for positions one (1),
 160 two (2), and three (3), respectively. GR means "generic rule," while AS and N-AS represent aqua-salt-culture
 161 and non-aqua-salt culture pixels.

Rule	Input (Year)			Output		
	T1	T2	T3	T1	T2	T3
GR	AS	N-AS	AS	AS	AS	AS
GR	N-AS	AS	N-AS	N-AS	N-AS	N-AS

162

163 Later, a spatial filter was applied. To avoid unwanted modifications to the edges of the pixel
 164 groups (blobs), a spatial filter was built based on the "connectedPixelCount" function. Native to the
 165 GEE platform, this function locates connected components (neighbors) that share the same pixel
 166 value. Thus, only pixels that do not share connections to a predefined number of identical neighbors
 167 are considered isolated, as shown in Figure 4. In this filter, at least ten connected pixels are needed to
 168 reach the minimum connection value. Consequently, the minimum mapping unit is directly affected
 169 by the spatial filter applied, and it was defined as 10 pixels (~1 ha).

170



171

172 **Figure 4** – The spatial filter removes pixels that do not share neighbors of identical value. The
 173 minimum connection value was 10 pixels.

174 The accuracy assessment analysis was performed based on 900 independent samples taken at
 175 the Landsat pixel level for a pair of years, 1985 and 2019. A total of 1.800 samples were used, 900 per
 176 year. These samples were generated by a stratified random sampling, where the aqua-salt culture
 177 class's rareness compared to the non-aqua-salt culture class was taken into account. The validation
 178 samples' distribution occurred in three different strata: 1) inside a temporally unified aqua-salt-
 179 culture surface (the annual classifications were merged, forming a single aqua-salt culture surface)
 180 450 samples were taken. 2) the previously merged stratum was buffered and, 400 samples were taken
 181 inside the buffered zone, and 3) outside of the strata 1 and 2 boundaries 50 samples were acquired.
 182 The total sample size was calculated according to the equation below.

183

184

$$n = \frac{Nz^2p(1-p)}{(N-1)e^2 + z^2p(1-p)}$$

185

Where;

186

n , is the sample size.

187

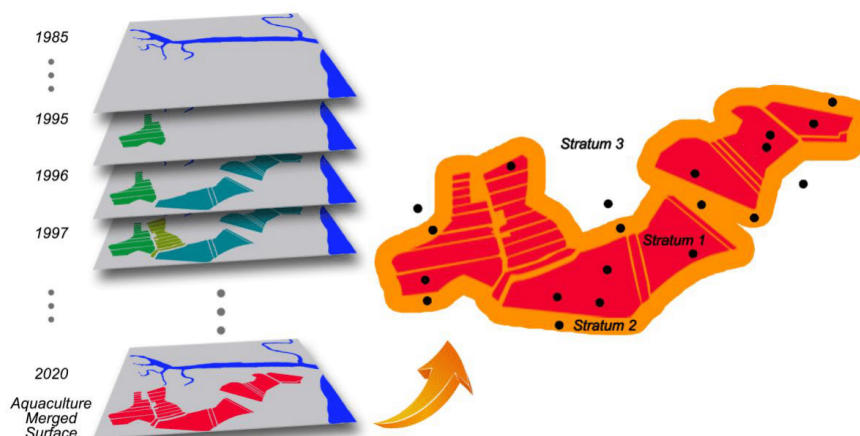
N , is the population size.

188

z , is the score of the normal distribution at a given confidence degree.

189 γ , is the confidence degree.
 190 p , is the population proportion to be estimated.
 191 e , the maximum error margin.

192 The validation samples were distributed to allow the possibility of evaluating the occurrence of
 193 both false positives and false negatives (Figure 5). A confidence interval of 95% and a maximum
 194 standard error of 5% were assumed to establish the sample size for each stratum, following the
 195 guidance of classical accuracy assessment practices [38–40].



196
 197 **Figure 5** – A total of 900 independent samples were taken at the Landsat pixel level in 1985 and 2019. 1.800
 198 samples in total. The sample distribution occurred in three different strata: 1) 450 samples inside a temporally
 199 merged aquaculture surface. 2) 400 samples inside the stratum 1 buffered zone, and 3) 50 samples outside of the
 200 strata 1 and 2 boundaries.

201 The accuracy assessment of the 1985 and 2019 classification was calculated based on an
 202 independent visual inspection. This visual evaluation inspected the same Landsat annual mosaics
 203 given as input to the U-net classifier. As decisive support data, the interpreters were also guided by
 204 Landsat multitemporal information, water indices, and high-resolution imagery available in Google
 205 Earth Pro.

206 Finally, an error matrix was produced for each of the inspected years. The following accuracy
 207 metrics were calculated per class, per year: Overall accuracy (OA), quantity and allocation
 208 disagreement (QD and AD), user and producer accuracy (CA and PA), omission and commission
 209 error (OE and CE), and the Kappa Coefficient (KC) [38–40].

210

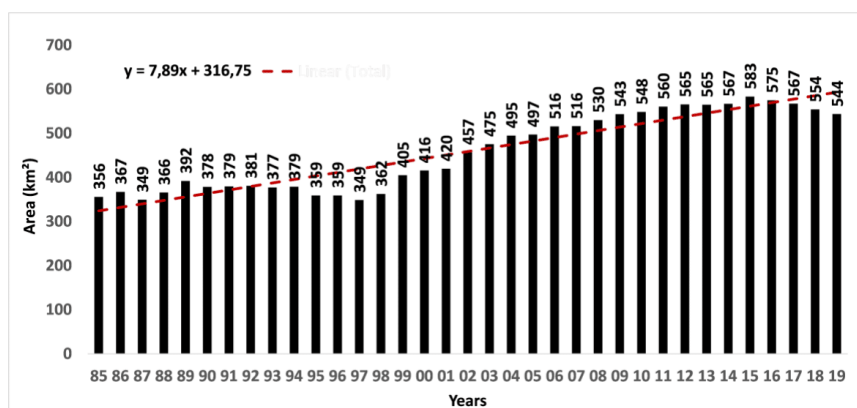
211 3. Results

212 3.1. Spatio-Temporal Changes of Aqua-Salt Culture Ponds In BCZ

213 The coastal aqua-salt-culture activity has been identified and quantified for the entire BCZ over
 214 34 years in a pioneering spatial and temporal coastal analysis. The aqua-salt culture features were

215 systematically and exhaustively mapped, producing digital spatial products and annual statistics
216 ranging from 1985 to 2019 (Figure 6).

217 In Brazil, aqua-salt culture ponds occupied ~356 km² in 1985 and ~544 km² in 2019, representing
218 an area expansion of ~188 km², ~52%. Two distinct periods can be observed: from 85 to 98, certain
219 stability is noticed, while from 1999 to 2019 a period of rapid growth appears. Considering the entire
220 time-span, the area associated with aqua-salt culture activity expanded 1.5x in 34 years (Figure 6).
221 The late 90s marked an accelerated development of aqua or salt-culture-related patterns. From 1997
222 to 2015, the area associated with this activity grew by a factor of ~1.7, jumping from 349 km² to 583
223 km², a 67% rise.



224

225 **Figure 6** – Annual variation of aqua-salt-culture area, in km², from 1985 to 2019. The area associated with this
226 activity expanded 1.5x in 34 years—an increase of ~51%. The black bars denote the aqua-salt culture area. The
227 red dashed line indicates the ascendant trend of aqua-salt culture evolution.

228 Regarding the presence or absence of aqua-salt-culture activity, out of 17 possible coastal states
229 (Figure 1), three did not present any signal of aquaculture or saline-related patterns, Amapá (AP),
230 Para (PA), and Parana (PR). The state of Maranhão (MA) presented the slightest, but a not null,
231 detected area. In MA, the detected ponds existed more prominently in the 80s, but the signals of
232 aquaculture or saline activities continuously decreased through time, reaching ~0.11 km² in 2019
233 (Figure 7A).

234 In terms of coastal distribution, Figure 7A, the states of Rio Grande do Norte and Ceará presents
235 the most extensive aqua-salt culture areas in the country, reaching ~376 km² (~70%) and ~ 82 km²
236 (~15%) of aqua-salt culture extension in Brazil, respectively, in 2019. Together, these two states
237 represent ~85% of Brazil's aqua-salt culture land-use.

238 Most Brazilian coastal states revealed an expansion of aqua-salt culture land-use in at least one
239 of the compared years, 1985 and 2019 (Figure 7A). Interestingly, the states of AP, PA, and MA in the
240 Northern sector, despite presenting extensive coastal zones and suitable conditions for developing
241 various aquaculture products, especially shrimp culture, show no relevant sign of aqua-salt culture
242 infrastructure.

243 On a regional scenario, the South sectors represent approximately 1% of the total aqua-salt
244 culture area, none of that on the North coast (Figure 7B). In contrast, the Northeast sector concentrates

245 93% of the coastal aqua-salt culture ponds (Figure 7C). The Southeast region is the second most
 246 relevant sector in terms of extension, contemplating 6% of the national coastal aqua-salt culture
 247 (Figure 7D). In comparison, the South sector contains only 1% of this land-use area (Figure 7E).
 248

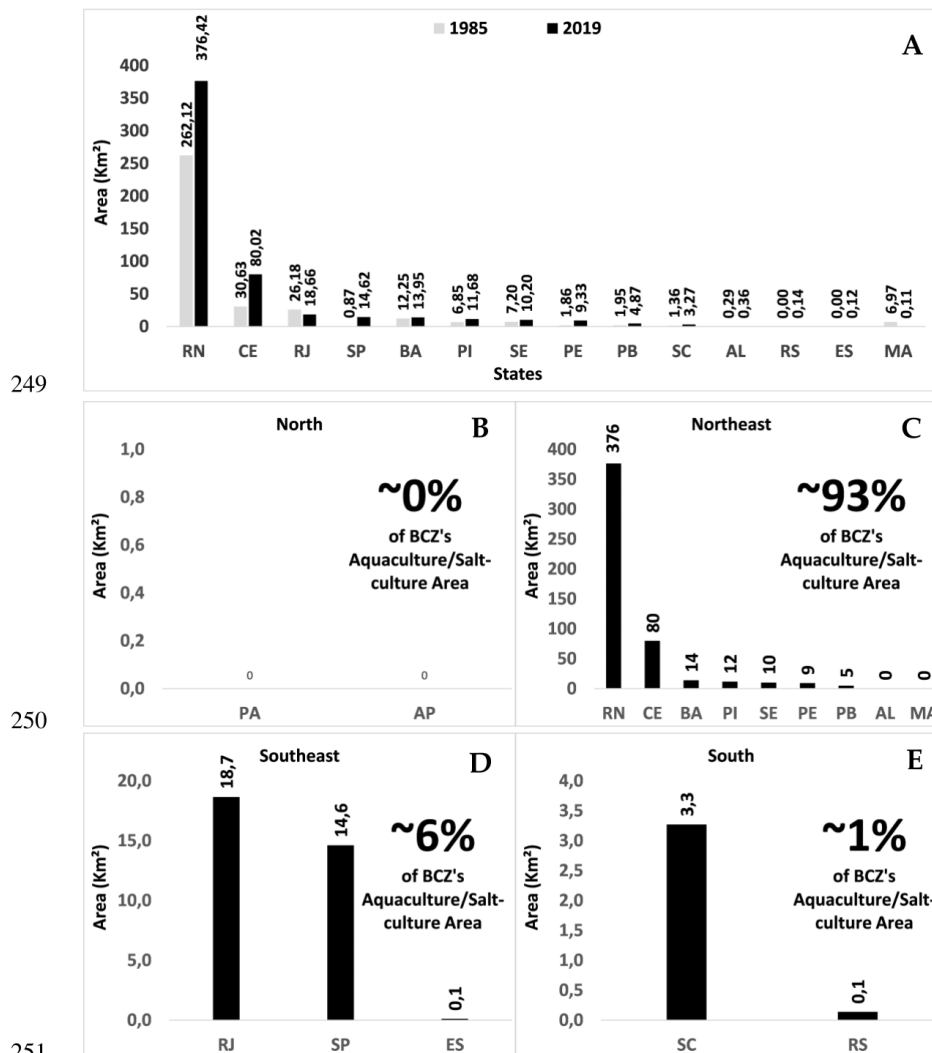
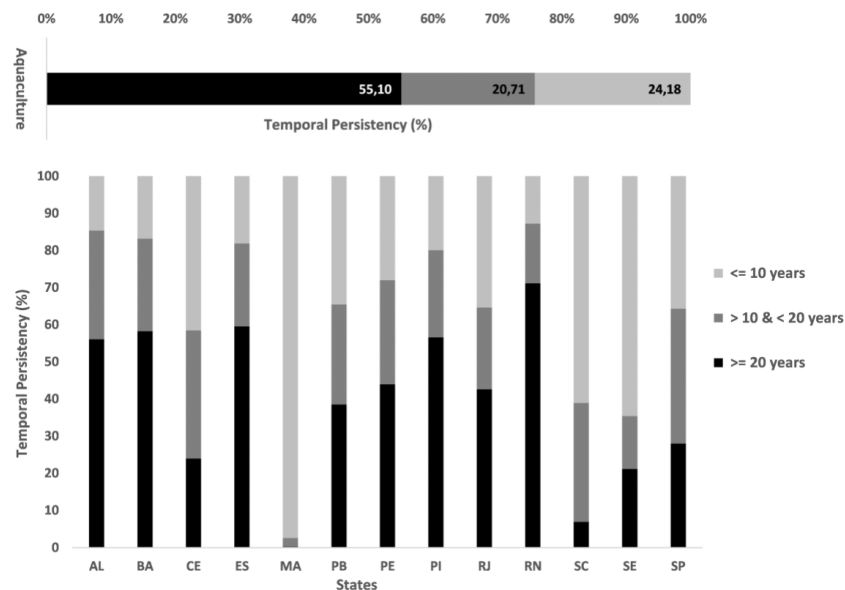


Figure 7 – A) Brazil's aqua-salt culture area per federal unit in km². Aqua-salt culture extension in km² and percentage per sector: B) North, C) Northeast, D) Southeast, and E) South are responsible, in sequence, by 0%, 93%, 6%, and 1% of the national aquaculture distribution.

255 From the perspective of temporal aquaculture persistence, ~55% of such coastal land-use
 256 remained unchanged for two decades or more, ~20% persisted between one and two decades, and
 257 ~24% remained stable for ten or fewer years. In this scenario, the state of Maranhão showed the lowest

258 temporal persistence level; near 97% of its aqua-salt culture ponds existed for less than ten years. In
 259 opposition, Rio Grande do Norte demonstrated the highest temporal ponds stability, with 71% of its
 260 aqua-salt culture patterns remained unchanged for 20 years or more, as illustrated in Figure 8.
 261



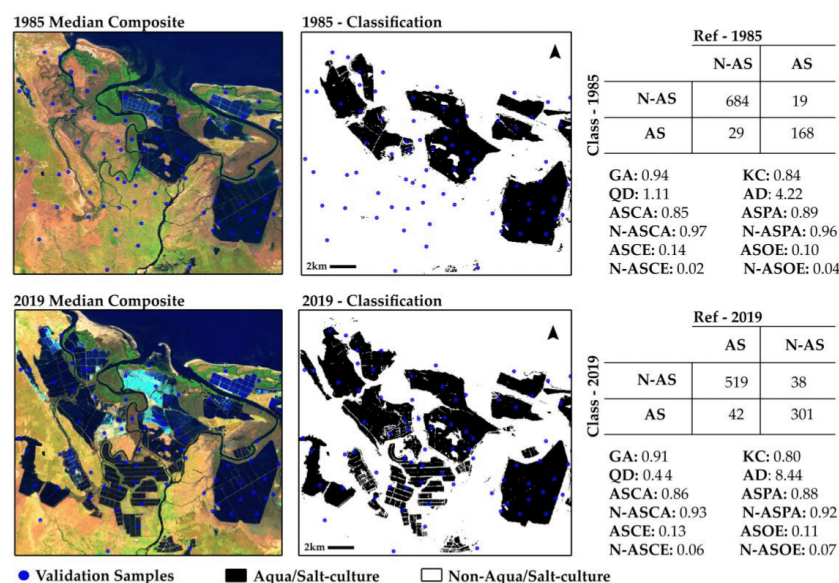
262
 263 **Figure 8** – Aqua-salt culture persistence at national and regional scales. The top bar shows the overall
 264 aqua-salt-culture persistence. The bottom graph shows the aqua-salt-culture persistence per state. The
 265 x-axis represents the state distributions, while the y-axis represents the mangrove cover temporal
 266 persistence percentages (%). Black represents 20 years or more of stability, and dark grey indicates
 267 stability between 10 and 20 years, and light grey represents stability for less than ten years

268 3.2. Aqua-Salt Culture Accuracy Assessment

269 The coastal spatial distribution of aqua-salt culture ponds was unknown until this publication.
 270 Thus, a visual accuracy assessment campaign was executed to determine U-net classified products'
 271 agreement levels compared to a human-interpreted reference. Details related to the accuracy
 272 assessment sampling method are available in Figure 5. The accuracy campaign involved the cross-
 273 comparison of two sets of data, 1985 and 2019. Figure 9 below shows the contingency tables for both
 274 years.

275 Compared to the 1985 reference, the classification developed herein achieves an overall
 276 agreement of 94% (OA = 0.94), quantity disagreement (QD), and allocation disagreement (AD) of
 277 1.11% and 4.22%, respectively, a Kappa Coefficient (KC) ~0.84, and presents a low proportion of false
 278 negatives (~0.10). It shows around 14% of false positives, representing roughly 1/6 of disagreement
 279 between what is referred to as N-AS, but is classified as AS (commission error).

280 Regarding 2019 data, the degree of global concordance is equally high, achieving an overall
 281 agreement of 91% (OA = 0.91). The QD and AD were 0.44% and 8.44%, respectively, with KC ~0.81.
 282 The presented proportions of aqua-salt culture's false positives and false negatives were smaller than
 283 15%, with AS commission errors (ASCE) ~0.13, while AS omission errors (ASOE) reached ~0.11. The
 284 N-AS class had omission and commission errors minor than 7% (N-ASCE ~0.06 and ASOE ~0.07).



285
 286 **Figure 9** – On the left are the 1985 and 2019 median composites. In the central portion, U-Net classifications from
 287 1985 and 2019. The contingency tables (right) show the agreement levels between the reference and classified
 288 data. Values on the main diagonal are the numbers of concordant pixels. On the off-diagonal, those above are
 289 commission differences, and those below are omission differences. OA stands for overall agreement. QD and
 290 AD for Quantity and Allocation Disagreements, KC denotes the Kappa coefficient. ASCA/N-ASCA means Aqua-
 291 Salt-culture/Non- Aqua-Salt-culture Consumer Accuracy, while AS/N-AS refers to the Aquaculture/Non-
 292 Aquaculture Producers Accuracy. Both classes' commission and omission errors are referred to as ASCE, N-
 293 ASCE, ASOE, N-ASOE.

294 4. Discussion

295 For the first time in the scientific literature, the Brazilian coastal aqua and salt-culture ponds
 296 were systematically and exhaustively mapped. Studies of such nature provide a better understanding
 297 of the country's coastal dynamics, allowing an annual analysis of two of the most classical coastal
 298 land-uses, directly subsidizing national and regional policymakers.

299 In the last three decades, two distinct moments can be highlighted in Brazil's coastal aqua-salt
 300 culture occupation (Figure 6). From 1985 to 1998, a moment of stability was noticed, in which
 301 variations of nearly $\pm 5\%$ alternately occur. On the other hand, from 1999 to 2019, a moment of rapid
 302 growth marked an accelerated expansion of the activity. This spatial-temporal behavior matches the
 303 history of marine salt production in Brazil [41–43], particularly in the states of RN and CE,
 304 responsible for ~85% of Brazil's aqua-salt culture area in 2019 (Figure 7).

305 The discrepant spatial distribution of the aqua-salt pond areas in the BCZ is primarily related to
306 geological and economic characteristics more than any other. Most of the national saline industry was
307 purposely implanted in areas where salt formation through solar evaporation naturally occurs
308 (hypersaline-shallow plains over dry climate) [4,32]. Moreover, physically dependent on solar
309 efficiency, evaporation ponds tend to be much larger and shallow than aquaculture ones. Together,
310 this set of characteristics helps comprehend the spatial distortions involving RN and CE compared
311 to the remaining coastal states.

312 Similarly, economic, geological, and climatic characteristics might be the most significant factors
313 influencing the absence of relevant areas of salines or aquaculture in the North of Brazil. Amapá,
314 Pará, and Maranhão's states present extensive coastal plains and suitable conditions for developing
315 various aqua-salt culture products. Although all three states lack economic competitiveness due to
316 the absence of requisite infrastructure to sustain large-scale aquaculture farming [29] and accentuated
317 precipitation regimes compromise marine salt-production [41,43–45]. Furthermore, the densest and
318 well-developed mangrove vegetation in the country acts as a natural barrier, challenging the
319 implementation of artificial ponds over the dense mangrove forest [8,18,21,46].

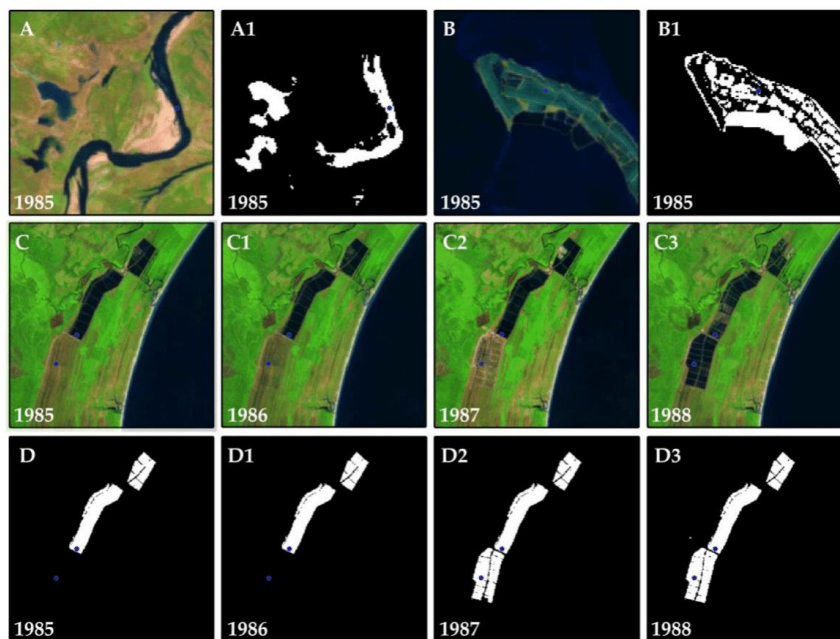
320 Regarding the concordance levels, when cross-compared to the references of 1985 and 2019, the
321 U-net classified product achieved high levels of overall agreement and kappa coefficients, OA >90%
322 and KC >0.80 for both years and a small proportion of quantity and allocation disagreements. The
323 worst QD measure case reached ~1% for 1985 data, and the max AD measure was not higher than
324 9%; AD ~8.5% in 2019. Moreover, both datasets presented relatively low AS commission and omission
325 errors. The omissions levels reached ~10%, while commissions were ~14%, Figure 7.

326 The inspected commission and omission errors were mainly associated with two types of
327 incongruences; 1) U-Net Errors (UE): for the U-net classifier, the context domain is given by the joint
328 use of spectral and spatial information. Thus, lack of training data may fail to successfully
329 transfer/learn the contextual knowledge necessary to decide upon aquaculture or non-aquaculture.
330 2) Filtering Errors (FE): aquaculture and salines ponds may suffer from fluctuations in their water
331 levels. Still, even without water for a short amount of time, the land-use related to an empty artificial
332 pond cannot be classified as non-aquaculture. Although combined with TE, an inaccurate
333 classification may be propagated through time due to the temporal-filter algorithm's action (Figure
334 8).

335 Figure 8 shows the two most representative types of misclassifications. The pair-wise A-A1 and
336 B-B1 demonstrate an inaccurate aqua-salt culture label due to TEs. In A-A1, a portion of a river and
337 natural lagoons were wrongly associated with the aqua-salt culture class. In B-B1, TE's resulted in
338 partially accurate classification. Inner parts of the aqua-salt-culture-dominated cuspaté spit were
339 classified as non-aquaculture. In the C sequence, the water presence/absence is evident through time.
340 In D, the temporal filter acts and uses temporal information to classify empty artificial ponds as
341 aquaculture/salt-culture. Although the temporal rule's simplicity may need to be revisited, seeing
342 that once helping to stabilize the expected water fluctuation, the temporal filter may also propagate
343 errors through time.

344 Thus, despite the overall data agreement and its relatively low commission and omission
345 metrics, there is still room for improvements in the data accuracy levels. In this sense, capturing a
346 more diverse group of training samples in terms of spectral and geometrical variability, increasing

347 the sample's representativity throughout the temporal series, and revisiting the post-classification
 348 algorithms are essential strategies to be further implemented.
 349



350
 351 **Figure 10** – The two most representative types of misclassifications. Due to UEs, the pairs A-A1 and B-B1 indicate
 352 erroneous classification. A river segment and natural lagoons were incorrectly labeled in A-A1. In B-B1, Inner
 353 parts of the aqua-salt culture dominated cusplate spit were classified as a non-aqua-salt culture. In the C and D
 354 sequences, the temporal filter operates and uses temporal information to classify empty artificial ponds as an
 355 aqua-salt culture. Although, while helping to stabilize the expected water fluctuation, the temporal filter may
 356 also propagate errors through time. Blue-dots are accuracy assessment samples.

357

358 The following steps of this research consist of a) improving the post-classification mechanism to
 359 identify empty artificial ponds more accurately, b) expand the analysis to the entire continental region
 360 of the country (including the aqua or salt-culture that may exist further inland), and c) use refined
 361 spatial resolution data. Considering the importance of the contextual domain, the better the data's
 362 spatial resolution is, the better it is for context recognition. Thus, our future works will incorporate
 363 Landsat panchromatic data and the 10m resolution Sentinel-2A full dataset.

364 5. Conclusion

365 Studies of this nature fill up a relevant information gap, exhaustively mapping the entire BCZ
 366 and distinguishing its aqua-salt culture spatial distribution. Using massive, publicly available,
 367 remotely sensed time series, this study contributes to the multi-scale understanding of aquacultures
 368 and salt-culture occupation across the BCZ, which provides a better understanding of Brazil's coastal
 369 dynamics, directly subsidizing national and regional policymakers.

370 This work's findings pave the path towards understanding the direct and indirect influences
 371 that aquaculture and marine salt production might have on the coastal environment's sustainability.
 372 The developed pipeline seems reasonably scalable, sufficiently generic, and suitable for large-scale
 373 aquaculture/saline automatic identification.

374 The method allowed a systematic and continuous delineation of aquaculture and saline patterns,
 375 supporting digital cartographic products and annual statistics, ranging from 1985 to 2019. The data
 376 specialized and quantified the contribution of each one of ZCB's states. Out of 17 possible coastal
 377 states, three did not present any signal of aqua-salt-culture-related patterns (AP, PA, and PR). The
 378 state of Maranhão (MA) presented the smallest detected area, where signs of aquaculture or saline
 379 ponds were more prominently in the 80s.

380 Further developments may allow its applicability to larger continental areas, such as the entire
 381 country, including continental aqua-salt culture or even South America in its entirety, and
 382 incorporate Landsat panchromatic data as well as the Sentinel-2 10m dataset.

383 All the data produced are available through the website <http://www.solved.eco.br/datasets/>
 384 and will be transferred to the MapBiomias project (www.mapbiomas.org) in MapBiomias Collection
 385 6.0.

386

387 **Authors Contribution:** Conceptualization, Cesar Diniz and Luiz Cortinhas; Data curation, Cesar Diniz, Luiz
 388 Cortinhas and Pedro Walfir M. Souza-Filho; Formal analysis, Cesar Diniz; Funding acquisition, Pedro Walfir M.
 389 Souza-Filho; Investigation, Cesar Diniz; Methodology, Cesar Diniz, Luiz Cortinhas and Maria Luize Pinheiro;
 390 Project administration, Pedro Walfir M. Souza-Filho; Supervision, Pedro Walfir M. Souza-Filho; Validation, Luis
 391 Sadeck, Alexandre S. F. Filho, Luis Baumann and Marcos Adami; Visualization, Luis Sadeck; Writing – original
 392 draft, Cesar Diniz; Writing – review & editing, Cesar Diniz, Marcos Adami and Pedro Walfir M. Souza-Filho.

393 **Funding:** his research was funded by the Brazilian National Council for Scientific and Technological
 394 Development (CNPq): 870005/1997-9, 310283/2019-1, 306334/2020-8 and the MapBiomias Project. The processing
 395 charges were covered by the Dean of Research and Graduate Studies (PROPESP) from the Federal University of
 396 Para (UFPA)

397 **Acknowledgments:** The authors are grateful to Tasso Azevedo, general coordinator of the MapBiomias Project
 398 and the SEEG initiative. Lastly, the authors acknowledge the Brazilian National Council for Scientific and
 399 Technological Development (CNPq) for the financial support received during this research and the Dean of
 400 Research and Graduate Studies (PROPESP) from the Federal University of Para (UFPA) for covering all the
 401 publication fees.

402 **Conflicts of Interest:** The authors declare no conflicts of interest.

403 6. References

- 404 1. Food And Agriculture Organization (FAO) *Fishery and Aquaculture Statistics - 2017*; Rome,
 405 2019; ISBN 9789251316696.
- 406 2. Food And Agriculture Organization (FAO) *The State of World Fisheries and Aquaculture 2018 -*
 407 *Meeting the sustainable development goals*; 2018; ISBN 9789251305621.
- 408 3. Food And Agriculture Organization (FAO) *The State of World Fisheries and Aquaculture 2020*;
 409 2020;
- 410 4. Ottinger, M.; Clauss, K.; Kuenzer, C. Aquaculture: Relevance, distribution, impacts and

- 411 spatial assessments – A review. *Ocean Coast. Manag.* **2016**, *119*, 244–266,
412 doi:<https://doi.org/10.1016/j.ocecoaman.2015.10.015>.
- 413 5. Giri, C.; Ochieng, E.; Tieszen, L. L.; Zhu, Z.; Singh, A.; Loveland, T.; Masek, J.; Duke, N. Status
414 and distribution of mangrove forests of the world using earth observation satellite data. *Glob.*
415 *Ecol. Biogeogr.* **2011**, *20*, 154–159, doi:10.1111/j.1466-8238.2010.00584.x.
- 416 6. Thomas, N.; Bunting, P.; Lucas, R.; Hardy, A.; Rosenqvist, A.; Fatoyinbo, T. Mapping
417 Mangrove Extent and Change: A Globally Applicable Approach. *Remote Sens.* **2018**, *10*.
- 418 7. Bunting, P.; Rosenqvist, A.; Lucas, M. R.; Rebelo, L.-M.; Hilarides, L.; Thomas, N.; Hardy, A.;
419 Itoh, T.; Shimada, M.; Finlayson, M. C. The Global Mangrove Watch—A New 2010 Global
420 Baseline of Mangrove Extent. *Remote Sens.* **2018**, *10*.
- 421 8. Diniz, C.; Cortinhas, L.; Nerino, G.; Rodrigues, J.; Sadeck, L.; Adami, M.; Souza-Filho, W. P.
422 Brazilian Mangrove Status: Three Decades of Satellite Data Analysis. *Remote Sens.* **2019**, *11*.
- 423 9. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A. S. High-resolution mapping of global surface
424 water and its long-term changes. *Nature* **2016**, *540*, 418–422.
- 425 10. Luijendijk, A.; Hagenaars, G.; Ranasinghe, R.; Baart, F.; Donchyts, G.; Aarninkhof, S. The State
426 of the World’s Beaches. *Sci. Rep.* **2018**, *8*, 6641, doi:10.1038/s41598-018-24630-6.
- 427 11. Worthington, T. A.; zu Ermgassen, P. S. E.; Friess, D. A.; Krauss, K. W.; Lovelock, C. E.;
428 Thorley, J.; Tingey, R.; Woodroffe, C. D.; Bunting, P.; Cormier, N.; Lagomasino, D.; Lucas, R.;
429 Murray, N. J.; Sutherland, W. J.; Spalding, M. A global biophysical typology of mangroves
430 and its relevance for ecosystem structure and deforestation. *Sci. Rep.* **2020**, *10*, 14652,
431 doi:10.1038/s41598-020-71194-5.
- 432 12. Lagomasino, D.; Fatoyinbo, T.; Lee, S.; Feliciano, E.; Trettin, C.; Simard, M. A Comparison of
433 Mangrove Canopy Height Using Multiple Independent Measurements from Land, Air, and
434 Space. *Remote Sens.* **2016**, *8*.
- 435 13. Lagomasino, D.; Fatoyinbo, T.; Lee, S.; Feliciano, E.; Trettin, C.; Shapiro, A.; Mangora, M. M.
436 Measuring mangrove carbon loss and gain in deltas. *Environ. Res. Lett.* **2019**, *14*, 25002,
437 doi:10.1088/1748-9326/aaf0de.
- 438 14. Duan, Y.; Li, X.; Zhang, L.; Chen, D.; Liu, S.; Ji, H. Mapping national-scale aquaculture ponds
439 based on the Google Earth Engine in the Chinese coastal zone. *Aquaculture* **2020**,
440 doi:10.1016/j.aquaculture.2019.734666.
- 441 15. Xia, Z.; Guo, X.; Chen, R. Automatic extraction of aquaculture ponds based on Google Earth
442 Engine. *Ocean Coast. Manag.* **2020**, *198*, 105348,
443 doi:<https://doi.org/10.1016/j.ocecoaman.2020.105348>.
- 444 16. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image
445 Segmentation. *CoRR* **2015**, *abs/1505.0*.

- 446 17. Arruda, V. L. S.; Piontekowski, V. J.; Alencar, A.; Pereira, R. S.; Matricardi, E. A. T. An
447 alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine,
448 and Deep Learning in the Brazilian Savanna. *Remote Sens. Appl. Soc. Environ.* **2021**, *22*, 100472,
449 doi:<https://doi.org/10.1016/j.rsase.2021.100472>.
- 450 18. Schaeffer-Novelli, Y.; Cintrón-Molero, G.; Adaime, R. R.; de Camargo, T. M.; Cintron-Molero,
451 G.; de Camargo, T. M. Variability of Mangrove Ecosystems along the Brazilian Coast. *Estuaries*
452 **1990**, doi:10.2307/1351590.
- 453 19. Rodrigues, S. W. P.; Souza-Filho, P. W. M. Use of Multi-Sensor Data to Identify and Map
454 Tropical Coastal Wetlands in the Amazon of Northern Brazil. *Wetlands* **2011**, *31*, 11–23,
455 doi:10.1007/s13157-010-0135-6.
- 456 20. Souza Filho, P. W. M. Costa de manguezais de macromaré da Amazônia: cenários
457 morfológicos, mapeamento e quantificação de áreas usando dados de sensores remotos. *Rev.*
458 *Bras. Geofísica* **2005**, *23*, 427–435, doi:10.1590/S0102-261X2005000400006.
- 459 21. Nascimento Jr, W. R.; Souza-Filho, P. W. M.; Proisy, C.; Lucas, R. M.; Rosenqvist, A. Mapping
460 changes in the largest continuous Amazonian mangrove belt using object-based classification
461 of multisensor satellite imagery. *Estuar. Coast. Shelf Sci.* **2013**, *117*, 83–93,
462 doi:<http://dx.doi.org/10.1016/j.ecss.2012.10.005>.
- 463 22. Diniz, C.; Cortinhas, L.; Nerino, G.; Rodrigues, J.; Sadeck, L.; Adami, M.; Souza-Filho, P. W.
464 M. Brazilian mangrove status: Three decades of satellite data analysis. *Remote Sens.* **2019**, *11*,
465 doi:10.3390/rs11070808.
- 466 23. Dominguez, J. M. L. The Coastal Zone of Brazil. In *Geology and Geomorphology of Holocene*
467 *Coastal Barriers of Brazil*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; pp. 17–51 ISBN
468 978-3-540-44771-9.
- 469 24. IBGE Sistema IBGE de recuperação automática (SIDRA) - *Produção da Aquicultura Brasileira*; 2019;
- 470 25. Pereira, J. M. P.; de Souza, E. N. V.; Candido, J. R. B.; Dantas, M. D. A.; Nunes, A. R. D.; Ribeiro,
471 K.; Teixeira, D. I. A.; Lanza, D. C. F. Alternative PCR Primers for Genotyping of Brazilian
472 WSSV Isolates. *J. Invertebr. Pathol.* **2019**, *162*, 55–63, doi:10.1016/j.jip.2019.02.003.
- 473 26. Bandeira, J. de T.; Morais, R. S. M. M. de; Silva, R. P. P. e; Mendes, E. S.; Silva, S. M. B. C. da;
474 Santos, F. L. dos First report of white spot syndrome virus in wild crustaceans and mollusks
475 in the Paraíba River, Brazil. *Aquac. Res.* **2019**, *50*, 680–684, doi:10.1111/are.13949.
- 476 27. Santos, R. N.; M. Varela, A. P.; Cibulski, S. P.; Esmail S. Lima, F.; Rosado Spilki, F.; Schemes
477 Heinzelmann, L.; Bordin da Luz, R.; Abreu, P. C.; Roehe, P. M.; Cavalli, L. S. A Brief History
478 of White Spot Syndrome Virus and its Epidemiology in Brazil. *VIRUS Rev. Res.* **2013**, *18*,
479 doi:10.17525/vrr.v18i1-2.63.
- 480 28. Roubach, R.; Correia, E. S.; Zaiden, S.; Martino, R. C.; Cavalli, R. O. Aquaculture in Brazil.

- 481 *WORLD Aquac. ROUGE-* **2003**, *34*, 28–35.
- 482 29. Tenório, G. S.; Souza-Filho, P. W. M.; Ramos, E. M. L. S.; Alves, P. J. O. Mangrove shrimp farm
483 mapping and productivity on the Brazilian Amazon coast: Environmental and economic
484 reasons for coastal conservation. *Ocean Coast. Manag.* **2015**, *104*, 65–77,
485 doi:<https://doi.org/10.1016/j.ocecoaman.2014.12.006>.
- 486 30. Rocha, I. de P. *Global Aquaculture Alliance*. Portsmouth, UK 2010, pp. 43–45.
- 487 31. Bueno, G. W.; Ostrensky, A.; Canzi, C.; de Matos, F. T.; Roubach, R. Implementation of
488 aquaculture parks in Federal Government waters in Brazil. *Rev. Aquac.* **2015**, *7*, 1–12,
489 doi:[10.1111/raq.12045](https://doi.org/10.1111/raq.12045).
- 490 32. Lima, L. B.; Oliveira, F. J. M.; Giacomini, H. C.; Lima-Junior, D. P. Expansion of aquaculture
491 parks and the increasing risk of non-native species invasions in Brazil. *Rev. Aquac.* **2018**, *10*,
492 111–122, doi:[10.1111/raq.12150](https://doi.org/10.1111/raq.12150).
- 493 33. USGS *Landsat 8 (L8) Data Users Handbook*; 2015;
- 494 34. Storey, J.; Choate, M.; Lee, K. Landsat 8 Operational Land Imager On-Orbit Geometric
495 Calibration and Performance. *Remote Sens.* **2014**, *6*.
- 496 35. Teillet, P. M.; Barker, J. L.; Markham, B. L.; Irish, R. R.; Fedosejevs, G.; Storey, J. C. Radiometric
497 cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets.
498 *Remote Sens. Environ.* **2001**, *78*, 39–54, doi:[https://doi.org/10.1016/S0034-4257\(01\)00248-6](https://doi.org/10.1016/S0034-4257(01)00248-6).
- 499 36. USGS Landsat Collection 1 Level 1 Product Definition. **2017**, *26*.
- 500 37. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water
501 features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033,
502 doi:[10.1080/01431160600589179](https://doi.org/10.1080/01431160600589179).
- 503 38. Olofsson, P.; Foody, G. M.; Herold, M.; Stehman, S. V.; Woodcock, C. E.; Wulder, M. A. Good
504 practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.*
505 **2014**, *148*, 42–57, doi:[10.1016/j.rse.2014.02.015](https://doi.org/10.1016/j.rse.2014.02.015).
- 506 39. Pontius, R. G.; Santacruz, A. Quantity, exchange, and shift components of difference in a
507 square contingency table. *Int. J. Remote Sens.* **2014**, *35*, 7543–7554,
508 doi:[10.1080/2150704X.2014.969814](https://doi.org/10.1080/2150704X.2014.969814).
- 509 40. Stehman, S. V. Estimating area and map accuracy for stratified random sampling when the
510 strata are different from the map classes. *Int. J. Remote Sens.* **2014**, *35*, 4923–4939,
511 doi:[10.1080/01431161.2014.930207](https://doi.org/10.1080/01431161.2014.930207).
- 512 41. Diniz, M. T. M.; Vasconcelos, F. P.; Diniz, M. T. M.; Vasconcelos, F. P. Natural Conditions for
513 the Sea Salt Production In Brazil. *Mercator* **2017**, *16*, 1–19, doi:[10.4215/rm2017.e16013](https://doi.org/10.4215/rm2017.e16013).

- 514 42. Costa, D. F. da S.; Silva, A. A. da; Medeiros, D. H. M.; Lucena Filho, M. A.; Rocha, R. D. M.;
515 Lillebo, A. I.; Soares, A. M. V. M. Breve revisão sobre a evolução histórica da atividade
516 salineira no estado do Rio Grande do Norte (Brasil) . *Soc. Nat.* 2013, 25, 21–34.
- 517 43. Diniz, M. T. M.; Vasconcelos, F. P.; Martins, M. B. Inovação Tecnológica na Produção Brasileira
518 se Sal Marinho e as Alterações Sócioterritoriais dela Decorrentes: Uma Análise Sob a Ótica da
519 Teoria do Empreendedorismo de Schumpeter . *Soc. Nat.* 2015, 27, 421–437.
- 520 44. Bhat, A. H.; Sharma, K. C.; Banday, U. J. Impact of Climatic Variability on Salt Production in
521 Sambhar Lake, a Ramsar Wetland of Rajasthan, India. *Middle-East J. Sci. Res.* 2015, 23, 2060–
522 2065.
- 523 45. Lima, K. C. Impactos Econômicos das Mudanças Climáticas sobre a Indústria de Sal Marinho
524 na Principal Região Produtora do Brasil. *Rev. Bras. Geogr. Física; v. 10, n. 2 Rev. Bras. Geogr.*
525 *Física* 2017, doi:10.5935/1984-2295.20170037.
- 526 46. ICMBio *Atlas dos manguezais do Brasil*; ICMBio, Ed.; 1st ed.; 2017; ISBN 978-85-61842-75-8.
- 527



© 2018 by the authors. Submitted for possible open access publication under the
terms and conditions of the Creative Commons Attribution (CC BY) license
(<http://creativecommons.org/licenses/by/4.0/>)

530

4. CONCLUSÕES, CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

Esta tese de doutorado detalhou um fluxo de processamento de dados, gerenciados pela plataforma Google Earth Engine e pela biblioteca TensorFlow, para discriminar florestas de manguezais da vegetação circundante, bem como a discriminação de padrões salineiros e aquícolas de outras superfícies costeiras à água relacionadas. O fluxo desenvolvido é escalável e adequado para análises costeiras em larga escala, permitindo o mapeamento sistemático e contínuo da cobertura de manguezais brasileiros, juntamente com dois dos usos costeiros mais clássicos da Terra, as lagoas artificiais de aquicultura e produção de sal.

Essa pesquisa produziu mapas e estatísticas anuais que vão de 1985 a 2019. Quando comparados a pesquisas anteriores em escala global e nacional, e a referências visualmente interpretadas, os dados produzidos apresentam altos níveis de concordância espacial e estatística. Todos os dados produzidos estão disponíveis para download no site <http://www.solved.eco.br/datasets> e foram transferidos para o projeto MapBiomias (www.mapbiomas.org) em sua Coleção 4.0. As atualizações aqui constituídas, relativas a superfícies de aquícolas e salineiras, bem como dos avanços científicos alcançados de processamento e pós-processamento de grandes volumes de dados, serão incorporados à Coleção 6.0 do projeto.

O índice espectral desenvolvido, MMRI, demonstrou ser robusto e aumentar a separabilidade espectral entre as classes de manguezal e não-manguezal. Aparentemente, a expansão da janela temporal para a construção de mosaicos mais amplos, a restrição do padrão climático (chuvoso ou seco) e a capacidade de discernir entre os regimes de marés predominantes, sejam altos ou baixos, podem ajudar a reduzir algumas das flutuações de área associadas a cobertura do manguezal.

Estudos desta natureza preenchem uma lacuna de informação relevante, ao mapear exaustivamente todo a BCZ e distinguir a seus parques aquícolas e salineiros, trouxe à luz um dado primário e de grande relevância costeira, que era até então, inexistente. Usando extensas séries temporais de dados orbitais de média resolução e publicamente disponíveis, esse estudo contribui para a compreensão, em várias escalas, da evolução da ocupação costeira pela atividade aquícola-salineira, o que fornece uma melhor compreensão da dinâmica costeira do Brasil, subsidiando diretamente os formuladores de políticas nacionais e regionais.

Os dados espacializaram e quantificaram a extensão das áreas costeiras de cada um dos estados da ZCB ocupadas por tanques de aquicultura/salicultura. Confirmando que, dos 17 estados costeiros do país, apenas três não apresentaram qualquer padrão relacionados à

atividade aquícola ou a produção de sal por evaporação (AP, PA e PR). Já o estado do Maranhão (MA), apresentou a menor área detectada, onde na década de 80 indícios de infraestrutura aquícola ou salineira eram mais marcantes.

Nossas descobertas ajudam a pavimentar o caminho para a compreensão das influências diretas e indiretas que a aquíicultura e a produção de sal marinho, podem ter na sustentabilidade do ambiente costeiro. Desenvolvimentos posteriores podem permitir sua aplicabilidade a áreas continentais maiores, como o país inteiro, incluindo a cultura de sal aquático continental ou mesmo a América do Sul em sua totalidade, e incorporar dados pancromáticos Landsat, bem como o conjunto de dados Sentinel-2 10m.

REFERÊNCIAS

- Adami M. *et al.* 2012. Remote sensing time series to evaluate direct land use change of recent expanded sugarcane crop in Brazil. *Sustainability*, **4** (4):574–585.
- Alencar A. *et al.* 2020. Mapping three decades of changes in the Brazilian savanna native vegetation using landsat data processed in the Google Earth Engine Platform. *Remote Sensing*, **12**(6): 924. Mar. <https://doi.org/10.3390/rs12060924>.
- Alongi D. M. 2002. Present state and future of the world's mangrove forests. *Environmental Conservation*, **29** (03): 331–349, 13 set.
- Alongi D. M. 2009. *The energetics of mangrove forests*. 1. ed. Dordrecht, Springer Netherlands.
- Alsaaidh B. *et al.* 2013. Mangrove forests mapping in the Southern Part of Japan using Landsat ETM+ with DEM. *Journal of Geographic Information System*, **05** (4): ID:35346. DOI:10.4236/jgis.2013.54035.
- ANM. *Sumário Mineral Brasileiro - Sal*. [s.l: s.n.]. Disponível em: https://www.gov.br/anm/pt-br/centrais-de-conteudo/publicacoes/serie-estatisticas-e-economia-mineral/sumario-mineral/pasta-sumario-brasileiro-mineral-2018/sal_sm_2018. Acesso em: 20 Novembro 2020.
- Arruda V. L. S. *et al.* 2021. An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna. *Remote Sensing Applications: Society and Environment*, **22**:100472.
- Baatz M. 2000. Multi resolution segmentation: an optimum approach for high quality multi scale image segmentation. In: Beutrage zum AGIT-Symposium. Salzburg, Heidelberg, *Anais[...]*, p. 12-23.
- Barbosa C. C. F., Moraes Novo E. M. L. de, Martins V. S.(eds). 2019. *Introdução ao sensoriamento remoto de sistemas aquáticos: princípios e aplicações*. 1. ed. São José dos Campos, Instituto Nacional de Pesquisas Espaciais.
- Belgiu M. & Drăguț L. 2016. Random forest in remote sensing: a review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, **114**: 24–31.
- Bogner C. *et al.* 2018. Classification of rare land cover types: distinguishing annual and perennial crops in an agricultural catchment in South Korea. *Plos One*, **13**: (1): e0190476, 25 jan.
- Breiman L. 2001. Random forests. *Machine Learning*, **45** (1): 5–32.
- Bunting P. *et al.* 2018. The global mangrove watch—a new 2010 global baseline of mangrove extent. *Remote Sensing*, **10** (10): 1669. <https://doi.org/10.3390/rs10101669>.
- Burke L. *et al.* *Pilot analysis of global ecosystems: coastal ecosystems*. Washington, DC (USA) WRI, , 2000. Disponível em: [http://agris.fao.org/agris-search/search.do?recordID=XF2015025829#.WTHIvEAWVgw.mendeley&title=Pilot analysis of global ecosystems: coastal ecosystems&description=Looking for Agricultural Science and Technology Information?](http://agris.fao.org/agris-search/search.do?recordID=XF2015025829#.WTHIvEAWVgw.mendeley&title=Pilot%20analysis%20of%20global%20ecosystems%3A%20coastal%20ecosystems&description=Looking%20for%20Agricultural%20Science%20and%20Technology%20Information?). Acesso em: 2 jun. 2017.

- Chander G., Markham B. L. Helder D. L. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, **113**(5): 893–903.
- Chen X. *et al.* 2017. Running climate model on a commercial cloud computing environment: A case study using Community Earth System Model (CESM) on Amazon AWS. *Computers & Geosciences*, **98** (Supplement C): 21–25.
- Costa D. F. da S. *et al.* 2013. Breve revisão sobre a evolução histórica da atividade salineira no estado do Rio Grande do Norte (Brasil). *Sociedade & Natureza*, **25**(1): 21–34.
- Diniz C. *et al.* 2019. Brazilian mangrove status: three decades of satellite data analysis. *Remote Sensing*, 11 (7): 808. <https://doi.org/10.3390/rs11070808>.
- Diniz M. T. M. *et al.* 2017. Natural conditions for the sea salt production in Brazil. *Mercator*, 16 (5): 1–19, 15 maio.
- Dominguez J. M. L. 2009. The Coastal Zone of Brazil. In: Dillenburg S. R. & Hesp P. A. *Geology and geomorphology of Holocene Coastal Barriers of Brazil*. Berlin, Heidelberg, Springer Berlin Heidelberg, p. 17–51. (Lecture Notes in Earth Sciences, 107).
- Donato D. C. *et al.* 2011. Mangroves among the most carbon-rich forests in the tropics. *Nature Geosci*, **4**(5): 293–297.
- Emery W. & Camps A. 2017. The history of satellite remote sensing. In: Emery W. Camps A. *Introduction to satellite remote - atmosphere, ocean, land and cryosphere applications*. Amsterdam, Netherlands, Elsevier, Chapter 1 , p. 1–42.
- Falk T. *et al.* 2019. U-Net: deep learning for cell counting, detection, and morphometry. *Nature Methods*, 16 (1): 67–70..
- Fei S. X., Shan C. H., Hua G. Z. 2011. Remote sensing of mangrove wetlands identification. *Procedia Environmental Sciences*, **10**: 2287–2293.
- Finegold, Y. *et al.* 2016. *Map accuracy assessment and area estimation: a practical guide*. Rome, Food and Agriculture Organization of the United Nations.
- Food and Agriculture Organization (FAO). 2018. *The state of world fisheries and aquaculture - meeting the sustainable development goals*. Rome, Italy (Série La situation mondiale des pêches et de l'aquaculture -SOFIA).
- Food and Agriculture Organization (FAO). 2020. *The state of world fisheries and aquaculture - sustainability in action*. Rome. <https://doi.org/10.4060/ca9229en>.
- Foody G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, **80** (1):185–201.
- Foody G. M. 2009. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sensing of Environment*, **113** (8):1658–1663.
- Forgy E. W. 1965. Cluster analysis of multivariate data: efficiency versus interpretability of

classifications. *Biometrics*, **21**: 768–769.

Fuller R. M., Groom G. B., Jones A. R. 1994. The land cover map of Great Britain: An automated classification of Landsat thematic mapper data. *Photogrammetric Engineering and Remote Sensing*, **60** (5): 553–562.

Gao B. 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, **58** (3): 257–266.

Ghimire B. *et al.* 2012. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing*, **49** (5): 623–643.

Giri C. *et al.* 2013. Global mangrove forests distribution, 2000 Palisades, Ny, Nynasa Socioeconomic Data and Applications Center (SEDAC). Disponível em: <http://dx.doi.org/10.7927/H4J67DW8>. Acesso em: 27 de Janeiro de 2019.

Giri C. 2016. Observation and monitoring of mangrove forests using remote sensing: opportunities and challenges. *Remote Sensing*, **8** (9): 783. <https://doi.org/10.3390/rs8090783>.

GIRI, C. *et al.* 2011. Status and distribution of mangrove forests of the world using earth observation satellite data. *Global Ecology and Biogeography*, **20** (1): 154–159.

Gómez C., White J. C., Wulder M. A. 2016. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116: 55–72.

Gonzalez R. C. & Woods R. E. 2008. *Digital image processing*. 3rd. Upper Saddle River, N.J., Prentice Hall. p. 743-747.

Gorelick N. *et al.* 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, **202** (Supplement C): 18–27, Dec.

Gu Y. *et al.* 2007. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical Research Letters*, **34** (6): L06407, doi:10.1029/2006GL02912.

Gupta K. *et al.* 2018. An index for discrimination of mangroves from non-mangroves using Landsat 8 OLI imagery. *MethodsX*, **5**: 1129–1139, 28 set.

Guzha A. C. *et al.* 2018. Impacts of land use and land cover change on surface runoff, discharge and low flows: Evidence from East Africa. *Journal of Hydrology: Regional Studies*, **15**: 49–67.

Hansen M. C. *et al.* 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, **21**(6–7): 1331–1364.

Instituto Brasileiro de Geografia e Estatística - IBGE. 2019a. *Biomass e sistema costeiro-marinho do Brasil* : compatível com a escala 1:250000. Rio de Janeiro. p. 125-138 ID: 101676.

Instituto Brasileiro de Geografia e Estatística - IBGE. 2019b. *Sistema IBGE de recuperação automática (SIDRA)* - Produção da Aquicultura Brasileira. [S.l.].

Instituto Chico Mendes de Conservação da Biodiversidade - ICMBIO. 2017. *Atlas dos manguezais do Brasil*. 1. ed. Brasília,DF.

Jähne B. 2012. *Digital image processing and image formation*. Heidelberg, Neckar, Springer Berlin.

Jensen J. R. 2014. *Remote sensing of the environment: an earth resource perspective*. 2nd ed. [S.l.], Pearson Education India. v. 1

Jin B. *et al.* 2019. Object-oriented method combined with deep convolutional neural networks for land-use-type classification of remote sensing images. *Journal of the Indian Society of Remote Sensing*, **47**(6): 951–965.

Lary D. J. *et al.* 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, **7**(1): 3–10.

Lima K. C. 2017. Impactos econômicos das mudanças climáticas sobre a indústria de sal marinho na principal região produtora do Brasil. *Revista Brasileira de Geografia Física*, **10**(2): 584-596.

Liu H. Q. & Huete A. 1995. Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, **33** (2): 457 - 465. Mar.

Magris R. A. & Barreto R. 2010. Mapping and assessment of protection of mangrove habitats in Brazil. *Pan-American Journal of Aquatic Sciences*, **5**(4): 546–556.

Marinho R. R. *et al.* 2021. Spatiotemporal dynamics of suspended sediments in the Negro river, Amazon basin, from in situ and sentinel-2 remote sensing data. *ISPRS International Journal of Geo-Information*, **10**(2): 86; <https://doi.org/10.3390/ijgi10020086>.

Marinho R. R., Filizola Junior N. P., Cremon É. H. 2020. Analysis of suspended sediment in the Anavilhanas archipelago, rio Negro, Amazon basinwater. *Water*, **12**(4): 1073; <https://doi.org/10.3390/w12041073>.

Markham B. L. *et al.* 2018. 1.03 - Landsat program. In: Liang S. B. T. (ed.). *Comprehensive remote sensing*. Oxford, Elsevier, p. 27–90.

Masek J. G. *et al.* 2008. North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, **112** (6): 2914–2926.

Masek J. G. *et al.* 2020. Landsat 9: empowering open science and applications through continuity. *Remote Sensing of Environment*, **248**: 111968.

Maxwell A. E., Warner T. A., Fang F. 2018. Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*, **39** (9):2784–2817, 3 maio.

McConnell W. J. 2015. Land change: the merger of land cover and land use dynamics A2. In: Wright J. D. *International Encyclopedia of the Social & Behavioral Sciences*. 2nd ed. Oxford, Elsevier, p. 220–223.

- Mcfeeters S. K. 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, **17**(7): 1425–1432, maio.
- Mertes L. A. K., Smith M. O., Adams J. B. 1993. Estimating suspended sediment concentrations in surface waters of the Amazon River wetlands from Landsat images. *Remote Sensing of Environment*, **43** (3): 281–301.
- Murdiyarso D. *et al.* 2015. The potential of Indonesian mangrove forests for global climate change mitigation. *Nature Clim. Change*, **5** (12): 1089–1092.
- Nardin W. *et al.* 2016. Dynamics of a fringe mangrove forest detected by Landsat images in the Mekong River Delta, Vietnam. *Earth Surface Processes and Landforms*, **41** (14): 2024–2037.
- Nascimento Jr. W. R. *et al.* 2013. Mapping changes in the largest continuous Amazonian mangrove belt using object-based classification of multisensor satellite imagery. *Estuarine, Coastal and Shelf Science*, **117**: 83–93.
- Neumann B. *et al.* 2015. Future coastal population growth and exposure to sea-level rise and coastal flooding - a global assessment. *Plos One*, **10** (3): e0118571, 11 mar.
- Nicolodi J. L. & Petermann R. M. 2010. Potential vulnerability of the Brazilian coastal zone in its environmental, social, and technological aspects. *Pan-American Journal of Aquatic Sciences*, **5**(2): 12-32.
- Novo E. M. L. de M. 2010. *Sensoriamento remoto: principios e aplicações*. São Paulo (SP), Blucher.
- Oertel C. *et al.* 2016. Greenhouse gas emissions from soils—a review. *Geochemistry*, **76** (3):327–352.
- Olofsson P. *et al.* 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, **148**: 42–57.
- Ottinger M., Clauss K., Kuenzer C. 2016. Aquaculture: Relevance, distribution, impacts and spatial assessments – a review. *Ocean & Coastal Management*, **119**: 244–266.
- Parente L. *et al.* 2019. Next generation mapping: combining deep learning, cloud computing, and big remote sensing data. *Remote Sensing*, **11**(23): 2881. <https://doi.org/10.3390/rs11232881>.
- Pekel J.-F. *et al.* 2016. High-resolution mapping of global surface water and its long-term changes. *Nature*, **540** (7633): 418–422, 15 dez.
- Phiri D. & Morgenroth J. 2017. Developments in landsat land cover classification methods: a review. *Remote Sensing*, **9**(9): 967; <https://doi.org/10.3390/rs9090967>..
- Pontius R. G. & Millones M. 2011. Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, **32** (15): 4407–4429, 10 ago.
- Pontius R. G. & SantaCruz A. 2014. Quantity, exchange, and shift components of difference in

- a square contingency table. *International Journal of Remote Sensing*, **35** (21): 7543–7554.
- Queiroz L. *et al.* 2013. Shrimp aquaculture in the federal state of Ceará, 1970–2012: trends after mangrove forest privatization in Brazil. *Ocean & Coastal Management*, **73**: 54–62.
- Rocchio L. E. P. *et al.* 2018. *Landsat benefiting society for fifty years*. Nasa. Disponível em: <http://pubs.er.usgs.gov/publication/70201780>. Acesso em: ????
- Rogers A. S. & Kearney M. S. 2004. Reducing signature variability in unmixing coastal marsh Thematic Mapper scenes using spectral indices. *International Journal of Remote Sensing*, **25**(12): 2317–2335.
- Ronneberger O., Fischer P., Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation. *In: International Conference on Medical image computing and computer-assisted intervention- CoRR*. Cham, Springer, p. 234–241. v. abs/1505.0. (LNCS, ,9351).
- Rosa M. R. *et al.* 2021. Hidden destruction of older forests threatens Brazil’s Atlantic Forest and challenges restoration programs. *Science Advances*, **7** (4): eabc4547, 1 jan.
- Rudorff B. F. T. *et al.* 2010. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo state (Brazil) using landsat data. *Remote Sensing*, **2**: 1057–1076.
- Saenger P., Hegerl E. J., Davie J. D. S. (ed.). 1983. *Global status of mangrove ecosystems*. Gland, Switzerland, International Union for Conservation of Nature and Natural Resources. (Commission on Ecology Papers Number 3).
- Salazar A. *et al.* 2015. Land use and land cover change impacts on the regional climate of non-Amazonian South America: a review. *Global and Planetary Change*, **128**: 103–119.
- Sanderman J. *et al.* 2018. A global map of mangrove forest soil carbon at 30 m spatial resolution. *Environmental Research Letters*, **13** (5): 55002, 1 maio.
- Saraiva M. *et al.* 2020. Automatic mapping of center pivot irrigation systems from satellite images using deep learning. *Remote Sensing*, **12**(3): 558; <https://doi.org/10.3390/rs12030558>.
- Sasmito S. D. *et al.* 2019. Effect of land-use and land-cover change on mangrove blue carbon: A systematic review. *Global Change Biology*, **25** (12): 4291–4302, 1 dez.
- Schaeffer-Novelli Y. *et al.* 1990. Variability of mangrove ecosystems along the Brazilian coast. *Estuaries*, **13**(2): 204–218.
- Small C. & Nicholls R. J. 2003. A global analysis of human settlement in coastal zones. *Journal of Coastal Research*, **19**(3): 584–599.
- Souza C. M. *et al.* 2019. Long-term annual surface water change in the brazilian amazon biome: potential links with deforestation, infrastructure development and climate change. *Water*, **11**(3): 566. <https://doi.org/10.3390/w11030566>.
- Souza C. M. *et al.* 2020. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing*, **12**(17): 2735. <https://doi.org/10.3390/rs12172735>.

- Souza Filho P. W. M. 2005. Costa de manguezais de macromaré da Amazônia: cenários morfológicos, mapeamento e quantificação de áreas usando dados de sensores remotos. *Revista Brasileira de Geofísica*, **23** (4): 427–435, dez.
- Spalding M., Kainuma M., Collins L. 2010. *World atlas of mangroves*. [S.l.], Taylor & Francis Group.
- Stehman S. V. 2014. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *International Journal of Remote Sensing*, **35** (13): 4923–4939, 3 jul.
- Storey J., Choate M., Lee K. 2014. Landsat 8 operational land imager on-orbit geometric calibration and performance. *Remote Sensing*, **6**(11): 11127-11152.
- Tahmasebi P. *et al.* 2020. Machine learning in geo- and environmental sciences: From small to large scale. *Advances in Water Resources*, **142**: 103619.
- Teillet P. M. *et al.* 2001. Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sensing of Environment*, **78** (1):39–54,
- Tenório G. S. *et al.* 2015. Mangrove shrimp farm mapping and productivity on the Brazilian Amazon coast: Environmental and economic reasons for coastal conservation. *Ocean & Coastal Management*, **104**: 65–77.
- Tong P. H. S. *et al.* 2004. Assessment from space of mangroves evolution in the Mekong Delta, in relation to extensive shrimp farming. *International Journal of Remote Sensing*, **25** (21): 4795–4812.
- Tucker C. J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, **8** (2): 127–150.
- U.S. Geological Survey USGS. ano. Landsat 8 (L8) data users handbook. Earth Resources Observation and Science (EROS) Center. [S.l.].
- U.S. Geological Survey USGS. 2017.Landsat collection 1 level 1 product definition. Earth Resources Observation and Science (EROS) Center. [S.l.], p. 26.
- Vogelmann J. E. *et al.* 2001. Completion of the 1990s national land cover data set for the conterminous united states from landsat thematic mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, **67** (6): 650–662.
- Vuuren D. P. van *et al.* 2017. Energy, land-use and greenhouse gas emissions trajectories under a green growth paradigm. *Global Environmental Change*, **42** : 237–250.
- Xu H. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, **27** (14): 3025–3033, 20 jul.
- Zhang, C. *et al.* An object-based convolutional neural network (OCNN) for urban land use classification. **Remote Sensing of Environment**, v. 216, p. 57–70, 2018.
- Zhu Z., Wang S., Woodcock C. E. 2015. Improvement and expansion of the Fmask algorithm:

cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. **Remote Sensing of Environment**, 159 (Supplement C): 269–277, 15 mar.



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOLOGIA E GEOQUÍMICA

PARECER
Sobre a Defesa Pública da Tese de Doutorado de
CESAR GUERREIRO DINIZ

A banca examinadora da Tese de Doutorado de **CESAR GUERREIRO DINIZ** orientando do Prof. Dr. Pedro Walfir Martins e Souza Filho (UFPA), composta pelos professores doutores David Lagomasino – ECU/NASA, Carlos Souza Junior - Imazon, Arnaldo Queiroz da Silva – UFPA e Nils E. Asp Neto - UFPA, após apresentação da sua tese intitulada **“TRÊS DÉCADAS DE MUDANÇAS NA PLANÍCIE COSTEIRA BRASILEIRA: O STATUS DOS MANGUEZAIS, DA AQUICULTURA E SALICULTURA A PARTIR DE SÉRIES TEMPORAIS LANDSAT E TÉCNICAS DE APRENDIZADO DE MÁQUINA”**, emite o seguinte parecer:

A banca examinadora reconhece a importante contribuição do trabalho para o mapeamento de manguezais, aquicultura e salicultura utilizando séries temporais de dados Landsat, algoritmos de aprendizagem de máquina e de aprendizagem profunda (*Deep Learning*) em computação em nuvem. O estudo traz um novo índice espectral para o mapeamento de manguezais. Essas contribuições já foram publicadas em periódicos especializados.

A banca recomenda que a tese destaque nas conclusões a relevância científica dos resultados deste estudo e suas potenciais aplicações na área ambiental e no sensoriamento remoto. Recomenda-se incluir também uma melhor contextualização dos processos geológicos e oceanográficos em escala continental e global.

Finalmente, a banca examinadora decidiu por unanimidade aprovar a tese de doutorado.

Belém, 31 de março de 2021.

Prof. Dr. Pedro Walfir Martins e Souza Filho
(Orientador – UFPA)

Prof. Dr. David Lagomasino
(Membro externo – East Carolina University/NASA)

Prof. Dr. Arnaldo de Queiroz da Silva
(Membro externo-UFPA)

Prof. Dr. Carlos Souza Junior
(Membro externo-Imazon)

Prof. Dr. Nils E. Asp Neto
(Membro interno-UFPA)