

## **META 4: VISUALIZAÇÃO E ANÁLISE DE DADOS GEOESPACIAIS VIA GOOGLE EARTH ENGINE**

### **4.1 Desenvolver uma ferramenta de visualização e análise de dados geoespaciais, utilizando a tecnologia Google Earth Engine**

#### **4.1.1 Ferramentas de Desenvolvimento**

A principal ferramenta de desenvolvimento utilizada no CASSIE é o Google Earth Engine (GEE), que possibilita o acesso aos dados necessários e oferece todo o potencial de processamento paralelo em nuvem da Google Cloud. Os datasets disponíveis no GEE possibilitam à comunidade científica realizar análises em escala global com disponibilidade de dados de décadas.

O CASSIE foi construído como uma interface gráfica amigável que realiza a comunicação entre o usuário e o GEE. É caracterizado como uma aplicação web, sendo programado com ECMAScript (popularmente conhecido como Javascript), utilizando a biblioteca modular *React* para construir os componentes visuais e o Google Maps como plataforma para visualização de mapa e imagens de satélite, servindo como fonte principal de entrada e saída de dados de regiões espaciais e formas geométricas.

Adicionalmente, foram utilizadas as bibliotecas Redux, para gerenciar o estado da aplicação, possibilitando que o acesso aos dados pelos componentes do CASSIE seja simplificado e; Redux-Saga para controlar os eventos assíncronos, como a comunicação com o Google Earth Engine.

Utilizou-se, ainda, uma biblioteca da Mapbox, a *shp-write*, para possibilitar que os dados construídos e analisados com o CASSIE sejam exportáveis para o formato da ESRI, a *shapefile*, um formato popular de dados geoespaciais em forma de vetor utilizados por Sistemas de Informações Geográficas (*GIS*).

#### **4.1.2 Funcionalidades**

Nesta seção, são explorados as funcionalidades implementadas no CASSIE, abrangendo desde a importação dos Datasets do Google Earth Engine até a criação de transectos e análise de linha de costa.

#### 4.1.2.1 Importação de Datasets

O CASSIE possibilita o acesso a todas as imagens SR (Reflectância de Superfície) Tier 1 do Landsat 5 (TM), Landsat 7 (ETM+) e Landsat 8 (OLI) e está no processo de implementação da coleção dos produtos Sentinel 2 (MSI) Level-1C. A Tabela 1 resume os dados de satélite disponíveis no GEE para as missões supracitadas.

As coleções do Landsat SR corrigem efeitos da geometria atmosférica e de iluminação e visualização, enquanto que os produtos do Sentinel 2 Level-1C são oferecidos como reflectância do Topo da Atmosfera (TOA) e possuem parâmetros para transformação destes em radiância.

Missão	Produto	Período	Ciclo (dias)	Resolução (m)
Landsat 5 (TM)	SR	1984-2013	15	30
Landsat 7 (ETM+)	SR	1999-presente	15	30
Landsat 8 (OLI)	SR	2013-presente	15	30
Sentinel 2 (MSI)	TOA	2015-presente	5	10

Tabela 1. Datasets importados do GEE

As missões de satélite possuem diferenças nas propriedades e bandas de imagens, bem como diferentes procedimentos para realização de composição e aquisição, que impossibilitam que missões diferentes sejam tratadas uniformemente. O CASSIE contorna estas diferenças através da implementação do mecanismo de abstração de herança da Programação Orientada a Objetos (POO). O diagrama de classes referente aos Landsats é observado na Figura 1.

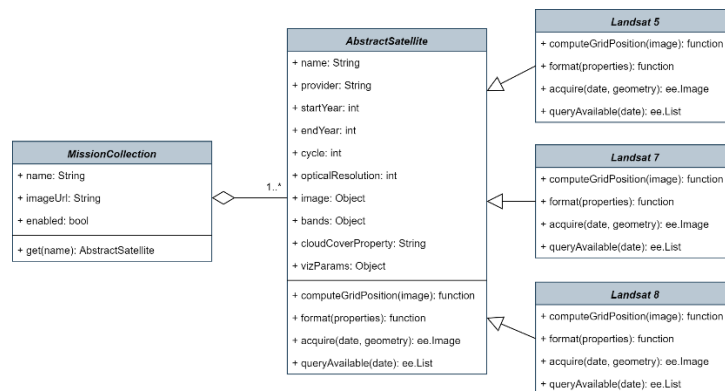


Figura 1 Diagrama de Classes das coleções de satélite implementadas no CASSIE.

Através da herança, os algoritmos utilizados no pré-processamento, processamento e operações com as bandas da imagem, não precisam conhecer os procedimentos específicos da missão, pois estes foram abstraídos.

#### *4.1.2.2 Definição de Área de Interesse*

O tamanho das requisições de imagens no GEE é reduzido pelo CASSIE através da definição do escopo da análise, uma área de interesse (AOI) demarcada pelo usuário que é utilizada para realizar o corte das imagens ainda no GEE, reduzindo o tempo de processamento e o uso de rede.

O CASSIE utiliza a biblioteca Drawing presente no Google Maps para permitir que o usuário defina a AOI por meio de um conjunto de ferramentas de desenho. Após a inserção, são armazenadas as coordenadas geográficas no estado da aplicação para posterior uso no pré-processamento das imagens.

#### *4.1.2.3 Pré-processamento*

O pré-processamento consiste no conjunto de etapas para realizar a aquisição das imagens num formato específico e adequado para possibilitar que qualquer região no planeta possa ser analisada e para que o usuário possa escolher quais imagens devem ser utilizadas no processo de extração e análise de linha de costa. Os resultados do pré-processamento são imagens que passaram pelo processo de registro, composição e detecção de nuvens; que correspondem exatamente à área de interesse e possuem um indicador de cobertura de nuvens.

##### *4.1.2.3.1 Registro*

A operação de registro é uma transformação geométrica que relaciona as coordenadas da imagem (linha e coluna) com coordenadas geográficas (longitude e latitude) de um mapa. A transformação resulta na eliminação de distorções existentes na imagem causadas no processo de formação da imagem.

O registro de imagens é realizado no CASSIE utilizando um par de funções da API do GEE: *displacement* e *displace*. A primeira função é responsável por calcular o movimento horizontal e vertical que deve ser aplicado a cada pixel com base em uma imagem de referência. A segunda função recebe o conjunto de movimentos e efetua a distorção da imagem.

Para um conjunto de imagens, é necessário eleger uma como referência e caminhar pela coleção, calculando o movimento e aplicando a distorção.

#### 4.1.2.3.2 Composição

A composição de imagens no CASSIE refere-se ao procedimento também chamado de mosaicamento, que consiste na montagem espacial de um conjunto de dados de imagens para produzir uma imagem espacialmente contínua.

A cobertura planetária de uma missão de satélite é realizada através da captura de imagens sobre uma extensão geográfica limitada. Cada imagem refere-se a somente uma célula (*tile*) de uma grade de cobertura. Portanto, torna-se necessário compôr mais de uma imagem quando a área de interesse é extensa ou simplesmente abrange mais de uma célula por estar sobre bordas da grade.

O algoritmo de composição desenvolvido no CASSIE particiona o conjunto de imagens em períodos de revisitação do satélite e, sem seguida, realiza a verificação de presença de todas as imagens necessárias para preencher a área de interesse; se uma das imagens não estiver disponível, a partição referente a data da imagem é descartada. A montagem espacial das imagens é realizada através de uma função disposta pelo GEE sobre conjuntos de imagens (*ImageCollection mosaic*).

#### 4.1.2.3.3 Detecção de Nuvens

As imagens do conjunto de dados presentes no GEE já estão avaliadas com a porcentagem de cobertura de nuvens. A informação de cobertura é, entretanto, para a imagem inteira referente à uma célula de captura. Ao analisar imagens no CASSIE, o utilizador raramente selecionará a região exata de uma célula, ao invés disto, é comum realizar análises em uma região menor que a célula ou em regiões em intersecções de células, como descrito na seção de composição.

Para analisar a cobertura de nuvens de imagens da área de interesse, portanto, o CASSIE realiza a razão da contagem de pixels classificados como nuvem pela quantidade total de pixels, para cada uma das imagens.

A classificação do pixel como nuvem é um processo que utiliza uma banda denominada *Quality Assessment* (*pixel\_qa* nos Landsats), calculada pelo algoritmo *Function of Mask* em C para cada pixel da imagem. O pixel consiste em um valor de

máscara de bits, cujas posições estão preenchidas com algum aspecto avaliado pelo algoritmo, detalhado na Tabela 2.

Bit	Aspecto
0	Nenhum
1	Limpo
2	Água
3	Sombra de Nuvem
4	Neve
5	Nuvem
6 a 7	00: nenhuma confiança 01: confiança baixa 10: confiança média 11: confiança alta

*Figura 2 Aspectos da máscara de bits do Fmask*

O pixel é classificado com existência de nuvem (valor 1) caso haja confiança média ou alta e o se o terceiro ou quinto bit tiverem o valor alto. A porcentagem de cobertura de nuvens da imagem é inferida, portanto, através da razão da soma da classificação de todos os pixels (com valores de 0 ou 1) pela quantidade de pixels existentes na imagem, obtida pela divisão da área do polígono pela escala do pixel. O resultado da operação é salvo como uma propriedade na imagem, para que seja utilizado posteriormente em filtros realizados pelo usuário.

#### *4.1.2.4 Filtro de imagens*

O CASSIE dispõe o resultado do pré-processamento ao usuário para realização de um filtro em duas etapas:

1. Seleção do intervalo de tempo desejado e o limiar de cobertura de nuvens desejado.
2. Remoção manual de imagens indesejadas, de acordo com o julgamento do usuário sobre as miniaturas das mesmas.

As imagens que atenderem os critérios estabelecidos na primeira etapa e as que não forem descartadas na segunda etapa são mantidas no estado do CASSIE para

interação futura do usuário (visualização e operação com bandas) e o processo de extração e análise de linhas de costa.

#### 4.1.2.5 Definição de linha de base

A linha de base permite que o CASSIE identifique qual região da área de interesse terá as linhas de costa analisadas, para formulação dos transectos e corte na AOI.

É utilizada novamente a biblioteca Drawing da Google Maps, porém com limitação à apenas um tipo de forma geométrica: uma *polyline* (*linestring*). Após a inserção do usuário, o CASSIE armazena a linha de base em seu estado, para que seja utilizada na etapa de processamento.

#### 4.1.2.6 Processamento

O processamento corresponde à execução de uma série de etapas sequenciais, com o objetivo de extrair as linhas de costa para um determinado conjunto de imagens, melhorar a qualidade das linhas extraídas e calcular estatísticas.

##### 4.1.2.6.1 Cálculo de NDWI

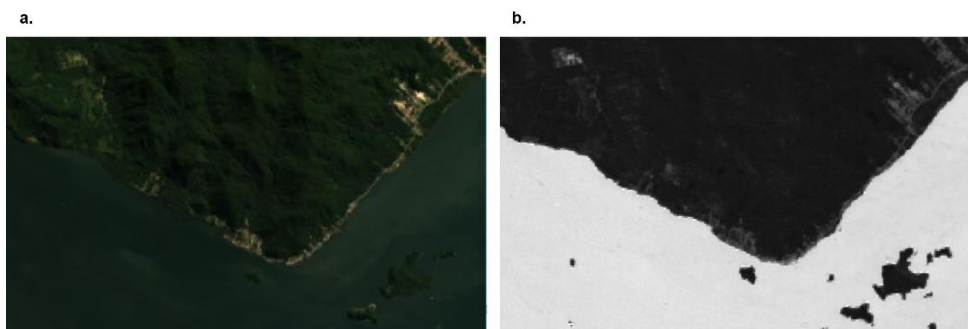
O NDWI (*Normalized Difference Water Index*), em imagens multiespectrais, consiste em realçar feições de água e minimizar o restante dos grupos. As bandas utilizadas dos satélites são aquelas que correspondem à *green* e *NIR*. Devido à abstração apresentada na seção de *Importação de Datasets*, o algoritmo não precisa diferir qual missão específica está sendo tratada em um dado momento.

A expressão realizada no cálculo do NDWI é observada na Equação 1.

$$NDWI = \frac{(Green - NIR)}{(Green + NIR)}$$

*Equação 1. Expressão de NDWI*

O CASSIE solicita para o GEE a aplicação da expressão na Equação 1 através do método *expression* nas imagens, informando as duas bandas equivalentes. Na Figura 3.a é possível observar a composição RGB de uma imagem multiespectral de uma região da Baía da Babitonga; e na Figura 3.b observa-se o produto do índice NDWI sobre a imagem.



*Figura 3 Comparação entre uma composição RGB e o resultado do índice NDWI sobre sua respectiva imagem multiespectral*

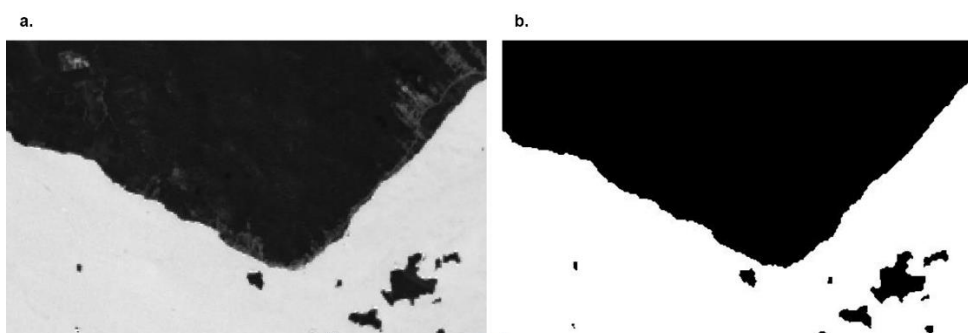
As imagens resultantes desta etapa possuem valores de píxel entre -1 e 1. As feições de água são representadas por pixels com valores próximos de 1, enquanto demais objetos possuem um valor próximo a -1.

#### *4.1.2.6.2 Limiarização*

A limiarização é um processo de segmentação de imagens que se baseia na diferença dos níveis de cinza que compõe objetos de uma imagem. A partir das características do objeto que se deseja isolar, estabelece-se um limiar para segmentar a imagem.

As imagens podem ser limiarizadas no CASSIE de acordo com um limiar específico, estabelecido pelo utilizador; ou por um limiar calculado pelo método de seleção de limiar para escalas de cinza desenvolvido por Otsu (1979).

O objetivo desta etapa é separar as feições de água da imagem resultante do cálculo de NDWI, que geralmente possuem valores de pixel maiores que 0. Na Figura 4.a observa-se a imagem resultante do cálculo de NDWI e na Figura 4.b o produto da limiarização.



*Figura 4 Limiarização de uma imagem em escala de cinza*

#### 4.1.2.6.3 Extração de linhas de costa

A extração de linhas de costa é dada pelo processo de transformação da imagem matricial limiarizada em uma imagem vetorial e, em seguida, realizar a extração dos segmentos exteriores (contorno) do polígono.

O polígono resultante da transformação para uma imagem vetorial é equivalente ao corpo de água observado na imagem, conforme demonstrado na Figura 5.

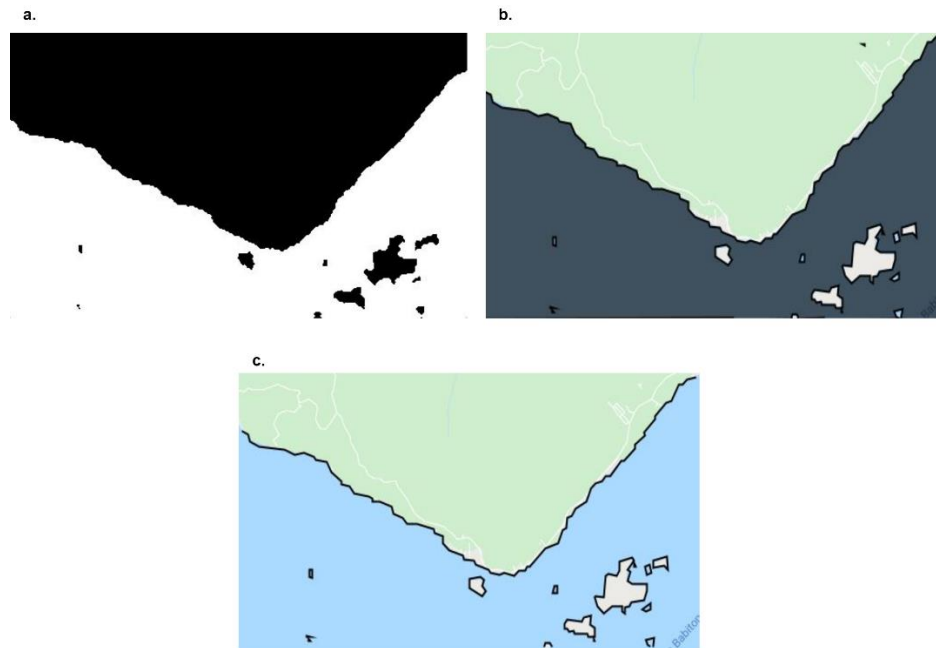


Figura 5 Extração de linhas de costa

O passo final para a extração das linhas de costa é a transformação do polígono obtido em linhas (*linestrings*), realizada por uma operação de intersecção com um buffer de redução da área de interesse. Na Figura 5.c é possível observar um exemplo do resultado desta operação.

#### 4.1.2.6.4 Criação dos transectos

A criação de transectos consiste na inserção de *linestrings* sobre a linha de base. Os transectos proporcionam uma amostragem sobre pontos determinados sobre as linhas de costa.

O CASSIE automaticamente gera transectos perpendiculares para cada segmento da linha de base, de acordo com uma extensão e espaçamento configurados pelo usuário, conforme observado na Figura 6.



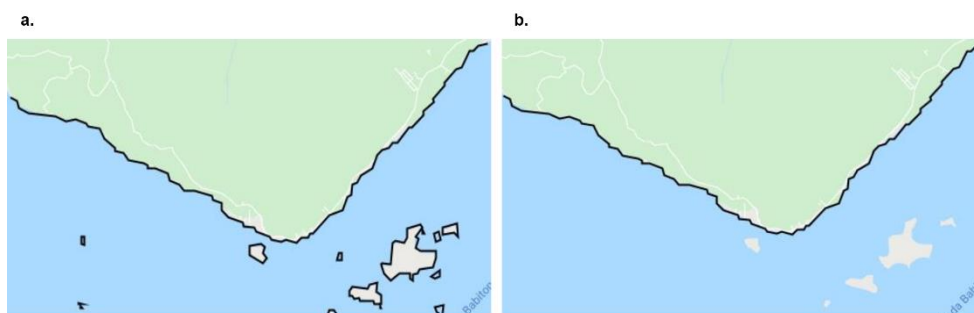


*Figura 6 Criação de transectos*

#### *4.1.2.6.5 Redução de ruídos*

As linhas de costas extraídas geralmente apresentam diversos tipos de ruídos (Figura 7.a), como polígonos muito pequenos e objetos inseridos no meio de feições de água que não são interessantes para o utilizador.

Desta forma, o CASSIE utiliza dos transectos criados para determinar qual dos objetos identificados como linha de costa deve ser analisado. O processo mantém apenas a linestring que tiver o maior número de intersecções com os transectos (Figura 7.b).



*Figura 7 Remoção de ruídos através de intersecção com transectos*

#### *4.1.2.6.6 Filtro Gaussiano*

A resolução da imagem de satélite afeta a qualidade da linha de costa extraída, contribuindo com diversos tipos de comportamento, como segmentos muito abruptos, que não condizem com o estado real da linha de costa, sendo necessário realizar uma suavização para obter uma linha de costa gradual.

O procedimento implementado no CASSIE é o filtro Gaussiano 1D (com tamanho de kernel 3 e desvio padrão 1), que realiza uma passagem sobre as coordenadas da linha de costa. O produto é uma linha de costa gradual e sem ruídos (Figura 8.b).

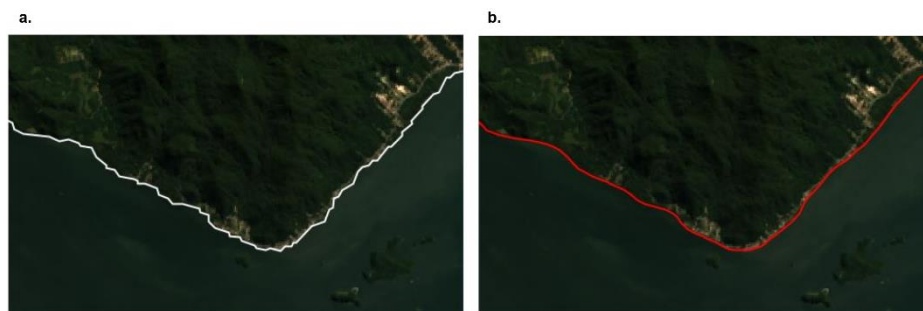


Figura 8 Suavização da linha de costa

#### 4.1.2.6.7 Cálculo de estatísticas

Após a extração e tratamento das linhas de costa, o CASSIE realiza o cálculo de diversas estatísticas, que comumente são utilizadas na análise de linha de costa, para cada um dos transectos.

O conjunto de estatísticas calculadas é composto por: *Shoreline Change Envelope (SCE)*, determinado pela distância entre a linha de costa mais distante e a mais próxima da linha de base; *Net Shoreline Movement (NSM)*, calculado pela distância entre a linha de costa mais recente e a linha de costa mais antiga; *End Point Rate (EPR)*, taxa de variação calculada pela razão do NSM e a diferença de tempo entre as linhas de costa (mais recente e mais antiga); *Linear Regression Rate (LRR)*, taxa de variação calculada por regressão linear e; Coeficiente de Correlação.

O cálculo foi implementado por métodos próprios do CASSIE utilizando as respectivas equações, com exceção do LRR e do Coeficiente de Correlação que utilizaram métodos do GEE (*linearRegression* e *pearsonsCorrelation*).

#### 4.1.2.7 Visualização de imagens

O CASSIE permite que o usuário opte por visualizar as imagens adquiridas durante a utilização, adicionado as que desejar ao mapa. A inserção necessita, entretanto, da data da imagem desejada.

Assim que uma requisição de visualização é recebida, o CASSIE solicita as bandas necessárias (RGB ou outra específica) para o GEE e instancia uma camada (Layer) para a imagem no Google Map, adicionando-a ao mapa e permitindo que o usuário possa realizar operações sobre as bandas desta imagem. A sobreposição de imagens no mapa prioriza as imagens adicionadas mais recentemente.

#### 4.1.2.8 Operações com bandas de imagem

O CASSIE implementa uma forma de aplicar quaisquer tipo de expressões sobre as bandas das imagens adquiridas, desde que estejam formatadas corretamente.

A operação é realizada através do método *expression* da classe Image do GEE, assim como é feito no cálculo do NDWI, na extração de linhas de costa. A diferença, entretanto, é a expressão que não é fixa. O método recebe uma cadeia de texto, inserida pelo usuário, referente a operação a ser realizada, além de um dicionário contendo a identificação de bandas utilizadas na expressão, que é mantido pelo próprio CASSIE através da encapsulação das bandas. É possível realizar operações com as bandas RED, GREEN, BLUE, NIR ou SWIR dos satélites.

#### 4.1.2.9 Exportação

Todos os dados sobre a linha de base, transectos e linhas de costa do CASSIE são exportáveis. O CASSIE disponibiliza três formatos de exportação: CSV, GeoJSON e ESRI Shapefile.

Os formatos CSV e GeoJSON não utilizam outras bibliotecas para a adição da funcionalidade. Para CSV, o CASSIE simplesmente converte os atributos dos objetos em colunas separadas por vírgulas e; para GeoJSON, que é o formato recebido na comunicação com o GEE, não é necessária operação nenhuma além da ativação do download.

Exportar para Shapefile, entretanto, necessitou da importação de uma biblioteca adicional, a *shp-write* da Mapbox. Esta biblioteca é responsável pela escrita em baixo nível da Shapefile (das formas geométricas e atributos), mas é limitada quanto ao formato de saída (uma pasta compactada por shapefile). Desta forma, a exportação não utiliza diretamente da biblioteca e sim de uma classe adaptadora sobre a biblioteca, que permite a exportação de diversas shapefiles em apenas uma pasta compactada.