



# GEOTECNOLOGIAS E SIG

Fundamentos, Análise Espacial e  
Aplicações Ambientais

Luiz Diego Vidal Santos

Universidade Estadual de Feira de Santana

2026 • CC BY-NC-SA 4.0

# **Geotecnologias e Sistemas de Informação Geográfica**

**Fundamentos, Análise Espacial e Aplicações Ambientais**

Luiz Diego Vidal Santos

15/02/2026

# Índice

<b>Sobre este livro</b>	<b>1</b>
Para quem é . . . . .	1
Como citar . . . . .	2
Licença . . . . .	2
<b>Prefácio</b>	<b>3</b>
<b>I. Parte I — Fundamentos de Geoinformação e Análise Espacial</b>	<b>4</b>
<b>1. Representação do Espaço Geográfico e Modelos de Dados</b>	<b>5</b>
1.1. Ontologias espaciais . . . . .	5
1.2. Modelo vetorial . . . . .	6
1.3. Modelo matricial . . . . .	6
1.4. Referenciamento espacial . . . . .	7
1.5. Geometria computacional . . . . .	8
1.6. Estruturas de indexação espacial . . . . .	8
1.7. Integração vetor-raster . . . . .	9
<b>2. Álgebra de Mapas, Inferência e Geoestatística Espacial</b>	<b>10</b>
2.1. Álgebra de mapas . . . . .	10
2.2. Overlay vetorial . . . . .	10
2.3. Modelos digitais de terreno . . . . .	11
2.4. Análise de redes . . . . .	11
2.5. Inferência e estatística espacial . . . . .	12
2.5.1. LISA (indicadores locais de associação espacial) . . . . .	13
2.6. Geoestatística e krigagem . . . . .	13
2.6.1. Krigagem ordinária . . . . .	14
2.7. Modelagem difusa (Fuzzy Sets) . . . . .	14
<b>3. Sensoriamento Remoto e Mudanças de Uso e Cobertura</b>	<b>16</b>
3.1. Interação radiação-matéria . . . . .	16
3.2. Índices de vegetação . . . . .	16
3.3. Classificadores de imagens . . . . .	17
3.4. Detecção de mudanças . . . . .	18
3.5. Séries temporais de imagens . . . . .	18
3.6. PRODES, DETER e MapBiomas . . . . .	19
3.7. Plataformas de processamento em nuvem . . . . .	20
<b>4. Classificação Digital e Avaliação de Incerteza</b>	<b>21</b>
4.1. Fontes de erro na classificação . . . . .	21
4.2. Amostragem de referência . . . . .	21
4.3. Matriz de confusão . . . . .	22
4.4. Índice Kappa e suas críticas . . . . .	23
4.5. Protocolo de validação área-ponderada . . . . .	23

4.6. Mapas de probabilidade . . . . .	24
4.7. Google Earth Engine na prática . . . . .	25
<b>II. Parte II — Solos, Erosão e Degradação Ambiental</b>	<b>26</b>
<b>5. Intemperismo, Erosão e Formação de Solos</b>	<b>27</b>
5.1. Fatores de formação do solo . . . . .	27
5.2. Processos de intemperismo . . . . .	27
5.2.1. Intemperismo químico . . . . .	27
5.2.2. Balanço de massa (Brimhall-Chadwick) . . . . .	28
5.3. Pedogênese e horizontação . . . . .	28
5.4. Dinâmica erosiva e RUSLE . . . . .	29
5.5. Integração solo-paisagem . . . . .	29
5.6. Conservação e engenharia de mitigação . . . . .	30
<b>6. Degradação do Solo, Desertificação e Mudanças Climáticas</b>	<b>31</b>
6.1. Degradação do solo . . . . .	31
6.2. Desertificação . . . . .	31
6.3. Neutralidade da degradação das terras . . . . .	32
6.4. Retroalimentação erosão-clima . . . . .	33
6.5. Salinização e colapso da estrutura do solo . . . . .	33
6.6. Mecanismo de Charney . . . . .	34
6.7. Indicadores bioquímicos de degradação . . . . .	34
6.8. Engenharia de restauração . . . . .	35
<b>III. Parte III — Hidrologia e Recursos Hídricos</b>	<b>36</b>
<b>7. Recursos Hídricos e Gestão por Bacia Hidrográfica</b>	<b>37</b>
7.1. Caracterização morfométrica da bacia . . . . .	37
7.2. Balanço hídrico . . . . .	38
7.3. Lei 9.433/1997 (Política Nacional de Recursos Hídricos) . . . . .	38
7.4. Governança e participação . . . . .	39
7.5. Desafios contemporâneos . . . . .	39
<b>8. Monitoramento Hidrológico e Geotecnologias</b>	<b>41</b>
8.1. Monitoramento convencional . . . . .	41
8.2. Telemetria e IoT hidrológica . . . . .	41
8.3. Curva-chave e hidrometria . . . . .	42
8.4. Altimetria espacial por radar . . . . .	42
8.5. Precipitação e umidade do solo por satélite . . . . .	42
8.6. Evapotranspiração e bacias virtuais . . . . .	43
<b>9. Geotecnologias, Recursos Hídricos e Gestão da Seca</b>	<b>45</b>
9.1. Crise hídrica e paradigma tecnológico . . . . .	45
9.2. Monitoramento de reservatórios e aquíferos . . . . .	45
9.3. Índices de seca . . . . .	46
9.3.1. SPI (Standardized Precipitation Index) . . . . .	46
9.3.2. SPEI (Standardized Precipitation-Evapotranspiration Index) . . . . .	46
9.3.3. Flash droughts . . . . .	46
9.4. Índices biofísicos de seca agrícola . . . . .	46
9.5. Evapotranspiração e seca hidrológica . . . . .	47

9.6. Governança estratégica e resiliência . . . . .	47
<b>IV. Parte IV — Aplicações Setoriais e Modelagem Avançada</b>	<b>49</b>
<b>10. Geoprocessamento e Monitoramento Ambiental da Mineração</b>	<b>50</b>
10.1. Monitoramento contínuo como paradigma . . . . .	50
10.2. InSAR e monitoramento de deformação . . . . .	50
10.3. Balanço volumétrico por LiDAR e fotogrametria . . . . .	51
10.4. Recuperação ambiental e histerese vegetacional . . . . .	51
10.5. Transporte de sedimentos e qualidade da água . . . . .	52
10.6. Gestão de risco e governança ambiental . . . . .	53
<b>11. SIG Aplicado à Exploração Mineral e Modelagem Preditiva</b>	<b>54</b>
11.1. Sistemas minerais e modelagem prospectiva . . . . .	54
11.2. Dados aerogeofísicos . . . . .	54
11.2.1. Aeromagnetometria . . . . .	54
11.2.2. Gamaespectrometria . . . . .	54
11.3. Sensoriamento remoto hiperespectral . . . . .	55
11.4. Modelagem multicritério e AHP . . . . .	56
11.5. Weights of Evidence (WofE) . . . . .	56
11.6. Modelagem por aprendizado de máquina . . . . .	57
11.7. Restrições ambientais e planejamento integrado . . . . .	58
<b>12. Inteligência Artificial e Qualidade da Pesquisa Ambiental</b>	<b>59</b>
12.1. O quarto paradigma da ciência . . . . .	59
12.2. Modelagem híbrida . . . . .	59
12.3. Validação espacial . . . . .	60
12.4. Explicabilidade e confiança . . . . .	60
12.5. Princípios FAIR e reprodutibilidade . . . . .	61
12.6. Green AI e eficiência computacional . . . . .	62
12.7. Human-in-the-loop e IA responsável . . . . .	62
<b>Referências</b>	<b>64</b>

## Sobre este livro

Este é o site do livro “Geotecnologias e Sistemas de Informação Geográfica: Fundamentos, Análise Espacial e Aplicações Ambientais”, escrito por [Luiz Diego Vidal Santos](#).

As geotecnologias constituem o conjunto de técnicas e métodos computacionais voltados à aquisição, armazenamento, processamento, análise e visualização de dados geoespaciais. No centro desse arsenal encontra-se o Sistema de Informação Geográfica (SIG), que integra modelos de dados vetoriais e matriciais com operações de álgebra de mapas, análise de redes, geoestatística e inferência espacial para transformar dados brutos em informação territorial acionável. A convergência recente entre sensoriamento remoto orbital, computação em nuvem, inteligência artificial e Internet das Coisas (IoT) ampliou a escala e a velocidade com que o monitoramento ambiental pode ser conduzido, tornando indispensável a formação de profissionais capazes de operar nesse ecossistema tecnológico.

Este livro apresenta de forma integrada os fundamentos teóricos, os métodos de análise espacial e as aplicações ambientais das geotecnologias, com ênfase em problemas relevantes para regiões tropicais semiáridas. Cada capítulo combina base conceitual, equações fundamentais, exemplos com dados reais e referências à legislação e aos sistemas de monitoramento operacionais no Brasil.

O conteúdo está organizado em quatro partes:

1. Fundamentos de Geoinformação e Análise Espacial — Representação do espaço geográfico, modelos de dados, álgebra de mapas, geoestatística, sensoriamento remoto e validação de classificações.
2. Solos, Erosão e Degradação Ambiental — Intemperismo, pedogênese, modelagem hidrosedimentológica (RUSLE), degradação, desertificação e retroalimentação climática.
3. Hidrologia e Recursos Hídricos — Gestão por bacia hidrográfica, monitoramento hidrológico com geotecnologias, índices de seca e governança hídrica.
4. Aplicações Setoriais e Modelagem Avançada — Monitoramento ambiental da mineração, exploração mineral com SIG e inteligência artificial aplicada à pesquisa ambiental.

## Para quem é

- Estudantes de graduação e pós-graduação em Engenharia Agrônômica, Ambiental, de Minas, Florestal e Geografia
- Profissionais de geoprocessamento, sensoriamento remoto e monitoramento ambiental
- Pesquisadores em recursos hídricos, solos e planejamento territorial
- Gestores de bacias hidrográficas, mineração e licenciamento ambiental

*Sobre este livro*

## **Como citar**

Santos, L. D. V. (2026). *Geotecnologias e Sistemas de Informação Geográfica: Fundamentos, Análise Espacial e Aplicações Ambientais*. Disponível em: <https://diegovidalcv.com.br/books/geotecnologias-sig/>

## **Licença**

Este livro é disponibilizado sob licença [Creative Commons Atribuição-NãoComercial-CompartilhaIgual 4.0 Internacional \(CC BY-NC-SA 4.0\)](#).

# Prefácio

O espaço geográfico é, simultaneamente, suporte físico e registro da ação humana sobre a superfície terrestre. Compreender sua organização, monitorar suas transformações e antecipar seus cenários de mudança exige ferramentas capazes de capturar a complexidade espacial e temporal dos processos ambientais. As geotecnologias, entendidas como o conjunto de técnicas de aquisição, processamento e análise de dados georreferenciados, consolidaram-se nas últimas décadas como infraestrutura indispensável para a gestão ambiental, o ordenamento territorial e a tomada de decisão baseada em evidências espaciais.

Este livro nasceu da experiência acumulada em mais de uma década de ensino, pesquisa e extensão em geoprocessamento aplicado ao semiárido brasileiro e à zona costeira do Nordeste. A convivência com problemas concretos de monitoramento de recursos hídricos em bacias intermitentes, mapeamento de degradação de terras em paisagens de Caatinga, avaliação de riscos geotécnicos em áreas de mineração e gestão da seca no Polígono das Secas me ensinou que a tecnologia geoespacial só se torna útil quando articulada a um arcabouço conceitual robusto, à validação rigorosa e ao conhecimento íntimo do território.

Ao longo dos doze capítulos, o leitor encontrará os fundamentos matemáticos da análise espacial (da álgebra de mapas à geoestatística), as bases físicas do sensoriamento remoto (da interação radiação-matéria aos índices de vegetação e seca), os modelos hidrossedimentológicos e hidrológicos que conectam processos de superfície ao ordenamento do território (RUSLE, balanço hídrico, SEBAL), as técnicas de análise multicritério e modelagem preditiva (AHP, lógica fuzzy, pesos de evidência, machine learning) e os marcos regulatórios que fundamentam a governança ambiental no Brasil (Lei 9.433/1997, CONAMA 357/2005, Código Florestal).

Agradeço aos colegas do grupo de pesquisa PLANeT-Inova (UEFS) e aos estudantes das disciplinas de Geotecnologias e SIG que, com suas perguntas e projetos, me obrigaram a refinar continuamente o conteúdo aqui apresentado.

**Luiz Diego Vidal Santos**

Feira de Santana, Bahia

Fevereiro de 2026

**Parte I.**

**Parte I — Fundamentos de  
Geoinformação e Análise Espacial**

# 1. Representação do Espaço Geográfico e Modelos de Dados

## 1.1. Ontologias espaciais

A representação computacional do espaço geográfico exige, antes de qualquer implementação técnica, uma decisão ontológica sobre a natureza dos fenômenos que se pretende modelar. Duas ontologias fundamentais competem na ciência da geoinformação, e a escolha entre elas determina não apenas a estrutura de dados, mas o repertório analítico disponível e as conclusões alcançáveis.

A ontologia de objetos discretos concebe o espaço geográfico como um conjunto de entidades distinguíveis (rios, estradas, edificações, lotes) com fronteiras definidas e identidade individual. Cada objeto possui geometria (posição, forma e dimensão), topologia (relações de adjacência, conectividade e continência com outros objetos) e atributos alfanuméricos (nome, tipo, área, proprietário). Essa concepção é formalizada pelo modelo vetorial, que representa o espaço por meio de primitivas geométricas (pontos, linhas e polígonos) associadas a tabelas de atributos.

A ontologia de campos contínuos, em contraste, concebe o espaço como uma superfície sobre a qual uma ou mais variáveis assumem valores em cada localização, sem fronteiras discretas entre regiões. Temperatura, elevação, precipitação, concentração de poluentes e umidade do solo são exemplos de fenômenos que variam continuamente no espaço e cuja representação natural é o modelo matricial (raster), no qual cada célula (pixel) de uma grade regular armazena o valor da variável naquela posição.

A Tabela 1.1 sintetiza as implicações dessa escolha ontológica para a modelagem espacial.

Tabela 1.1.: Comparativo entre as ontologias espaciais e suas implicações para modelagem em SIG.

Aspecto	Objetos discretos (vetor)	Campos contínuos (raster)
Natureza do fenômeno	Entidades com fronteiras definidas	Variável contínua no espaço
Primitivas Armazenamento	Pontos, linhas, polígonos Coordenadas + atributos	Células (pixels) de grade regular Matrizes numéricas
Resolução	Definida pela geometria	Definida pelo tamanho do pixel
Operações típicas	SQL espacial, overlay, redes	Álgebra de mapas, interpolação
Software exemplar	PostGIS, QGIS vetorial	GRASS GIS, Google Earth Engine

A escolha entre essas ontologias não é meramente técnica, mas reflete a compreensão do fenômeno em estudo. Um mapeamento de uso e cobertura do solo pode ser implementado tanto em formato vetorial (polígonos de classes) quanto raster (classificação pixel a pixel), e cada abordagem

carrega vantagens e limitações distintas. Na prática, projetos de geoprocessamento maduros integram ambos os modelos, convertendo entre formatos conforme a necessidade analítica.

### 1.2. Modelo vetorial

O modelo vetorial representa entidades geográficas por meio de três primitivas geométricas fundamentais, cuja hierarquia dimensional determina o tipo de fenômeno que cada uma pode representar.

O ponto (dimensão 0) é definido por um par de coordenadas  $(x, y)$  ou, em representações tridimensionais, por uma tripla  $(x, y, z)$ . Pontos representam fenômenos cuja extensão espacial é desprezível na escala de análise, como estações pluviométricas, poços de monitoramento, ocorrências minerais ou localidades. A linha (dimensão 1) é definida por uma sequência ordenada de pontos conectados, representando feições com extensão longitudinal dominante, como rios, estradas, linhas de transmissão e divisores de água. O polígono (dimensão 2) é definido por uma sequência fechada de linhas que delimita uma área, representando feições com extensão bidimensional significativa, como parcelas agrícolas, reservatórios, municípios e manchas de vegetação.

Além da geometria, o modelo vetorial exige uma estrutura topológica que codifique as relações espaciais entre objetos. A topologia é o ramo da matemática que estuda as propriedades que se preservam sob transformações contínuas (rotação, translação, escala), e no contexto de SIG permite responder a perguntas como “quais municípios são adjacentes?”, “esta estrada cruza este rio?” ou “este ponto está dentro deste polígono?”. O modelo arc-node, implementado em sistemas como ArcGIS e PostGIS, codifica a topologia por meio de arcos (segmentos de linha) conectados em nós (pontos de interseção), de modo que cada arco conhece os polígonos à sua esquerda e à sua direita, e cada nó conhece os arcos que nele convergem.

#### **i** Topologia e integridade

A ausência de topologia (modo “espaguete”) permite sobreposições, lacunas e inconsistências entre polígonos que passam despercebidas visualmente mas corrompem análises como cálculo de área, overlay e roteamento. A construção de topologia é, portanto, um passo obrigatório na preparação de dados vetoriais para análise espacial.

O padrão aberto da OGC (Open Geospatial Consortium) especifica três serviços fundamentais para interoperabilidade de dados vetoriais e matriciais entre sistemas: o WMS (Web Map Service), que fornece imagens renderizadas de mapas; o WFS (Web Feature Service), que fornece feições vetoriais com geometria e atributos; e o WCS (Web Coverage Service), que fornece coberturas raster. Esses padrões permitem que dados produzidos em diferentes instituições e softwares sejam consumidos em qualquer SIG compatível, viabilizando a Infraestrutura Nacional de Dados Espaciais (INDE) e plataformas como o i3Geo e o GeoServer.

### 1.3. Modelo matricial

O modelo matricial (raster) discretiza o espaço em uma grade regular de células retangulares, onde cada célula armazena um valor numérico que representa a grandeza medida naquela posição. A resolução espacial (tamanho do pixel) define o nível de detalhe capturável: imagens Sentinel-2 com resolução de 10 m distinguem talhões agrícolas individuais, enquanto dados MODIS com resolução de 250 m capturam padrões regionais de vegetação.

O trade-off fundamental do modelo raster é a relação inversa entre resolução e volume de dados. Uma imagem que cobre  $100 \times 100$  km com pixels de 10 m contém  $10^8$  (cem milhões) de células por banda espectral. Para um satélite como o Sentinel-2 com 13 bandas, isso equivale a 1,3 bilhão de valores por cena. Essa escala de dados exige infraestrutura computacional robusta e motiva o uso de plataformas de processamento em nuvem como o Google Earth Engine (Gorelick et al., 2017).

Cada pixel de uma imagem raster pode armazenar diferentes tipos de dados: inteiros sem sinal (0–255 para imagens de 8 bits), inteiros com sinal, ponto flutuante (para dados contínuos como temperatura ou elevação) ou valores categóricos (para mapas de uso do solo). A profundidade radiométrica (número de bits por pixel) determina a quantidade de níveis de cinza distinguíveis: uma imagem de 8 bits resolve 256 níveis, enquanto uma de 12 bits (como as do Sentinel-2) resolve 4.096 níveis, permitindo diferenciar variações sutis de reflectância que uma imagem de 8 bits não captaria.

## 1.4. Referenciamento espacial

A localização de qualquer fenômeno na superfície terrestre exige um Sistema de Referência de Coordenadas (SRC) que defina a forma matemática da Terra (elipsoide de referência), a posição e orientação desse elipsoide em relação ao centro de massa da Terra (datum geodésico) e a projeção cartográfica que transforma coordenadas curvilíneas (latitude/longitude) em coordenadas planas ( $E, N$ ) sobre um mapa.

O datum geodésico é o parâmetro mais crítico, pois datums diferentes posicionam o mesmo ponto em coordenadas distintas. No Brasil, a transição do SAD-69 (anterior) para o SIRGAS 2000 (atual, geocêntrico e compatível com GPS) introduziu deslocamentos de 60 a 70 m em algumas regiões do país. Projetar dados em SAD-69 sobre uma base em SIRGAS 2000 sem transformação de datum gera erros de posicionamento que podem comprometer análises de sobreposição, cálculo de distâncias e delimitação de áreas de proteção.

As projeções cartográficas introduzem distorções inevitáveis ao transformar uma superfície curva em um plano. A indicatriz de Tissot fornece uma visualização quantitativa dessas distorções: em projeções conformes (como a UTM), a indicatriz é sempre circular (preserva ângulos mas distorce áreas); em projeções equivalentes (como a de Albers), a indicatriz preserva a área do círculo mas distorce sua forma; e em projeções equidistantes, a indicatriz preserva distâncias ao longo de direções privilegiadas.

A Tabela 1.2 resume as propriedades das projeções mais utilizadas em geoprocessamento ambiental no Brasil.

Tabela 1.2.: Projeções cartográficas e suas propriedades para geoprocessamento no Brasil.

Projeção	Propriedade preservada	Uso típico	Distorção principal
UTM (Universal Transversa de Mercator)	Ângulos (conforme)	Topografia, cadastro, alta precisão local	Área (até 0,04% no eixo central do fuso)
Cônica de Albers	Área (equivalente)	Mapeamento temático, uso do solo	Forma em altas latitudes

Projeção	Propriedade preservada	Uso típico	Distorção principal
Policônica	Compromisso	Cartas do IBGE, séries históricas	Distorções crescentes nas bordas
Lat/Long (geográfica)	Nenhuma especificamente	Intercâmbio de dados, GPS	Distorce tudo; não é projeção

**! Rastreabilidade metrológica**

Toda análise espacial deve documentar explicitamente o SRC utilizado (código EPSG), o datum geodésico e a fonte dos dados altimétricos. A metadocumentação é requisito da INDE (Infraestrutura Nacional de Dados Espaciais) e condição para reprodutibilidade científica. A omissão do SRC é uma das causas mais frequentes de erro em projetos de geoprocessamento.

## 1.5. Geometria computacional

A manipulação eficiente de dados vetoriais exige algoritmos de geometria computacional que resolvam problemas como interpolação de superfícies, particionamento do espaço e busca de vizinhos mais próximos.

A triangulação de Delaunay constrói uma malha de triângulos não sobrepostos a partir de um conjunto de pontos, com a propriedade de que nenhum ponto se encontra no interior do circuncírculo de qualquer triângulo. Essa propriedade maximiza o menor ângulo de todos os triângulos, evitando triângulos muito alongados (“finos”) que comprometem a qualidade da interpolação. A triangulação de Delaunay é a base dos Modelos Numéricos de Terreno (MNTs) do tipo TIN (Triangulated Irregular Network), que representam a topografia com resolução variável (mais triângulos em terreno acidentado, menos em terreno plano).

O diagrama de Voronoi (ou Thiessen) é o dual geométrico da triangulação de Delaunay: para cada ponto do conjunto, define uma região poligonal contendo todas as localizações mais próximas daquele ponto do que de qualquer outro. Os polígonos de Voronoi são amplamente utilizados em hidrologia para calcular a precipitação média em uma bacia a partir de estações pluviométricas distribuídas irregularmente (método de Thiessen), e em ecologia para delimitar áreas de influência de espécies ou recursos.

## 1.6. Estruturas de indexação espacial

A consulta eficiente em bases de dados geoespaciais com milhões de feições exige estruturas de indexação que evitem a varredura sequencial de todos os registros. A R-tree (e suas variantes R\*-tree e R+-tree) organiza os retângulos envolventes mínimos (MBR) das geometrias em uma árvore balanceada, permitindo que consultas do tipo “quais rios cruzam este município?” ou “quais poços estão a menos de 5 km desta rodovia?” sejam resolvidas em tempo logarítmico em vez de linear.

O PostGIS, extensão espacial do PostgreSQL, implementa indexação por GiST (Generalized Search Tree), que suporta consultas espaciais complexas com desempenho adequado para bases de dados com dezenas de milhões de feições. A eficiência da indexação espacial é o que torna

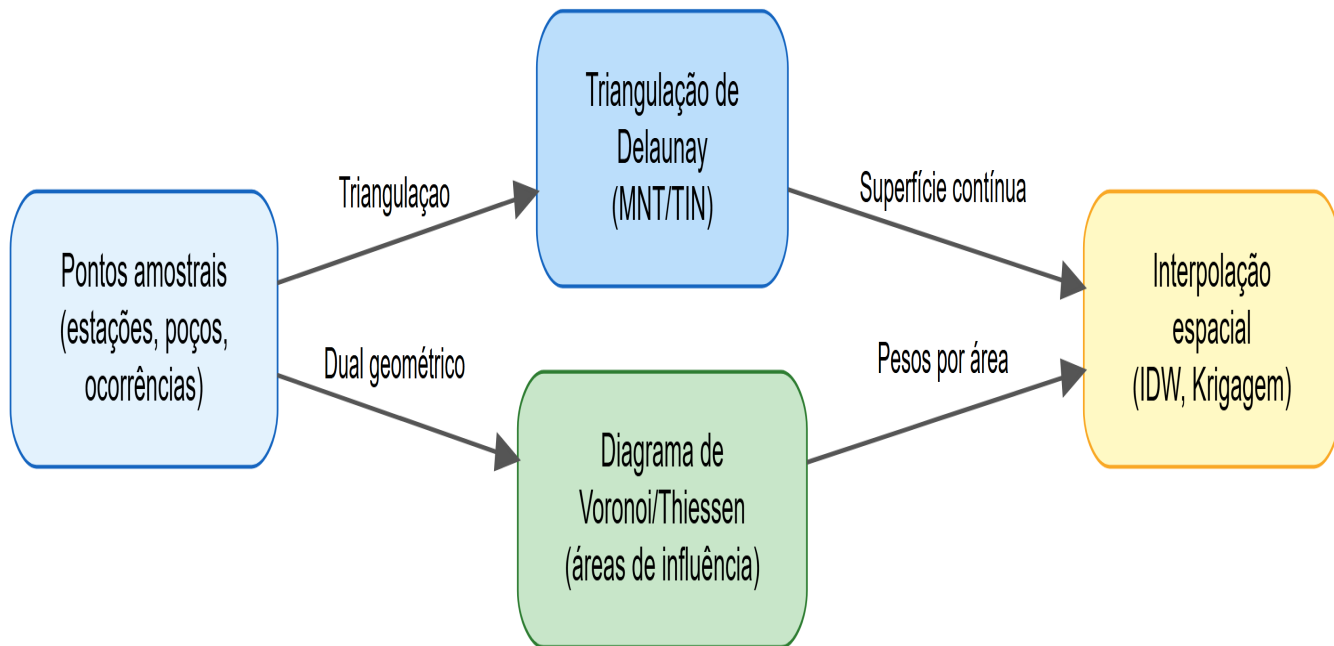


Figura 1.1.: Relação dual entre triangulação de Delaunay e diagrama de Voronoi para um conjunto de pontos amostrais.

viável operar sobre a INDE ou sobre bases de dados como o cadastro ambiental rural (CAR), que contém mais de 6 milhões de imóveis rurais georreferenciados.

#### 💡 Na prática

Ao carregar uma camada vetorial no QGIS ou PostGIS, a criação de índice espacial (GIST ou R-tree) deve ser o primeiro passo antes de qualquer análise. Sem índice, operações de overlay, junção espacial e seleção por localização podem ser ordens de magnitude mais lentas.

## 1.7. Integração vetor-raster

Na prática de geoprocessamento, a maioria dos projetos exige conversões entre os modelos vetorial e matricial. A rasterização transforma feições vetoriais em pixels, atribuindo a cada célula o valor do atributo do polígono que a cobre; é necessária quando se deseja realizar álgebra de mapas com camadas temáticas originalmente vetoriais (uso do solo, geologia, pedologia). A vetorização transforma regiões contíguas de pixels com mesmo valor em polígonos; é necessária quando se deseja extrair estatísticas zonais ou realizar análise topológica sobre resultados de classificação de imagens.

Ambas as conversões introduzem erros. A rasterização gera efeito “escada” em bordas diagonais e pode atribuir erroneamente a uma célula a classe do polígono que cobre a maior fração de sua área, mesmo que outros polígonos cubram porções significativas. A vetorização gera bordas anguladas que não correspondem às fronteiras reais dos fenômenos. A compreensão desses artefatos é fundamental para interpretar corretamente os resultados e estimar a incerteza associada a cada conversão.

## 2. Álgebra de Mapas, Inferência e Geostatística Espacial

### 2.1. Álgebra de mapas

A álgebra de mapas, formalizada por Tomlin (1990), constitui o arcabouço operacional para manipulação de dados raster em SIG. Trata-se de um conjunto de operações aritméticas, lógicas e estatísticas aplicadas sobre grades regulares de pixels, organizadas em quatro categorias funcionais segundo a vizinhança de cálculo envolvida.

As operações locais (ou pontuais) calculam o valor de cada pixel de saída exclusivamente a partir dos valores dos pixels correspondentes nas camadas de entrada. Exemplos incluem a soma, subtração, razão e combinação lógica de bandas espectrais, como o cálculo do NDVI ( $\frac{\rho_{NIR}-\rho_{RED}}{\rho_{NIR}+\rho_{RED}}$ ), que é uma operação local aplicada pixel a pixel sobre duas bandas de uma imagem de satélite. As operações focais (ou de vizinhança) calculam o valor de cada pixel com base em uma janela de vizinhança ao seu redor, como filtros de média, mediana, desvio-padrão e operadores morfológicos (erosão, dilatação). As operações zonais calculam estatísticas agregadas para regiões definidas por uma camada de zonas (por exemplo, a precipitação média por sub-bacia ou o NDVI médio por classe de uso do solo). As operações globais consideram toda a grade para computar o resultado, como o cálculo de fluxo acumulado a partir de um modelo digital de elevação.

A Tabela 2.1 resume as quatro categorias e suas aplicações em geoprocessamento ambiental.

Tabela 2.1.: Categorias da álgebra de mapas de Tomlin e suas aplicações ambientais.

Categoria	Vizinhança	Exemplo ambiental	Operação SIG típica
Local	Pixel individual	NDVI, razão de bandas, reclassificação	Raster Calculator
Focal	Janela $n \times n$	Filtro de média (suavização), declividade, curvatura	r.neighbors (GRASS), Focal Statistics (ArcGIS)
Zonal	Regiões (zonas)	NDVI médio por município, precipitação por bacia	Zonal Statistics
Global	Grade inteira	Fluxo acumulado, distância euclidiana, custo acumulado	r.watershed, Cost Distance

### 2.2. Overlay vetorial

Enquanto a álgebra de mapas opera sobre grades raster, o overlay vetorial realiza operações de conjunto (interseção, união, diferença simétrica, clip) sobre geometrias de polígonos. A interseção de uma camada de aptidão agrícola com uma camada de áreas protegidas, por exemplo, produz uma nova camada cujos polígonos herdam atributos de ambas as entradas, permitindo identificar

conflitos de uso. A união combina as geometrias mantendo todos os segmentos de fronteira, e a diferença remove de uma camada as regiões cobertas pela outra.

Essas operações são computacionalmente custosas para bases de dados grandes (dezenas de milhares de polígonos complexos) e exigem topologia consistente nas camadas de entrada. Erros topológicos (polígonos com auto-interseção, lacunas, sobreposições) geram falhas silenciosas no overlay que podem propagar erros para toda a análise subsequente.

## 2.3. Modelos digitais de terreno

Os Modelos Digitais de Terreno (MDTs) são representações matriciais da superfície terrestre que codificam a elevação em cada célula da grade. A partir do MDT, derivam-se variáveis topográficas fundamentais para hidrologia, geomorfologia e planejamento ambiental.

A declividade ( $S$ ) quantifica a taxa de variação da elevação no espaço:

$$S = \arctan \left( \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \right)$$

onde  $z$  é a elevação e  $x, y$  são as coordenadas planas. O aspecto (orientação da vertente) indica a direção de máxima declividade, influenciando a insolação, a evapotranspiração e a distribuição da vegetação.

O fluxo acumulado, calculado pelo algoritmo D8 (que direciona toda a água de cada célula para a célula vizinha de menor elevação) ou pelo algoritmo D-infinity (Tarboton, 1997) (que distribui o fluxo proporcionalmente entre as duas vizinhas de descida mais íngreme), é a base para a delimitação automática de bacias hidrográficas e redes de drenagem a partir de MDTs. O fluxo acumulado em uma célula estima, de forma simplificada, a área de contribuição a montante daquele ponto, e quando multiplicado pela declividade, resulta no índice de potência de escoamento, indicador de susceptibilidade à erosão linear.

## 2.4. Análise de redes

A análise de redes modela o espaço como um grafo composto por nós (interseções, localidades) e arestas (segmentos de estrada, trechos de rio) com pesos associados (distância, tempo de percurso, custo). Algoritmos clássicos de teoria dos grafos são aplicados diretamente sobre dados vetoriais de linhas para resolver problemas logísticos e ambientais.

O algoritmo de Dijkstra encontra o caminho de menor custo entre dois nós, sendo utilizado para roteamento veicular, definição de rotas de escoamento agrícola e planejamento de estradas rurais. O algoritmo A\* é uma variante heurística de Dijkstra que incorpora uma estimativa de distância ao destino, acelerando a busca em grafos grandes. O problema de fluxo máximo (max-flow), resolvido pelo algoritmo de Ford-Fulkerson, modela a capacidade de transporte de uma rede de drenagem ou de distribuição de água.

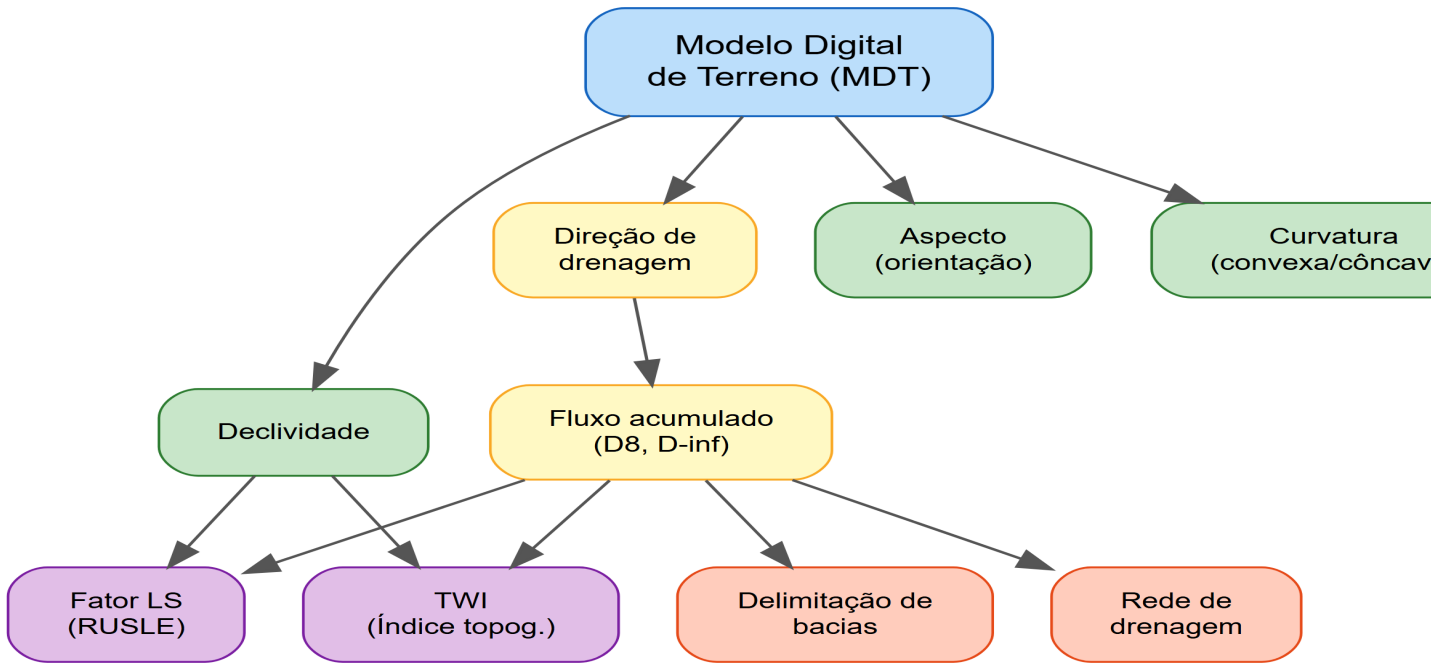


Figura 2.1.: Produtos deriváveis de um MDT e suas aplicações ambientais.

## 2.5. Inferência e estatística espacial

A estatística espacial difere da estatística clássica por incorporar explicitamente a localização das observações no modelo. A Primeira Lei da Geografia de Tobler (1970) postula que “tudo está relacionado com tudo o mais, mas coisas próximas estão mais relacionadas do que coisas distantes”. Essa dependência espacial (autocorrelação) viola a premissa de independência das observações que fundamenta a maioria dos testes estatísticos clássicos e exige métodos próprios de análise.

O Índice Global de Moran quantifica a autocorrelação espacial de uma variável contínua:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

onde  $n$  é o número de observações,  $x_i$  é o valor observado na localização  $i$ ,  $\bar{x}$  é a média global e  $w_{ij}$  são os pesos da matriz de vizinhança espacial. Valores de  $I$  próximos de +1 indicam autocorrelação positiva (clusters de valores similares), valores próximos de -1 indicam autocorrelação negativa (padrão xadrez) e valores próximos de zero indicam ausência de padrão espacial (distribuição aleatória).

O teste de significância do Índice de Moran compara o valor observado com a distribuição esperada sob a hipótese nula de aleatoriedade espacial (permutação aleatória dos valores entre as localizações). A rejeição da hipótese nula ( $p < 0,05$ ) indica que o padrão espacial observado é improvável sob distribuição aleatória, sugerindo a existência de processos espacialmente estruturados.

### 2.5.1. LISA (indicadores locais de associação espacial)

Enquanto o Moran Global fornece uma medida resumo para toda a distribuição, os indicadores LISA (Local Indicators of Spatial Association) propostos por Anselin (1995) decompõem a autocorrelação global em contribuições locais, identificando onde se localizam os clusters e as anomalias espaciais.

Cada observação  $i$  recebe uma classificação em uma de quatro categorias: High-High (valor alto cercado por valores altos, cluster quente), Low-Low (valor baixo cercado por valores baixos, cluster frio), High-Low (valor alto cercado por valores baixos, outlier positivo) e Low-High (valor baixo cercado por valores altos, outlier negativo). Essa classificação, quando mapeada, produz o mapa de clusters LISA, ferramenta visual poderosa para identificar hotspots de degradação ambiental, clusters de incidência de doenças, concentrações de pobreza ou zonas de alta produtividade agrícola.

#### ⚠ MAUP (Problema da Unidade de Área Modificável)

O MAUP adverte que os resultados de análises espaciais com dados agregados (por município, bacia, setor censitário) dependem da forma e do tamanho das unidades de agregação. Uma variável que apresenta correlação significativa com outra no nível municipal pode não apresentá-la no nível estadual, e vice-versa. Esse problema é inerente a qualquer análise com dados zonais e deve ser reconhecido explicitamente na interpretação dos resultados.

## 2.6. Geoestatística e krigagem

A geoestatística, desenvolvida por Georges Matheron a partir dos trabalhos de Daniel Krige na mineração sul-africana, é a disciplina que modela a variabilidade espacial de variáveis regionalizadas (variáveis que apresentam continuidade espacial mas também flutuações locais). A ferramenta central da geoestatística é o semivariograma, que quantifica a dependência espacial como função da distância entre observações.

O semivariograma experimental é calculado como:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i + h)]^2$$

onde  $N(h)$  é o número de pares de observações separados pela distância  $h$ ,  $Z(s_i)$  é o valor da variável na localização  $s_i$  e  $Z(s_i + h)$  é o valor na localização separada por  $h$ . O semivariograma cresce da origem até estabilizar-se em um patamar, e três parâmetros descrevem seu comportamento.

O efeito pepita ( $C_0$ ) é o valor do semivariograma na origem (mas não em  $h = 0$ ), representando a variabilidade em escalas menores que a resolução amostral somada ao erro de medição. O patamar ( $C_0 + C$ ) é o valor no qual o semivariograma se estabiliza, correspondendo à variância total da variável. O alcance ( $a$ ) é a distância na qual o semivariograma atinge o patamar, indicando o raio de dependência espacial; observações separadas por distâncias superiores ao alcance são estatisticamente independentes.

A Tabela 2.2 apresenta os três modelos teóricos mais utilizados para ajuste ao semivariograma experimental.

Tabela 2.2.: Modelos teóricos de semivariograma e suas características.

Modelo	Equação $\gamma(h)$	Comportamento na origem	Uso típico
Esférico	$C_0 + C \left[ \frac{3h}{2a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right]$ para $h \leq a$	Linear	Solos, geologia, variáveis ambientais
Exponencial	$C_0 + C \left[ 1 - \exp \left( -\frac{h}{a} \right) \right]$	Côncavo para cima	Propriedades contínuas, concentrações
Gaussiano	$C_0 + C \left[ 1 - \exp \left( -\frac{h^2}{a^2} \right) \right]$	Parabólico	Variáveis suaves, topografia

### 2.6.1. Krigagem ordinária

A krigagem ordinária é o interpolador geoestatístico que estima o valor de uma variável em localizações não amostradas como combinação linear ponderada dos valores observados:

$$\hat{Z}(s_0) = \sum_{i=1}^n \lambda_i \cdot Z(s_i)$$

onde  $\hat{Z}(s_0)$  é o valor estimado na localização não amostrada  $s_0$ ,  $Z(s_i)$  são os valores observados nas  $n$  localizações amostrais e  $\lambda_i$  são os pesos de krigagem, determinados pela resolução de um sistema de equações lineares que minimiza a variância da estimativa sob a restrição de não-viés (os pesos somam 1). Os pesos são obtidos a partir do modelo de semivariograma ajustado, de modo que observações mais próximas e em posições geometricamente favoráveis recebem pesos maiores.

A krigagem possui uma propriedade única entre os interpoladores: além da estimativa, fornece a variância de krigagem em cada ponto, que quantifica a incerteza da predição. A superfície de variância de krigagem revela onde a rede amostral é insuficiente (alta variância) e onde é redundante (baixa variância), orientando o planejamento otimizado de campanhas de campo.

#### Krigagem vs. IDW

O Inverso da Distância Ponderada (IDW) é um interpolador determinístico que pondera as observações pelo inverso da distância elevada a uma potência ( $p$ , tipicamente 2). Embora computacionalmente simples, o IDW não modela a estrutura de dependência espacial (não usa semivariograma), não fornece estimativa de incerteza e é sensível à configuração geométrica da amostragem. A krigagem deve ser preferida sempre que houver dados suficientes para ajustar um semivariograma robusto (mínimo de 30–50 observações).

## 2.7. Modelagem difusa (Fuzzy Sets)

A lógica difusa (fuzzy), proposta por Zadeh em 1965, permite representar pertinências parciais a classes, um avanço fundamental para análises ambientais nas quais as fronteiras entre categorias são gradacionais e não rígidas. Enquanto a lógica booleana classifica cada localização como “pertencente” (1) ou “não pertencente” (0) a uma classe, a lógica difusa atribui um grau de pertinência  $\mu \in [0, 1]$  que reflete o quanto cada localização satisfaz os critérios da classe.

O operador Fuzzy Gamma combina mapas de pertinência por meio de uma ponderação entre os operadores Fuzzy OR (união, que maximiza) e Fuzzy AND (interseção, que minimiza):

$$\mu_{\gamma} = (\mu_{OR})^{\gamma} \cdot (\mu_{AND})^{1-\gamma}$$

onde  $\gamma \in [0, 1]$  controla a tendência do modelo: valores próximos de 1 produzem resultados mais inclusivos (próximos ao OR) e valores próximos de 0 produzem resultados mais restritivos (próximos ao AND). Na prática,  $\gamma = 0,7$  a  $0,9$  é comumente utilizado em mapeamento de aptidão e prospectividade, pois representa um compromisso entre capturar todas as regiões potenciais (inclusividade) e evitar falsos positivos (seletividade).

## 3. Sensoriamento Remoto e Mudanças de Uso e Cobertura

### 3.1. Interação radiação-matéria

O sensoriamento remoto fundamenta-se na detecção e medição da radiação eletromagnética refletida ou emitida pela superfície terrestre. A assinatura espectral de cada alvo (solo, vegetação, água, rocha, superfície urbana) é determinada por suas propriedades físico-químicas, que condicionam a absorção, reflexão e transmissão da radiação em diferentes comprimentos de onda. A exploração dessas diferenças espectrais é o que permite discriminar alvos na superfície terrestre a partir de imagens orbitais.

Na região do visível (400–700 nm), a vegetação verde absorve intensamente a radiação azul e vermelha (pelos pigmentos fotossintéticos, principalmente clorofilas *a* e *b*, e carotenoides) e reflete parcialmente o verde, o que confere a coloração característica às folhas. Na transição para o infravermelho próximo (NIR, 700–1300 nm), a reflectância da vegetação sadia aumenta abruptamente (o chamado “red-edge”), atingindo valores de 40–50%, devido ao espalhamento múltiplo da radiação no mesófilo esponjoso da folha, uma estrutura de paredes celulares com interfaces ar-água que funcionam como espelhos microscópicos. No infravermelho de ondas curtas (SWIR, 1300–2500 nm), a reflectância diminui em comprimentos de onda centrados em 1450 e 1940 nm, correspondentes às bandas de absorção da água líquida; quanto maior o conteúdo de água da folha, mais profundas essas feições de absorção.

#### **i** Caatinga e deciduidade

Na Caatinga, a deciduidade sazonal (perda completa das folhas no período seco) transforma radicalmente a assinatura espectral da paisagem: a vegetação passa de valores altos de reflectância no NIR (período chuvoso) para valores dominados pelo solo exposto e serrapilheira (período seco), com drástica redução do contraste espectral entre vegetação e solo. Essa dinâmica exige cuidado na interpretação de séries temporais de índices de vegetação, pois reduções de NDVI na Caatinga não indicam necessariamente degradação, mas podem refletir o ciclo fenológico natural.

### 3.2. Índices de vegetação

Os índices de vegetação são combinações matemáticas de bandas espectrais projetadas para realçar propriedades biofísicas da vegetação enquanto minimizam efeitos de fundo (solo, atmosfera, geometria de iluminação). O NDVI (Normalized Difference Vegetation Index), proposto por Rouse et al. (1974), é o índice mais amplamente utilizado:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$$

onde  $\rho_{NIR}$  e  $\rho_{RED}$  são as reflectâncias nas bandas do infravermelho próximo e do vermelho, respectivamente. O NDVI varia de  $-1$  a  $+1$ , com valores negativos indicando água, valores próximos de zero indicando solo exposto ou vegetação rala e valores acima de  $0,3$  indicando vegetação ativa. A normalização pela soma das bandas confere ao NDVI relativa insensibilidade a variações de iluminação, mas o índice satura em coberturas densas ( $IAF > 3$ ) e é sensível ao substrato solo em coberturas esparsas.

O EVI (Enhanced Vegetation Index) foi desenvolvido para superar as limitações do NDVI, incorporando uma banda azul para correção atmosférica e um fator de ajuste de solo:

$$EVI = G \times \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + C_1 \cdot \rho_{RED} - C_2 \cdot \rho_{BLUE} + L}$$

onde  $G = 2,5$  é o fator de ganho,  $C_1 = 6$  e  $C_2 = 7,5$  são coeficientes de correção de aerossóis e  $L = 1$  é o fator de ajuste de dossel. O EVI apresenta menor saturação em vegetação densa e maior sensibilidade em regiões de alta biomassa, sendo preferido para monitoramento de florestas tropicais.

A Tabela 3.1 compara os principais índices de vegetação e suas aplicações.

Tabela 3.1.: Índices de vegetação e suas propriedades comparativas.

Índice	Fórmula simplificada	Vantagem	Limitação	Aplicação
NDVI	$(NIR - RED)/(NIR + RED)$	Simplicidade, série histórica longa	Satura em alta biomassa	Monitoramento regional, séries temporais
EVI	$2.5 \times (NIR - RED)/(NIR + 6RED - 7.5BLUE + 1)$	Menor saturação, correção atmosférica	Requer banda azul	Florestas tropicais, áreas úmidas
SAVI	$(NIR - RED)/(NIR + RED + L) \times (1 + L)$	Reduz efeito do solo	$L$ fixo (0,5) pode não ser ótimo	Vegetação esparsa, semiárido
NDFI	Fração de degradação	Detecta degradação florestal sutil	Complexo (requer Spectral Mixture)	Manejo florestal, REDD+

### 3.3. Classificadores de imagens

A classificação digital de imagens transforma uma imagem multiespectral (valores de reflectância em cada pixel e banda) em um mapa temático (classes de uso e cobertura do solo). Três famílias de classificadores dominam a prática atual do sensoriamento remoto.

O Random Forest (Breiman, 2001) é um ensemble de árvores de decisão que classifica cada pixel por voto majoritário entre centenas de árvores treinadas em subconjuntos aleatórios das amostras e variáveis. Sua robustez a ruído, capacidade de lidar com variáveis correlacionadas e de estimar a importância relativa de cada banda o tornaram o classificador mais popular em mapeamento de uso e cobertura. Uma configuração típica utiliza 100 a 500 árvores (a acurácia melhora pouco além de 500) e  $\sqrt{p}$  variáveis candidatas por divisão (onde  $p$  é o número total de bandas e índices).

### 3. Sensoriamento Remoto e Mudanças de Uso e Cobertura

O SVM (Support Vector Machine) busca o hiperplano de separação ótima entre classes no espaço multiespectral, maximizando a margem (distância entre o hiperplano e os exemplos mais próximos de cada classe). O kernel RBF (Radial Basis Function) permite ao SVM resolver problemas não linearmente separáveis ao projetar os dados em um espaço de dimensionalidade superior. O SVM é particularmente eficaz com amostras de treinamento limitadas e em problemas com muitas bandas espectrais (reduz o risco de “maldição da dimensionalidade”).

A U-Net é uma arquitetura de rede neural convolucional (CNN) com estrutura encoder-decoder e conexões residuais que produz segmentação semântica pixel a pixel. Projetada originalmente para imagens biomédicas, a U-Net foi adaptada com sucesso para classificação de imagens de satélite, onde sua capacidade de capturar contexto espacial (padrões de vizinhança) supera classificadores que tratam cada pixel isoladamente. Entretanto, a U-Net exige volumes significativos de dados de treinamento rotulados e recursos computacionais (GPU), o que limita sua aplicação em contextos com poucos dados de referência.

#### 3.4. Detecção de mudanças

A detecção de mudanças de uso e cobertura do solo ao longo do tempo pode seguir duas abordagens fundamentais. Na detecção pré-classificação (álgebra de mapas), realizam-se operações matemáticas diretamente sobre as imagens de duas datas (diferença, razão, análise de componentes principais de mudança), identificando pixels com variação espectral significativa sem atribuir classes temáticas. Na detecção pós-classificação, classificam-se independentemente as imagens de cada data e compara-se o par de mapas resultantes por tabulação cruzada (matriz de transição), identificando quais classes sofreram conversão.

As matrizes de transição de Markov são frequentemente empregadas para modelar e projetar as trajetórias de mudança. A matriz de probabilidades de transição  $P_{ij}$  quantifica a probabilidade de um pixel na classe  $i$  no tempo  $t_1$  ser convertido para a classe  $j$  no tempo  $t_2$ , e pode ser utilizada para simular cenários futuros de uso do solo sob a premissa (frequentemente violada) de estacionariedade das taxas de transição.

#### 3.5. Séries temporais de imagens

A análise de séries temporais de imagens de satélite permite detectar tendências, sazonalidade e rupturas abruptas na dinâmica da cobertura do solo ao longo de anos ou décadas. Dois algoritmos são amplamente utilizados.

O BFAST (Breaks for Additive Season and Trend), proposto por Verbesselt et al. (2010), decompõe uma série temporal de NDVI (ou EVI) em três componentes aditivas (tendência, sazonalidade e resíduo) e detecta quebras estruturais (breakpoints) na tendência e na sazonalidade. Cada breakpoint corresponde a um evento de mudança (desmatamento, queimada, abandono, regeneração), e a magnitude e a direção da quebra informam sobre o tipo de mudança. O critério BIC (Bayesian Information Criterion) seleciona automaticamente o número ótimo de breakpoints, evitando sobreajuste.

O LandTrendr (Landsat-based Detection of Trends in Disturbance and Recovery), proposto por Kennedy et al. (2010), opera sobre a família de segmentos lineares que melhor se ajusta à série temporal de cada pixel, identificando episódios de distúrbio (segmentos de queda abrupta), recuperação (segmentos de aumento gradual) e estabilidade (segmentos horizontais). O algoritmo é particularmente eficaz para detectar eventos de distúrbio pontual (corte raso, queimada) e quantificar a velocidade de recuperação espectral.

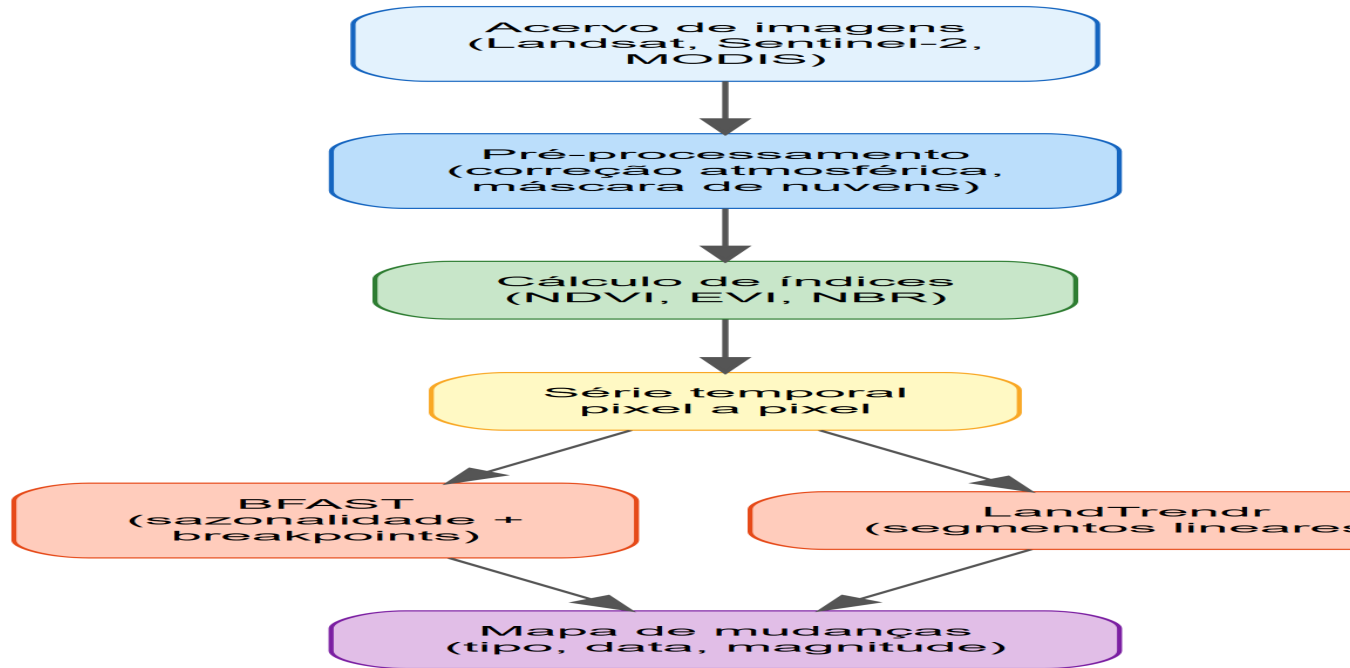


Figura 3.1.: Fluxo de processamento de séries temporais de imagens para detecção de mudanças.

### 3.6. PRODES, DETER e MapBiomias

O Brasil opera três sistemas complementares de monitoramento de mudanças de cobertura florestal que exemplificam a aplicação operacional do sensoriamento remoto.

O PRODES (Programa de Cálculo do Desflorestamento da Amazônia), operado pelo INPE desde 1988, quantifica anualmente o desmatamento por corte raso na Amazônia Legal usando imagens Landsat (30 m). Com acurácia global de 0,89 e Kappa de 0,86, o PRODES é a referência oficial para políticas públicas de combate ao desmatamento e é reconhecido internacionalmente como um dos sistemas de monitoramento florestal mais confiáveis do mundo.

O DETER (Sistema de Detecção de Desmatamento em Tempo Real) utiliza imagens de resolução moderada (MODIS, 250 m, e Sentinel-2, 10 m) para emitir alertas de desmatamento a cada 5 dias, com área mínima de detecção de 6,25 ha. Embora de menor acurácia que o PRODES (em função da resolução e frequência temporal), o DETER é o instrumento que viabiliza a fiscalização em tempo quase real pelo IBAMA.

O MapBiomias é uma iniciativa multi-institucional que produz mapas anuais de uso e cobertura do solo para todo o Brasil desde 1985, utilizando classificação por Random Forest no Google Earth Engine com resolução de 30 m e 27 classes temáticas. A acurácia global do MapBiomias evoluiu de ~0,75 na Collection 1 (2015) para ~0,87 na Collection 8 (2023), refletindo avanços metodológicos e aumento do volume de amostras de treinamento.

#### ! Complementaridade dos sistemas

O PRODES (anual, alta precisão), o DETER (quinzenal, fiscalização) e o MapBiomias (anual, reconstrução histórica, todo o país, todas as classes) não são redundantes; cada sistema atende a necessidades distintas de informação. O uso conjunto dos três permite reconstruir a trajetória histórica de mudanças (MapBiomias), quantificar o desmatamento com precisão legal (PRODES) e reagir em tempo quase real a eventos ilegais (DETER).

### **3.7. Plataformas de processamento em nuvem**

O Google Earth Engine (GEE), lançado em 2010, revolucionou o sensoriamento remoto ao oferecer acesso a petabytes de imagens de satélite (Landsat, Sentinel, MODIS, entre outros), algoritmos de processamento escaláveis e infraestrutura de computação em nuvem gratuita para pesquisa (Gorelick et al., 2017). O GEE permite análises que antes exigiam semanas de download, armazenamento e processamento local sejam executadas em minutos sobre servidores do Google, democratizando o acesso a análises de grande escala e tornando viáveis estudos planetários como o Global Forest Change (Hansen et al., 2013) e o MapBiomass.

A programação no GEE utiliza JavaScript (Editor online) ou Python (API earthengine-api), operando sobre coleções de imagens como objetos no servidor (avaliação preguiçosa), de modo que as operações são definidas localmente mas executadas na nuvem apenas quando solicitado. Essa arquitetura elimina o gargalo de transferência de dados e permite iterações rápidas durante a análise exploratória.

## 4. Classificação Digital e Avaliação de Incerteza

### 4.1. Fontes de erro na classificação

A classificação digital de imagens de satélite é um processo intrinsecamente sujeito a erros, e a avaliação de incerteza é parte indissociável de qualquer mapeamento temático rigoroso. As fontes de erro se acumulam ao longo de toda a cadeia de processamento, desde a aquisição da imagem até a produção do mapa final.

Na fase de aquisição, a resolução espacial, espectral e temporal do sensor limita os fenômenos detectáveis; pixels mistos (que contêm mais de uma classe de cobertura) são inevitáveis quando o tamanho do pixel excede o tamanho da menor feição de interesse. Na fase de pré-processamento, erros de correção atmosférica, registro geométrico imperfeito entre imagens multitemporais e falhas na remoção de nuvens e sombras propagam artefatos para as fases subsequentes. Na fase de classificação propriamente dita, a qualidade e a representatividade das amostras de treinamento são tipicamente os fatores mais críticos: amostras insuficientes, mal distribuídas no espaço ou no espaço espectral, ou rotuladas incorretamente comprometem qualquer algoritmo, por mais sofisticado que seja.

### 4.2. Amostragem de referência

O design amostral para validação de classificações deve seguir princípios estatísticos rigorosos para que as métricas de acurácia sejam estimativas não enviesadas da verdadeira acurácia do mapa. Olofsson et al. (2014) recomendam amostragem probabilística (aleatória simples ou estratificada por classe) com número mínimo de amostras determinado pela precisão desejada.

A fórmula de Cochran para amostragem aleatória simples estima o tamanho amostral necessário:

$$n = \frac{Z^2 \times p(1 - p)}{e^2}$$

onde  $Z$  é o valor crítico da distribuição normal para o nível de confiança desejado (1,96 para 95%),  $p$  é a proporção esperada de classificação correta (tipicamente 0,5 para maximizar o tamanho amostral, caso conservador) e  $e$  é a margem de erro tolerada (por exemplo, 0,05 para  $\pm 5\%$ ). Para  $p = 0,5$  e  $e = 0,05$  com 95% de confiança,  $n = 385$  amostras no total.

Na amostragem estratificada, o esforço amostral é alocado proporcionalmente à área de cada classe (alocação proporcional) ou de forma otimizada para reduzir a variância das estimativas de acurácia do produtor e do usuário (alocação ótima de Neyman). Congalton e Green (2019) recomendam um mínimo de 25–30 amostras por classe para estimativas estáveis das métricas por classe, e 50 amostras por classe quando a heterogeneidade intraclasse é elevada.

! Independência amostral

As amostras de validação devem ser espacialmente independentes das amostras de treinamento. Utilizar os mesmos pontos para treinar e validar o classificador (ou pontos muito próximos que compartilham autocorrelação espacial) inflaciona artificialmente as métricas de acurácia, produzindo avaliações irrealisticamente otimistas.

### 4.3. Matriz de confusão

A matriz de confusão (ou matriz de erro) é a ferramenta central para avaliação quantitativa da acurácia de classificações temáticas. Trata-se de uma tabela cruzada que confronta a classe atribuída pelo mapa (linhas) com a classe observada em campo ou por interpretação de referência (colunas).

A diagonal principal da matriz contém os pixels corretamente classificados. As células fora da diagonal representam erros de classificação, que se dividem em dois tipos. Os erros de comissão (falsos positivos) ocorrem quando o mapa atribui uma classe a pixels que na realidade pertencem a outra classe; são quantificados pela proporção de pixels incorretamente incluídos em cada linha. Os erros de omissão (falsos negativos) ocorrem quando o mapa falha em detectar pixels que na realidade pertencem a uma classe; são quantificados pela proporção de pixels incorretamente excluídos de cada coluna.

Três métricas derivam diretamente da matriz de confusão. A acurácia global (OA) quantifica a fração total de pixels corretamente classificados:

$$OA = \frac{\sum_{i=1}^k n_{ii}}{N}$$

onde  $n_{ii}$  é o número de amostras na diagonal (classe  $i$  mapeada e referência coincidentes) e  $N$  é o total de amostras. A acurácia do produtor (PA) para cada classe  $j$  quantifica a probabilidade de um pixel da classe  $j$  no campo ser corretamente classificado no mapa:

$$PA_j = \frac{n_{jj}}{\sum_{i=1}^k n_{ij}}$$

A acurácia do usuário (UA) para cada classe  $i$  quantifica a probabilidade de um pixel mapeado como classe  $i$  ser realmente dessa classe no campo:

$$UA_i = \frac{n_{ii}}{\sum_{j=1}^k n_{ij}}$$

💡 Interpretação prática

A PA responde à pergunta “quanto da classe real foi capturado pelo mapa?”, enquanto a UA responde “quanto do que o mapa diz é verdade?”. Um mapa pode ter PA alta para vegetação (detecta quase toda a vegetação existente) mas UA baixa (classifica erroneamente muitos pixels de solo como vegetação). Ambas as métricas devem ser avaliadas conjuntamente para cada classe.

## 4.4. Índice Kappa e suas críticas

O índice Kappa ( $\kappa$ ) de Cohen quantifica a concordância entre o mapa e a referência após descontar a concordância esperada por acaso:

$$\kappa = \frac{OA - P_e}{1 - P_e}$$

onde  $P_e$  é a concordância esperada sob classificação aleatória (calculada a partir dos totais marginais da matriz de confusão). A escala qualitativa de Landis e Koch classifica Kappa  $< 0,20$  como pobre,  $0,21-0,40$  como razoável,  $0,41-0,60$  como moderada,  $0,61-0,80$  como substancial e  $> 0,80$  como excelente.

Apesar de sua ampla utilização, o Kappa recebeu críticas substanciais. Pontius e Millones (2011) demonstraram que o Kappa é redundante com a OA (ambos contêm a mesma informação), mistura dois tipos de erro qualitativamente distintos (quantidade e alocação) e depende da prevalência das classes (um mapa dominado por uma única classe pode ter Kappa artificialmente elevado). Como alternativa, esses autores propuseram duas métricas mais informativas: o Quantity Disagreement (diferença entre as proporções de cada classe no mapa e na referência) e o Allocation Disagreement (erros de posicionamento uma vez igualadas as proporções).

A Tabela 4.1 compara as métricas de acurácia e suas propriedades.

Tabela 4.1.: Métricas de acurácia para classificação temática.

Métrica	O que mede	Vantagem	Limitação
OA	Fração global de acertos	Simples, intuitiva	Insensível a erros em classes raras
PA	Compleitude por classe	Revela omissões	Não informa sobre comissões
UA	Confiabilidade por classe	Revela comissões	Não informa sobre omissões
Kappa	Concordância ajustada ao acaso	Tradição, comparabilidade	Redundante, mistura erros
Quantity D.	Erro de proporção	Separa tipos de erro	Não localiza o erro espacialmente
Allocation D.	Erro de posição	Separa tipos de erro	Não localiza o erro espacialmente
F1-Score	Média harmônica $PA \times UA$	Balço entre PA e UA	Por classe, requer agregação

## 4.5. Protocolo de validação área-ponderada

Olofsson et al. (2014) propuseram um protocolo de validação rigoroso para estimativas de área baseadas em classificação. O protocolo reconhece que os mapas de uso e cobertura do solo são utilizados não apenas para visualização, mas para calcular áreas de cada classe (por exemplo, área desmatada, área de cultivo), e que essas estimativas de área devem incorporar a informação da matriz de confusão para corrigir o viés introduzido pelos erros de classificação.

A estimativa de área corrigida para a classe  $k$  é calculada como:

$$\hat{A}_k = \sum_{i=1}^q p_{ik} \cdot A_T$$

onde  $p_{ik}$  é a proporção estimada de pixels mapeados como classe  $i$  que na realidade pertencem à classe  $k$  (obtida da matriz de confusão estratificada),  $A_T$  é a área total mapeada e  $q$  é o número de classes. O intervalo de confiança da estimativa de área é calculado por propagação de variância, fornecendo limites superior e inferior que quantificam a incerteza.

### 4.6. Mapas de probabilidade

Os classificadores probabilísticos (como Random Forest e SVM com calibração) podem fornecer, para cada pixel, não apenas a classe mais provável, mas a probabilidade de pertencer a cada classe. O mapa de probabilidade máxima (que exibe para cada pixel a maior probabilidade entre todas as classes) é uma ferramenta de diagnóstico que identifica regiões onde o classificador é confiante (probabilidade > 0,8) e regiões onde há ambiguidade espectral entre classes (probabilidade < 0,5).

Essas regiões de baixa confiança são candidatas prioritárias para verificação de campo ou para coleta de amostras adicionais de treinamento. A integração do mapa de probabilidade com as métricas de acurácia por classe permite um diagnóstico espacialmente explícito da qualidade do mapa, superando as limitações das métricas globais (como OA e Kappa) que resumem a acurácia em um único número.

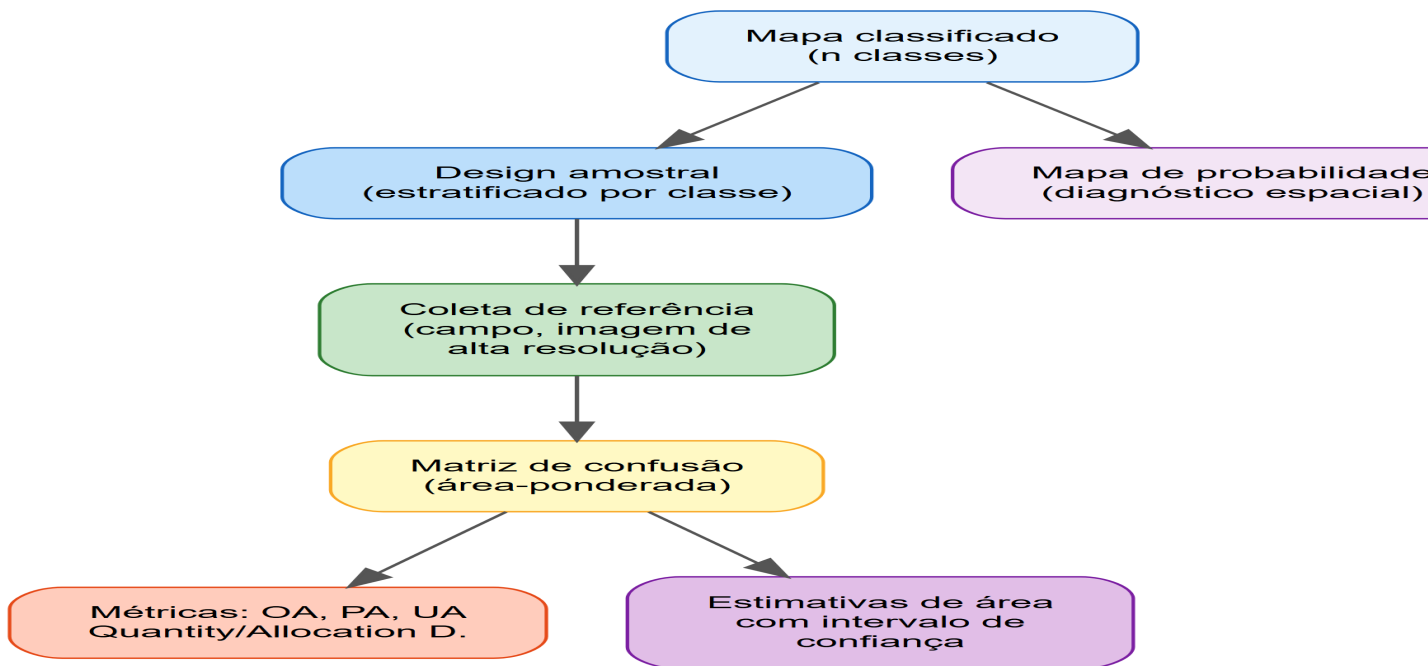


Figura 4.1.: Fluxo de validação de classificação digital conforme protocolo de Olofsson et al. (2014).

## 4.7. Google Earth Engine na prática

A implementação de um fluxo completo de classificação e validação no Google Earth Engine segue uma sequência padronizada que ilustra a integração dos conceitos apresentados neste capítulo. O processo inicia-se pela seleção de uma coleção de imagens (Sentinel-2, Landsat-8/9) filtrada por data, cobertura de nuvens e região de interesse. As bandas espectrais (B2 a B12 no caso do Sentinel-2) e índices derivados (NDVI, EVI, NDWI) compõem o espaço de feições para classificação. Um classificador Random Forest com 100 a 500 árvores é treinado sobre polígonos de referência e aplicado ao mosaico de imagens, produzindo simultaneamente o mapa classificado e o mapa de probabilidade por classe. A validação segue o protocolo de Olofsson com amostragem estratificada e cálculo de métricas área-ponderadas.

### Exercício aplicado

Um exercício completo de classificação no GEE, incluindo seleção de imagens, cálculo de índices, treinamento de Random Forest, geração de mapa de classes e avaliação de acurácia, pode ser executado em menos de 100 linhas de código JavaScript, demonstrando o poder de democratização que a plataforma trouxe ao sensoriamento remoto.

**Parte II.**

**Parte II — Solos, Erosão e Degradação  
Ambiental**

## 5. Intemperismo, Erosão e Formação de Solos

### 5.1. Fatores de formação do solo

O solo é um corpo natural tridimensional que resulta da ação combinada de cinco fatores de formação, expressos na equação conceitual de Jenny (1941):

$$S = f(cl, o, r, p, t)$$

onde *cl* representa o clima (precipitação e temperatura como motores termodinâmicos), *o* os organismos (vegetação, micro e macrofauna como agentes bioquímicos), *r* o relevo (declividade e posição na paisagem como condicionantes hidrológicos), *p* o material de origem (rocha-mãe que fornece o substrato mineral) e *t* o tempo (duração da pedogênese). A equação de Jenny não é uma função explícita resolvível, mas um modelo conceitual que orienta a interpretação das relações solo-paisagem e fundamenta a cartografia pedológica.

Na perspectiva termodinâmica, o solo é um sistema aberto e multifásico (sólida, líquida, gasosa e biológica) que troca energia e matéria com a atmosfera acima, a litosfera abaixo e os compartimentos adjacentes na paisagem. A pedogênese avança quando a energia disponível (chuva, temperatura) é suficiente para promover reações de hidrólise, oxidação e complexação que transformam minerais primários em minerais secundários (argilas e óxidos) e liberam íons para a solução do solo.

### 5.2. Processos de intemperismo

O intemperismo é o conjunto de processos físicos, químicos e biológicos que desagregam e decompõem rochas e minerais na superfície terrestre, produzindo o regolito (manto de alteração) que constitui o substrato para a pedogênese.

#### 5.2.1. Intemperismo químico

A série de estabilidade de Goldich (1938) ordena os minerais silicáticos em sequência decrescente de susceptibilidade ao intemperismo, que é essencialmente a inversa da série de cristalização de Bowen: olivina e plagioclásio cálcico (mais instáveis) intemperizam antes de muscovita e quartzo (mais estáveis). Essa sequência explica a composição mineralógica dos solos tropicais, nos quais minerais facilmente intemperizáveis (feldspatos, micas) foram quase completamente convertidos em caulinita, gibbsita e óxidos de ferro, restando quartzo como mineral resistato.

As principais reações de intemperismo químico incluem a hidrólise (ataque do H<sup>+</sup> da água aos silicatos, liberando cátions básicos e precipitando argilas), a carbonatação (dissolução de carbonatos por ácido carbônico, particularmente importante em rochas calcárias), a oxirredução (oxidação de Fe<sup>2+</sup> a Fe<sup>3+</sup>, formando goethita e hematita que conferem as colorações amarela e vermelha típicas dos solos tropicais) e a complexação orgânica (quelatagem de metais por ácidos orgânicos, promovendo a podzolização).

## 5. Intemperismo, Erosão e Formação de Solos

O Índice de Alteração Química (CIA) quantifica o grau de intemperismo de uma rocha ou solo:

$$CIA = \frac{Al_2O_3}{Al_2O_3 + CaO^* + Na_2O + K_2O} \times 100$$

onde  $CaO^*$  representa apenas o cálcio dos silicatos (excluído o dos carbonatos). Valores de CIA próximos de 50 indicam rochas frescas (feldspatos intactos), enquanto valores de 95–100 indicam intemperismo extremo (somente caulinita e óxidos), condição típica de Latossolos tropicais com mais de  $10^6$  anos de pedogênese.

### 5.2.2. Balanço de massa (Brimhall-Chadwick)

A análise de Brimhall et al. (1992) fornece um arcabouço quantitativo para rastrear ganhos, perdas e redistribuição de elementos durante o intemperismo, utilizando um elemento imóvel (tipicamente Ti ou Zr) como referência interna. O fluxo de massa ( $\tau_i$ ) de cada elemento  $i$  é calculado como:

$$\tau_i = \left( \frac{C_{i,w} \cdot \rho_w}{C_{i,p} \cdot \rho_p} \right) \cdot \frac{C_{j,p}}{C_{j,w}} - 1$$

onde  $C_{i,w}$  e  $C_{i,p}$  são as concentrações do elemento  $i$  no solo intemperizado e na rocha-mãe,  $\rho_w$  e  $\rho_p$  são as densidades correspondentes e  $j$  é o elemento imóvel. Valores de  $\tau_i > 0$  indicam enriquecimento residual (concentração por perda de outros elementos),  $\tau_i < 0$  indica depleção (perda do elemento por lixiviação) e  $\tau_i = 0$  indica imobilidade.

## 5.3. Pedogênese e horizonção

A diferenciação vertical do perfil de solo em horizontes é o resultado visível da pedogênese. O horizonte A (superficial) é enriquecido em matéria orgânica humificada e apresenta estrutura granular; o horizonte E (eluvial) é empobrecido em argila, ferro e alumínio por transloucação descendente; o horizonte Bt (textural) acumula a argila iluviada do horizonte E, apresentando cerosidade e relação textural B/A  $> 1,5$ ; e o horizonte C é o saprolito (rocha parcialmente intemperizada) que ainda preserva estrutura da rocha-mãe.

O Sistema Brasileiro de Classificação de Solos (SiBCS, EMBRAPA 2018) classifica os solos brasileiros em 13 ordens, com base na presença e nas propriedades dos horizontes diagnósticos. Os Argissolos (com horizonte B textural e gradiente de argila) e os Latossolos (com horizonte B latossólico, espesso, homogêneo e com predomínio de caulinita e óxidos) são as ordens dominantes nos ambientes tropicais brasileiros e apresentam comportamentos hidráulicos profundamente distintos.

A condutividade hidráulica saturada ( $K_s$ ), descrita pelo modelo de van Genuchten para a curva de retenção de água no solo, determina a capacidade de infiltração e o risco de geração de escoamento superficial. Em Latossolos bem estruturados,  $K_s$  pode atingir  $10^{-4}$  a  $10^{-3}$  m/s (infiltração rápida), enquanto em Argissolos com horizonte B impermeável,  $K_s$  no horizonte Bt pode ser  $10^{-7}$  a  $10^{-6}$  m/s (geração de escoamento subsuperficial lateral e saturação por rejeição).

## 5.4. Dinâmica erosiva e RUSLE

A Equação Universal de Perda de Solo Revisada (RUSLE) é o modelo empírico mais amplamente utilizado para estimativa de erosão laminar e em sulcos:

$$A = R \cdot K \cdot L \cdot S \cdot C \cdot P$$

onde  $A$  é a perda média anual de solo ( $\text{t ha}^{-1} \text{ano}^{-1}$ ),  $R$  é a erosividade da chuva ( $\text{MJ mm ha}^{-1} \text{h}^{-1} \text{ano}^{-1}$ ),  $K$  é a erodibilidade do solo ( $\text{t h MJ}^{-1} \text{mm}^{-1}$ ),  $L$  e  $S$  são os fatores de comprimento da rampa e declividade,  $C$  é o fator de uso e manejo e  $P$  é o fator de práticas conservacionistas.

O fator  $R$  é calculado a partir do índice de erosividade  $EI_{30}$ , que combina a energia cinética total da chuva com a intensidade máxima em 30 minutos. No semiárido brasileiro, a erosividade pode exceder  $7.000 \text{ MJ mm ha}^{-1} \text{h}^{-1} \text{ano}^{-1}$ , concentrada em poucos meses da estação chuvosa, quando eventos convectivos de alta intensidade geram a maior parte da perda de solo anual.

O fator  $K$  reflete a susceptibilidade intrínseca do solo à erosão, determinada por textura, estrutura, permeabilidade e teor de matéria orgânica. Neossolos Litólicos e Cambissolos rasos apresentam os maiores valores de  $K$  ( $0,04\text{--}0,06 \text{ t h MJ}^{-1} \text{mm}^{-1}$ ), enquanto Latossolos bem estruturados apresentam valores intermediários ( $0,01\text{--}0,03$ ).

O fator  $LS$  pode ser calculado a partir de MDTs utilizando o algoritmo D-infinity (Tarboton, 1997), que distribui o fluxo proporcionalmente entre as células vizinhas de descida, evitando artefatos do algoritmo D8 em vertentes planas. A integração do fator  $LS$  derivado de MDT com os demais fatores da RUSLE (obtidos de mapas temáticos) permite a espacialização da perda de solo em escala de bacia hidrográfica.

### ! Limitações da RUSLE

A RUSLE não modela erosão em ravinas, voçorocas nem deposição de sedimentos; estima apenas a erosão laminar e em sulcos para encostas retilíneas. Para erosão em canais concentrados, modelos fisicamente baseados como o WEPP (Water Erosion Prediction Project) ou o USPED (Unit Stream Power Erosion/Deposition) são necessários.

## 5.5. Integração solo-paisagem

A distribuição dos solos na paisagem segue padrões sistemáticos relacionados à posição topográfica, conceitualizados como catenas ou topossequências. Em paisagens tropicais típicas, o topo de chapada é ocupado por Latossolos (bem drenados, argilosos, profundos), a meia-encosta por Argissolos (com gradiente textural e susceptibilidade à erosão), as encostas íngremes por Neossolos Litólicos (rasos, sobre rocha) e as várzeas por Gleissolos (saturados, com acúmulo de matéria orgânica).

Essa organização catenary tem implicações diretas para o planejamento do uso da terra, pois cada posição no relevo impõe aptidões e restrições específicas. A conversão de vegetação nativa em pastagem, por exemplo, eleva o fator  $C$  da RUSLE de valores inferiores a 0,01 (vegetação densa) para valores de 0,2 a 0,6 (pastagem degradada), amplificando a perda de solo por um fator de 20 a 60 vezes. Quando essa conversão ocorre em posições de encosta com solos de gradiente textural (Argissolos), o risco de erosão linear (sulcos, ravinas) é multiplicado pela combinação de alta erodibilidade, alta declividade e redução da cobertura protetora.

## **5.6. Conservação e engenharia de mitigação**

A engenharia de mitigação da erosão opera pela manipulação dos fatores da RUSLE que são passíveis de controle humano (principalmente  $C$ ,  $P$  e, indiretamente,  $L$  e  $S$ ). As técnicas de conservação do solo incluem práticas vegetativas (plantio em curvas de nível, faixas de retenção, cobertura morta), práticas edáficas (calagem para promover agregação, incorporação de matéria orgânica) e práticas mecânicas (terraceamento, barraginhas, cordões de contorno).

A análise multicritério AHP (Analytic Hierarchy Process) de Saaty (1980) pode ser utilizada para priorização espacial de intervenções em microbacias, combinando fatores como declividade, erodibilidade, uso do solo, proximidade de corpos d'água e custo de intervenção em uma superfície de aptidão ponderada. Os pesos dos fatores são determinados por comparação pareada, e a consistência do julgamento é verificada pela razão de consistência ( $CR < 0,10$ ).

## 6. Degradação do Solo, Desertificação e Mudanças Climáticas

### 6.1. Degradação do solo

A degradação do solo compreende o conjunto de processos que reduzem a capacidade atual ou futura do solo de sustentar funções ecossistêmicas, incluindo produção de biomassa, filtragem e regulação hídrica, ciclagem de nutrientes e suporte físico à vegetação.

Diferentemente de uma simples perda de fertilidade química, a degradação envolve alterações estruturais, hidrológicas e biológicas frequentemente irreversíveis em escalas de tempo humanas, de modo que a reposição natural do material edáfico (taxas de formação de solo da ordem de 0,01–0,1 mm ano<sup>-1</sup> em regiões tropicais) é superada pelas taxas de remoção por erosão hídrica, eólica ou por processos de transporte em massa.

Os mecanismos primários de degradação podem ser agrupados em três domínios funcionais interconectados. O domínio físico engloba a erosão laminar e linear (sulcos, ravinas, voçorocas), a compactação subsuperficial (que eleva a densidade aparente e reduz a macroporosidade efetiva, diminuindo a condutividade hidráulica saturada  $K_s$ ), o selamento e encrostamento superficial (redução da infiltração inicial por colapso de agregados sob impacto de gotas de chuva) e a perda de profundidade efetiva do perfil. O domínio químico abrange a acidificação, a salinização, a sodificação (com conseqüente dispersão de argilas, conforme detalhado na seção sobre colapso estrutural), a lixiviação de bases trocáveis e a contaminação por metais pesados ou agrotóxicos. O domínio biológico compreende a redução da biomassa e diversidade microbiana, a queda da atividade enzimática (-glucosidase, fosfatase ácida, desidrogenase) e a diminuição do carbono orgânico do solo (COS), que opera como variável integradora entre os três domínios, uma vez que a matéria orgânica governa simultaneamente a estabilidade de agregados, a capacidade de troca catiônica e a atividade da biota edáfica.

Em termos quantitativos, o programa GLASOD (Global Assessment of Soil Degradation) estimou que aproximadamente 1,96 bilhão de hectares apresentavam algum grau de degradação induzida por atividades antrópicas, dos quais cerca de 60% resultavam de erosão hídrica e 28% de erosão eólica. No Brasil, a fragilidade dos Latossolos e Argissolos sob manejo inadequado (mecanização intensiva, superpastejo, ausência de cobertura vegetal permanente) configura taxas de erosão laminar que variam de 5 a 50 t ha<sup>-1</sup> ano<sup>-1</sup> em pastagens degradadas do Cerrado, valores que excedem em uma a duas ordens de magnitude as taxas de tolerância de perda de solo (geralmente entre 5 e 12 t ha<sup>-1</sup> ano<sup>-1</sup> para solos profundos tropicais).

### 6.2. Desertificação

A desertificação, conforme definida pela Convenção das Nações Unidas de Combate à Desertificação (UNCCD), consiste na degradação das terras em zonas áridas, semiáridas e subúmidas secas, resultante de variações climáticas e de atividades humanas. Essa definição circunscreve o fenômeno às regiões classificadas pelo Índice de Aridez ( $IA = P/ETP$ ) com valores inferiores

## 6. Degradação do Solo, Desertificação e Mudanças Climáticas

a 0,65, excluindo desertos hiperáridos ( $IA < 0,05$ ) e regiões úmidas, o que evidencia o caráter sinérgico entre susceptibilidade climática e pressão antrópica como condição necessária para a ocorrência do processo.

A desertificação não implica a transformação literal de áreas em desertos, mas sim a perda progressiva e potencialmente irreversível da produtividade biológica e econômica dos ecossistemas terrestres. Os vetores antrópicos dominantes incluem o desmatamento para expansão agrícola, o superpastejo (que remove a cobertura vegetal protetora e compacta a camada superficial do solo pelo pisoteio animal), a irrigação mal manejada (que promove salinização e sodificação) e a exploração excessiva de recursos hídricos subterrâneos (que rebaixa o nível freático e elimina a vegetação freatófita).

Esses vetores operam sobre uma matriz de vulnerabilidade definida pelas propriedades edáficas (textura, mineralogia, teor de matéria orgânica), pela magnitude e distribuição temporal da precipitação e pelo relevo.

No semiárido brasileiro, que abrange aproximadamente 982.563 km<sup>2</sup> e engloba porções de nove estados do Nordeste e do norte de Minas Gerais, a UNCCD e estudos nacionais identificaram quatro núcleos de desertificação em estágio avançado (Gilbués-PI, Irauçuba-CE, Seridó-RN/PB e Cabrobó-PE), caracterizados por intensa remoção da Caatinga, solos rasos e pedregosos com baixa capacidade de retenção hídrica, taxas de erosão superiores a 100 t ha<sup>-1</sup> ano<sup>-1</sup> em encostas desprotegidas e cobertura vegetal remanescente inferior a 30%.

A interação entre esses processos de degradação local e as projeções de aumento da aridez sob cenários de mudanças climáticas constitui a base conceitual para os mecanismos de retroalimentação e para o arcabouço de Neutralidade da Degradação das Terras discutidos nas seções seguintes.

### 6.3. Neutralidade da degradação das terras

O conceito de Neutralidade da Degradação das Terras (LDN, Land Degradation Neutrality), adotado pela UNCCD (Convenção das Nações Unidas de Combate à Desertificação) e incorporado à Meta 15.3 dos Objetivos de Desenvolvimento Sustentável (ODS), postula que a degradação de novas áreas deve ser compensada pela recuperação de áreas previamente degradadas, de modo que o balanço líquido de degradação seja zero ou positivo ao longo de um período definido.

A métrica tridimensional da LDN avalia simultaneamente três indicadores: a cobertura do solo (proporção e mudança de classes de uso), a produtividade primária da terra (tendência do NDVI ou EVI ao longo do tempo, medida por análise de séries temporais via BFAST ou Mann-Kendall) e o estoque de carbono orgânico do solo (COS, em toneladas de C por hectare nos primeiros 30 cm). A degradação é diagnosticada quando qualquer um dos três indicadores apresenta tendência negativa significativa.

O Índice de Aridez ( $IA$ ) classifica as terras pela relação entre precipitação e evapotranspiração potencial:

$$IA = \frac{P}{ETP}$$

onde  $P$  é a precipitação anual e  $ETP$  é a evapotranspiração potencial anual (Penman-Monteith FAO-56). Terras com  $IA < 0,65$  são classificadas como secas (áridas, semiáridas ou subúmidas secas) e estão sujeitas à Convenção de Combate à Desertificação. No semiárido brasileiro, o  $IA$

varia de 0,21 (núcleos de desertificação como Cabrobó-PE) a 0,65 (transição para subúmido), e modelos climáticos projetam expansão da área com  $IA < 0,50$  até o final do século XXI.

## 6.4. Retroalimentação erosão-clima

A relação entre erosão do solo e mudanças climáticas é bidirecional e amplificante. Por um lado, o aumento da intensidade das chuvas sob cenários de aquecimento global (IPCC AR6) eleva exponencialmente o fator  $R$  da RUSLE, pois a erosividade é proporcional à potência da intensidade de precipitação. Modelos climáticos projetam incrementos de 20–40% na erosividade em regiões tropicais até 2100, mesmo em cenários de emissão moderada (RCP 4.5).

Por outro lado, a erosão do solo retroalimenta o sistema climático ao remover o carbono orgânico da camada superficial (onde está concentrado 50–70% do COS) e expô-lo à oxidação microbiana, convertendo-o em  $CO_2$  atmosférico. Essa perda de COS não apenas contribui para o efeito estufa, mas degrada a estrutura do solo (pois a matéria orgânica é o principal agente cimentante dos agregados), reduzindo a infiltração ( $K_s$  diminui), elevando o escoamento superficial e aumentando a erodibilidade ( $K$  aumenta). Configura-se, assim, um ciclo vicioso no qual a perda de solo gera mais perda de solo e mais emissões de carbono.

### Ciclo vicioso erosão-carbono

A perda de 1 mm de solo superficial por erosão laminar equivale à remoção de aproximadamente 1,5 t C ha<sup>-1</sup> em solo rico em matéria orgânica. Em taxas de erosão de 10–50 t ha<sup>-1</sup> ano<sup>-1</sup> (comuns em pastagens degradadas no Cerrado), a perda anual de COS pode atingir 0,3–1,5 t C ha<sup>-1</sup>, valor que supera a capacidade de reposição pela vegetação em sistemas degradados.

## 6.5. Salinização e colapso da estrutura do solo

A salinização é o processo de acúmulo de sais solúveis na zona radicular ou na superfície do solo, afetando 20% das áreas irrigadas do mundo e representando uma das formas mais graves de degradação em regiões semiáridas. A Razão de Adsorção de Sódio (RAS) quantifica o risco de sodificação:

$$RAS = \frac{[Na^+]}{\sqrt{\frac{[Ca^{2+}] + [Mg^{2+}]}{2}}}$$

onde as concentrações iônicas são expressas em mmol<sub>c</sub> L<sup>-1</sup>. Valores de RAS superiores a 13 (conforme classificação de Richards (1954)) indicam risco elevado de sodificação, processo no qual o sódio trocável substitui o cálcio e o magnésio nas posições de troca das argilas, causando expansão da dupla camada difusa, dispersão dos agregados e formação de crostas superficiais impermeáveis.

O colapso da estrutura do solo por sodificação tem consequências hidrológicas severas: a condutividade hidráulica saturada pode ser reduzida em uma a duas ordens de magnitude ( $K_s$  de 10<sup>-5</sup> para 10<sup>-7</sup> m/s), eliminando a capacidade de infiltração e convertendo toda a precipitação em escoamento superficial. Em solos sodificados, a vegetação experimenta seca fisiológica mesmo sob irrigação, pois o potencial osmótico da solução do solo reduz a disponibilidade de água para as raízes.

## 6.6. Mecanismo de Charney

O mecanismo de Charney (1975) descreve a retroalimentação entre albedo de superfície e precipitação em regiões semiáridas, um processo que conecta degradação do solo a mudanças climáticas regionais. Quando a vegetação é removida e o solo exposto apresenta albedo mais alto (reflete mais radiação solar), a superfície se resfria, reduzindo a convecção e, conseqüentemente, a precipitação. A redução de precipitação intensifica a mortalidade vegetal, aumentando ainda mais o albedo. Essa retroalimentação positiva pode deslocar o equilíbrio climático local para um estado mais seco e estável (biestabilidade climática).

O modelo SEBAL (Surface Energy Balance Algorithm for Land), proposto por Bastiaanssen et al. (1998), permite estimar remotamente os componentes do balanço de energia na superfície:

$$R_n - G = H + \lambda ET$$

onde  $R_n$  é o saldo de radiação,  $G$  é o fluxo de calor no solo,  $H$  é o fluxo de calor sensível (aquecimento do ar) e  $\lambda ET$  é o fluxo de calor latente (evapotranspiração). Em superfícies degradadas,  $\lambda ET$  diminui drasticamente (pouca evapotranspiração),  $H$  aumenta (aquecimento excessivo) e a razão de Bowen ( $\beta = H/\lambda ET$ ) cresce de valores em torno de 0,2–0,5 (vegetação ativa) para 2–5 (solo exposto), alterando a dinâmica atmosférica local.

## 6.7. Indicadores bioquímicos de degradação

O monitoramento da qualidade do solo requer indicadores sensíveis e de resposta rápida que detectem a degradação antes que ela se torne irreversível. Os indicadores bioquímicos superam os indicadores físicos e químicos tradicionais em sensibilidade e tempo de resposta, pois refletem diretamente a atividade biológica do solo.

O carbono da biomassa microbiana ( $C_{bio}$ , em  $\text{g C g}^{-1}$  de solo) quantifica a massa de microrganismos vivos e é sensível a mudanças de uso do solo em escalas de meses a poucos anos. O quociente metabólico ( $qCO_2$ , em  $\text{g CO}_2\text{-C g}^{-1} C_{bio} \text{ h}^{-1}$ ) expressa a taxa de respiração por unidade de biomassa microbiana; valores elevados indicam estresse da comunidade microbiana sob condições de degradação. As enzimas -glucosidase (envolvida no ciclo do carbono) e fosfatase ácida (envolvida no ciclo do fósforo) são indicadoras da capacidade do solo de ciclagem de nutrientes, e sua atividade diminui em solos degradados.

A Tabela 6.1 sintetiza os principais indicadores e seus limiares de referência para solos tropicais.

Tabela 6.1.: Indicadores de qualidade do solo e limiares de referência para ecossistemas tropicais.

Indicador	Unidade	Solo preservado	Solo degradado	Resposta
$C_{bio}$	$\text{g C g}^{-1}$	200–600	< 100	Meses
$qCO_2$	$\text{g CO}_2\text{-C g}^{-1} C_{bio} \text{ h}^{-1}$	0,5–1,5	> 3,0	Meses
-glucosidase	$\text{g PNP g}^{-1} \text{ h}^{-1}$	50–200	< 30	Estações
Fosfatase ácida	$\text{g PNP g}^{-1} \text{ h}^{-1}$	100–400	< 60	Estações
COS (0–30 cm)	$\text{t C ha}^{-1}$	30–80	< 15	Anos–décadas

A integração desses indicadores bioquímicos com dados de sensoriamento remoto (NDVI, SAVI, albedo) e índices climáticos (SPI, SPEI) permite construir sistemas de alerta precoce para degradação de terras, nos quais tendências negativas de produtividade da vegetação detectadas

remotamente são confirmadas por indicadores de campo antes que a degradação atinja estágio irreversível.

## 6.8. Engenharia de restauração

A restauração de solos degradados exige intervenções que atuem simultaneamente sobre os processos causadores da degradação (erosão, salinização, compactação) e sobre a recomposição das funções ecossistêmicas (ciclagem de nutrientes, infiltração, armazenamento de carbono).

A recomposição vegetal é a estratégia primária, mas em solos severamente degradados, o estabelecimento de vegetação requer condicionamento prévio: subsolagem controlada para romper camadas compactadas (aumentando  $K_s$  e o volume de solo explorável pelas raízes), aplicação de gesso agrícola para corrigir a sodicidade subsuperficial (o cálcio do gesso desloca o sódio das posições de troca, restaurando a floculação das argilas), e incorporação de matéria orgânica (composto, biocarvão) para elevar a CTC, a estabilidade dos agregados e a atividade microbiana.

O Índice de Qualidade do Solo (IQS) integra múltiplos indicadores em um escore único por meio de funções de pertinência fuzzy que normalizam variáveis de diferentes unidades e escalas em um intervalo  $[0, 1]$ . A agregação por média ponderada (onde os pesos refletem a importância relativa de cada indicador para a função avaliada) produz um valor de IQS que pode ser espacializado por krigagem, gerando mapas de qualidade do solo que orientam a priorização espacial de intervenções em planos de recuperação de áreas degradadas (PRAD).

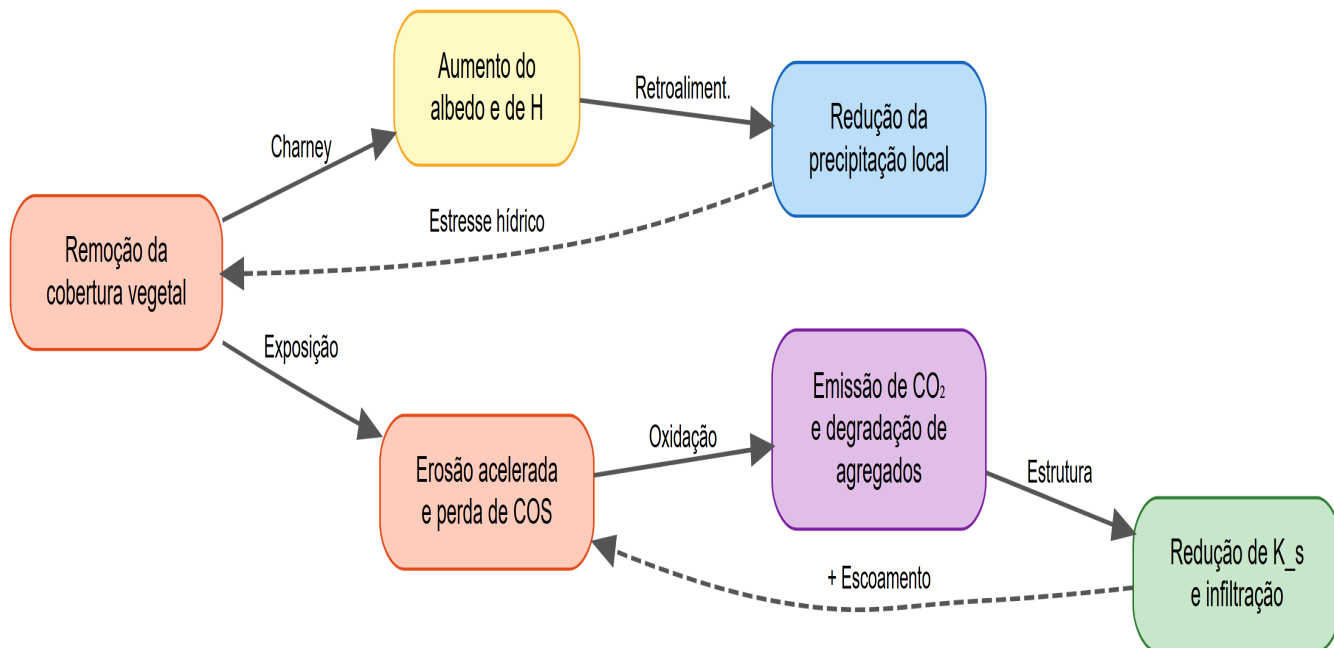


Figura 6.1.: Ciclo de retroalimentação entre degradação do solo, perda de carbono e mudanças climáticas regionais.

**Parte III.**

**Parte III — Hidrologia e Recursos  
Hídricos**

## 7. Recursos Hídricos e Gestão por Bacia Hidrográfica

### 7.1. Caracterização morfométrica da bacia

A bacia hidrográfica constitui a unidade natural de planejamento e gestão dos recursos hídricos, conforme estabelecido pela Lei 9.433/1997. A delimitação e a caracterização morfométrica de bacias, operacionalizadas em SIG a partir de Modelos Digitais de Elevação, fornecem os parâmetros físicos que condicionam a resposta hidrológica.

A hierarquia de drenagem de Strahler (1957) classifica os canais de uma rede fluvial por ordem crescente de magnitude: canais sem tributários são de primeira ordem, a confluência de dois canais de mesma ordem produz um canal de ordem superior, e a confluência de canais de ordens diferentes mantém a ordem do canal de maior magnitude. A ordem do canal principal de uma bacia é um indicador integrado da complexidade da rede de drenagem, da área de contribuição e do potencial de geração de vazão.

A Tabela 7.1 apresenta os principais índices morfométricos e suas implicações hidrológicas.

Tabela 7.1.: Índices morfométricos e suas implicações para a resposta hidrológica de bacias.

Índice	Fórmula	Significado hidrológico
Densidade de drenagem ( $D_d$ )	$D_d = \frac{L_{total}}{A}$	Disponibilidade de canais para escoamento. Valores altos indicam solos impermeáveis e resposta rápida
Fator de forma ( $K_f$ )	$K_f = \frac{A}{L^2}$	Susceptibilidade a enchentes. Bacias alongadas ( $K_f < 0,5$ ) têm menor risco
Índice de circularidade ( $IC$ )	$IC = \frac{4\pi A}{P^2}$	Proximidade à forma circular. $IC \rightarrow 1$ indica resposta rápida e concentrada
Sinuosidade ( $S_i$ )	$S_i = \frac{L_{riq}}{L_{vale}}$	Grau de meandrização. $S_i > 1,5$ indica canais sinuosos com dissipação de energia
Coefficiente de compacidade ( $K_c$ )	$K_c = 0,28 \cdot \frac{P}{\sqrt{A}}$	Relação com forma circular. $K_c \rightarrow 1$ indica maior concentração do escoamento

O tempo de concentração ( $t_c$ ), definido como o tempo necessário para que a água precipitada no ponto mais remoto da bacia atinja a seção de controle (exutório), é o parâmetro que conecta a morfometria da bacia ao dimensionamento de obras hidráulicas e sistemas de alerta. Diversas fórmulas empíricas estimam  $t_c$  a partir do comprimento do talvegue, da declividade média e da rugosidade do terreno, e a escolha da equação mais adequada depende das características da bacia e da disponibilidade de dados.

## 7.2. Balanço hídrico

O balanço hídrico de uma bacia integra todos os fluxos de entrada, saída e armazenamento de água:

$$P - E_a - Q = \frac{dS}{dt}$$

onde  $P$  é a precipitação,  $E_a$  é a evapotranspiração real,  $Q$  é a vazão (deflúvio) no exutório e  $dS/dt$  é a variação do armazenamento (em solo, aquíferos e reservatórios) ao longo do tempo. Em escala anual e para bacias em equilíbrio (sem variação de armazenamento entre anos), a equação simplifica-se para  $Q = P - E_a$ , revelando que a vazão é o resíduo entre precipitação e evapotranspiração.

Em regiões semiáridas, o balanço hídrico é caracterizado por elevados valores de  $E_a/P$  (70–90% da precipitação retorna à atmosfera por evapotranspiração), intermitência fluvial (rios secos por 4–8 meses ao ano) e forte sazonalidade (80–90% da precipitação concentrada em 3–5 meses). A cobertura vegetal exerce papel regulador ao interceptar precipitação (reduzindo o impacto das gotas sobre o solo), promover infiltração (via macroporos radiculares) e liberar água gradualmente por transpiração, atenuando picos de cheia e sustentando a vazão de base.

## 7.3. Lei 9.433/1997 (Política Nacional de Recursos Hídricos)

A Lei das Águas estabelece os fundamentos da gestão hídrica no Brasil, assentada sobre cinco princípios fundamentais: a água é um bem público de domínio da União ou dos estados; a água possui valor econômico e seu uso deve reconhecer esse valor; em situações de escassez, o abastecimento humano e a dessedentação animal têm prioridade; a gestão deve ser descentralizada com participação social; e a bacia hidrográfica é a unidade territorial de planejamento.

A lei institui cinco instrumentos de gestão, cuja operacionalização depende fortemente de geotecnologias.

O Plano de Recursos Hídricos define o diagnóstico da bacia, as metas de qualidade e quantidade, e as ações necessárias, todos baseados em dados geoespaciais (mapas de uso do solo, rede hidrográfica, localização de usuários). O enquadramento dos corpos d'água em classes de qualidade (CONAMA 357/2005) define metas de qualidade segundo o uso preponderante (abastecimento, irrigação, navegação, preservação), cuja espacialização em SIG permite identificar trechos críticos. A outorga de direito de uso autoriza a captação ou lançamento de efluentes, exigindo análise espacial da disponibilidade hídrica ( $Q_{7,10}$ , vazão mínima de 7 dias com recorrência de 10 anos). A cobrança pelo uso da água é o instrumento econômico que incentiva o uso racional, e o Sistema de Informações sobre Recursos Hídricos (SNIRH) centraliza os dados hidrológicos, meteorológicos e de qualidade da água para suporte decisório.

A equação de diluição, utilizada na análise de outorgas de lançamento de efluentes, calcula a concentração final de um poluente após a mistura com o corpo receptor:

$$C_{final} = \frac{Q_{rio} \cdot C_{rio} + Q_{efl} \cdot C_{efl}}{Q_{rio} + Q_{efl}}$$

onde  $Q_{rio}$  e  $C_{rio}$  são a vazão e a concentração do rio a montante,  $Q_{efl}$  e  $C_{efl}$  são a vazão e a concentração do efluente, e  $C_{final}$  deve atender ao limite da classe de enquadramento.

## 7.4. Governança e participação

A gestão participativa é operacionalizada pelos Comitês de Bacia Hidrográfica (CBHs), com composição tripartite (poder público, usuários e sociedade civil). As Agências de Água são os braços executivos dos comitês, responsáveis pela arrecadação da cobrança, manutenção do sistema de informações e elaboração de estudos técnicos.

### **i** Desafio da gestão em bacias intermitentes

Os instrumentos da Lei 9.433 foram concebidos para rios perenes. Em bacias intermitentes do semiárido, a outorga baseada em  $Q_{7,10}$  é frequentemente zero (o rio seca todo ano), exigindo abordagens alternativas como a alocação negociada de água, reservas estratégicas em aquíferos e a integração de fontes superficiais e subterrâneas em sistemas de gestão adaptativos.

## 7.5. Desafios contemporâneos

A gestão hídrica contemporânea opera nonexo água-energia-alimento, no qual as demandas concorrentes dos três setores devem ser harmonizadas sob restrições de disponibilidade hídrica e variabilidade climática. As secas multianuais que afetaram o Nordeste (2012–2017) e o Sudeste (2014–2015) demonstraram a vulnerabilidade do sistema hídrico brasileiro a eventos extremos cuja frequência e intensidade devem aumentar sob mudanças climáticas.

A modelagem do transporte e decaimento de poluentes em corpos hídricos, formalizada pela equação de Streeter-Phelps (que descreve o balanço entre consumo de oxigênio pela decomposição de matéria orgânica e reaeração superficial), pode ser espacializada ao longo da rede de drenagem em SIG, produzindo mapas de qualidade da água que orientam o enquadramento e a localização de pontos de monitoramento. A otimização multiobjetivo (maximizar qualidade hídrica e produção agrícola enquanto minimiza custos de tratamento) é formalizada como problema de programação não linear que integra modelos hidrológicos, modelos de qualidade da água e cenários de uso do solo.

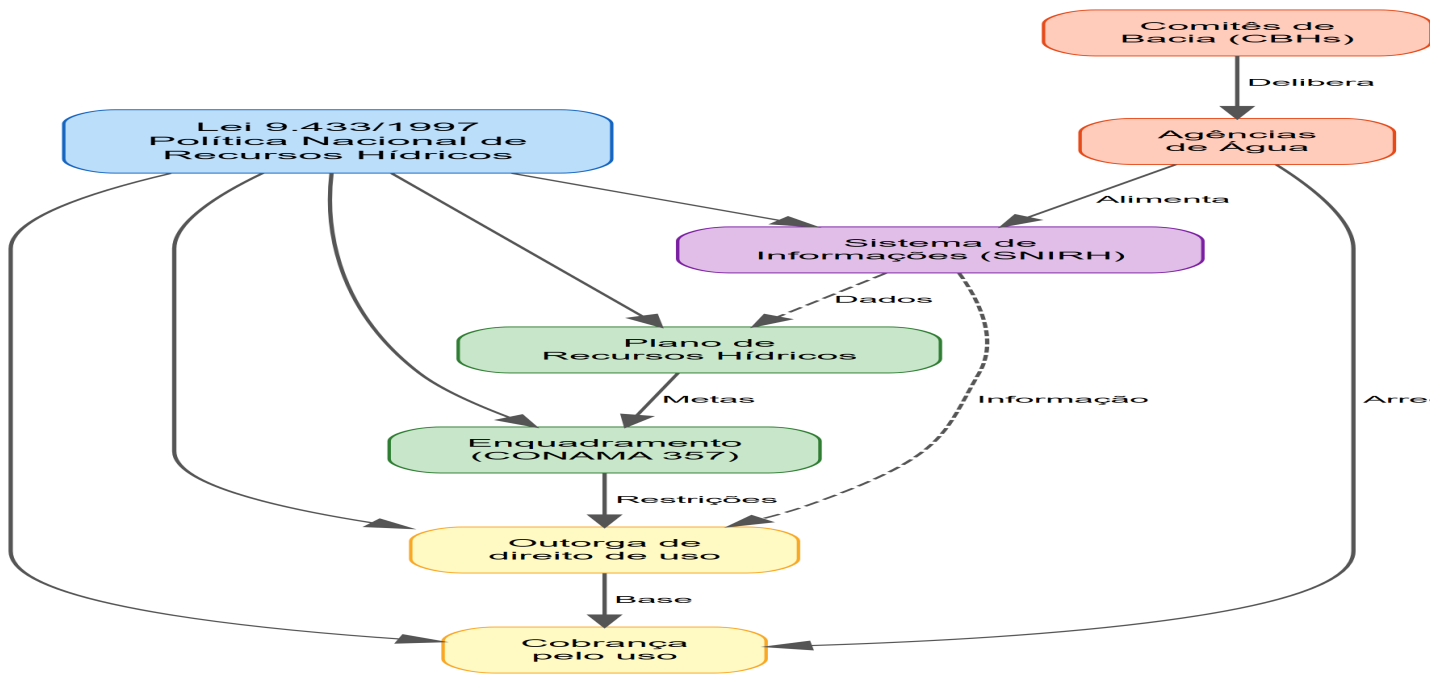


Figura 7.1.: Instrumentos da Política Nacional de Recursos Hídricos e suas inter-relações.

## 8. Monitoramento Hidrológico e Geotecnologias

### 8.1. Monitoramento convencional

A rede hidrométrica convencional brasileira, operada pela Agência Nacional de Águas (ANA) em parceria com estados e entidades delegatárias, constitui a infraestrutura fundamental para medição de variáveis hidrológicas. As estações fluviométricas medem nível d'água (régua limimétrica, limnógrafo de boia ou pressão) e vazão (molinetes hidrométricos, ADCPs acústicos), enquanto as estações pluviométricas medem precipitação (pluviômetros Ville de Paris, pluviógrafos automáticos).

O problema inverso da hidrologia (inferir variáveis distribuídas no espaço a partir de observações pontuais esparsas) é particularmente agudo em regiões tropicais com precipitação convectiva, onde células de chuva com diâmetro de 5–15 km podem gerar volumes de água que passam despercebidos por estações separadas por 50–100 km. A recomendação da WMO (2008) de uma estação pluviométrica por 250 km<sup>2</sup> (terreno plano) ou 25 km<sup>2</sup> (terreno montanhoso) está longe de ser atendida em grande parte do semiárido brasileiro, onde a densidade média da rede é inferior a uma estação por 1.000 km<sup>2</sup>.

### 8.2. Telemetria e IoT hidrológica

A transição do monitoramento manual para o automatizado e telemétrico transformou a capacidade de resposta a eventos hidrológicos extremos. A premissa operacional dos sistemas de alerta é que a latência (tempo entre a medição e a disponibilidade do dado) deve ser inferior ao tempo de concentração da bacia, permitindo que alertas sejam emitidos antes da chegada da onda de cheia.

Os sensores modernos de nível d'água incluem transdutores de pressão submersos (resolução milimétrica, aquisição a cada 5–15 min), sensores ultrassônicos (sem contato com a água) e sensores LiDAR (operação em condições de forte turbulência). Os pluviômetros automáticos de balança registram cada basculada (0,2 ou 0,25 mm), permitindo a reconstrução da hietograma com resolução sub-horária.

A transmissão de dados em tempo real utiliza protocolos de IoT (Internet of Things) dimensionados para o ambiente rural: LoRaWAN (alcance de 5–15 km em campo aberto, baixo consumo energético), redes celulares (4G/5G, em áreas com cobertura) e comunicação satelital (Iridium, GOES, para estações remotas). A convergência desses protocolos permite construir redes de monitoramento com latência inferior a 15 minutos a custos cada vez mais acessíveis.

#### **i** Latência e alerta

Em bacias urbanas com tempo de concentração inferior a 2 horas, a diferença entre 15 minutos de latência (telemetria) e 24 horas (leitura manual) determina se o sistema de alerta salva vidas ou é apenas um registro post-mortem do evento.

### 8.3. Curva-chave e hidrometria

A relação entre nível d'água ( $H$ ) e vazão ( $Q$ ) em uma seção fluviométrica é formalizada pela curva-chave, ajustada empiricamente a partir de medições simultâneas de nível e vazão em diferentes condições de escoamento:

$$Q = a(H - H_0)^b$$

onde  $a$  e  $b$  são parâmetros empíricos e  $H_0$  é o nível de referência (cota de vazão zero). A calibração requer medições cobrindo toda a faixa de variação de nível, incluindo estiagem extrema e cheias, o que em rios intermitentes do semiárido pode exigir décadas de monitoramento para capturar eventos raros.

A instabilidade da seção de medição (erosão ou assoreamento do leito, crescimento de macrófitas aquáticas, operação de barragens a montante) invalida a curva-chave e exige recalibração. A batimetria periódica (levantamento da geometria do leito por ADCP ou ecobatímetro) permite detectar mudanças na seção e atualizar os parâmetros da curva. A vazão  $Q_{7,10}$  (vazão mínima de 7 dias consecutivos com período de retorno de 10 anos), derivada da série histórica de vazões obtida via curva-chave, é o parâmetro fundamental para outorga de direito de uso.

### 8.4. Altimetria espacial por radar

A altimetria espacial por radar estende o monitoramento hidrológico a rios e reservatórios sem estações terrestres. Missões como Jason-2/3, Sentinel-3/6 e SWOT medem o nível d'água a partir da órbita por meio do tempo de ida e volta de pulsos de micro-ondas refletidos pela superfície, com precisão centimétrica após correções atmosféricas (troposfera úmida, ionosfera) e geofísicas (marés terrestres e oceânicas).

A fusão de dados orbitais (alta cobertura espacial, baixa frequência temporal) com dados terrestres (baixa cobertura espacial, alta frequência temporal) permite construir superfícies cota-tempo que descrevem a variação do nível d'água ao longo de toda a rede hidrográfica, superando a limitação fundamental das redes convencionais (que medem apenas nos pontos onde há estações). Essa abordagem é particularmente valiosa para bacias transfronteiriças e rios amazônicos, onde a instalação e manutenção de estações convencionais é logisticamente difícil.

### 8.5. Precipitação e umidade do solo por satélite

O GPM (Global Precipitation Measurement), lançado em 2014, combina radar de dupla frequência (Ku e Ka) com radiômetro de micro-ondas para estimar precipitação com resolução espaço-temporal de ~10 km e 30 minutos (produto IMERG). O CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) combina dados de infravermelho térmico de satélites geoestacionários com estações terrestres para produzir séries de precipitação com resolução de 0,05° (~5 km) e passo diário, desde 1981, sendo amplamente utilizado em regiões com baixa densidade de estações, como o Sahel e o semiárido brasileiro.

A umidade volumétrica do solo na camada superficial (0–5 cm) é estimada pela missão SMAP (Soil Moisture Active Passive, NASA) utilizando radiometria em Banda L (1,4 GHz), que penetra a cobertura vegetal moderada e é sensível ao conteúdo de água no solo. A resolução espacial do radiômetro SMAP é de ~40 km (downscaling para 9 km com dados auxiliares), e a resolução

temporal é de 2–3 dias, permitindo monitoramento quase contínuo da disponibilidade hídrica no solo para aplicações agrícolas e hidrológicas.

## 8.6. Evapotranspiração e bacias virtuais

A evapotranspiração, componente dominante do balanço hídrico em regiões semiáridas, pode ser estimada remotamente pelo algoritmo SEBAL (Surface Energy Balance Algorithm for Land) ou sua variante METRIC, que utilizam bandas termais de satélites (Landsat, ECOSTRESS) para calcular a temperatura de superfície e, a partir dela, particionar o saldo de radiação entre fluxo de calor sensível ( $H$ ) e latente ( $\lambda ET$ ).

O princípio dos pixels-âncora (um pixel “frio” com evapotranspiração máxima, tipicamente sobre vegetação irrigada ou corpo d’água, e um pixel “quente” com evapotranspiração nula, tipicamente sobre solo seco exposto) permite calibrar internamente a estimativa de  $H$  sem necessidade de dados meteorológicos locais, tornando o SEBAL aplicável a regiões com pouca infraestrutura terrestre.

A combinação de MDEs (SRTM, ALOS PALSAR), algoritmos de fluxo acumulado (D8, D-infinity) e dados de precipitação orbital (GPM, CHIRPS) permite a delimitação e modelagem de bacias hidrográficas “virtuais” (sem dados de campo), viabilizando estudos hidrológicos preliminares em qualquer região do globo. A validação dessas “bacias virtuais” contra dados observados, quando disponíveis, é condição para sua utilização em projetos de engenharia.

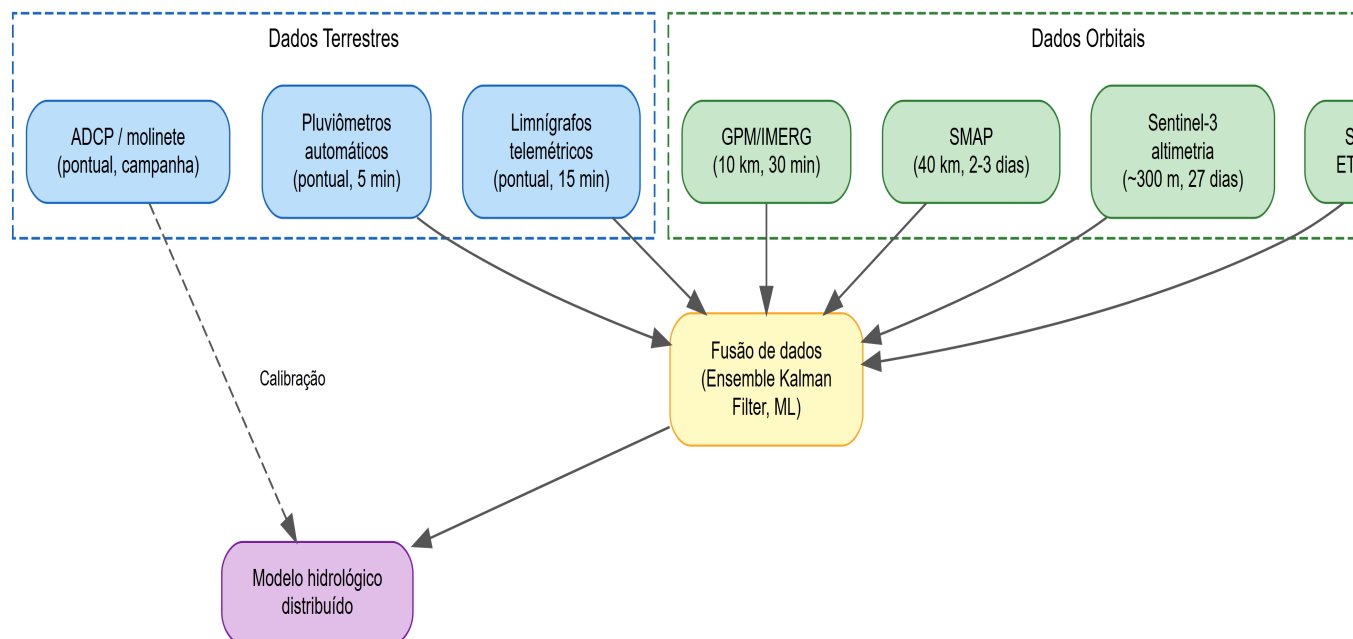


Figura 8.1.: Fontes de dados hidrológicos e suas escalas de resolução espaço-temporal.

### 💡 Ensemble Kalman Filter

O Ensemble Kalman Filter (EnKF) é a técnica mais utilizada para assimilação de dados em modelos hidrológicos distribuídos, combinando previsões do modelo com observações remotas (precipitação, umidade do solo, nível d’água) de forma sequencial e probabilística. O EnKF atualiza o estado do modelo a cada passo de tempo em que novas observações se

## 8. Monitoramento Hidrológico e Geotecnologias

tornam disponíveis, reduzindo progressivamente a incerteza da previsão.

## 9. Geotecnologias, Recursos Hídricos e Gestão da Seca

### 9.1. Crise hídrica e paradigma tecnológico

A gestão de secas no semiárido brasileiro operou historicamente sob um paradigma reativo (responder à crise após sua instalação) que está sendo progressivamente substituído por um paradigma preventivo (antecipar, monitorar e mitigar). Essa transição depende criticamente da capacidade de monitoramento contínuo e espacialmente distribuído que as geotecnologias proporcionam.

A revolução tecnológica que viabiliza essa transição combina cinco famílias de sensores orbitais: imageadores multiespectrais (Landsat, Sentinel-2, MODIS) para cobertura vegetal e albedo; radar de abertura sintética (Sentinel-1) para umidade do solo e deformação; altimetria radar (Sentinel-3, Jason-3) para nível de reservatórios e rios; gravimetria por satélite (GRACE/GRACE-FO) para variação de armazenamento total de água (incluindo aquíferos profundos); e radiometria de micro-ondas (SMAP, AMSR2) para umidade superficial e precipitação. A integração desses dados em plataformas de Big Data geoespacial (Google Earth Engine, Microsoft Planetary Computer) permite análises em escala continental com latência de dias a semanas.

### 9.2. Monitoramento de reservatórios e aquíferos

O monitoramento de reservatórios por altimetria radar (Sentinel-3, Jason-3) permite estimar variações de volume em reservatórios sem estações in situ, uma capacidade particularmente valiosa no semiárido, onde milhares de açudes de pequeno e médio porte não possuem monitoramento convencional. Quando combinada com a área superficial estimada por imagens ópticas (classificação água/não-água em Sentinel-2), a altimetria possibilita o cálculo de curva cota-área-volume e a estimativa de evaporação líquida.

A gravimetria orbital por GRACE (Gravity Recovery and Climate Experiment) e seu sucessor GRACE-FO mede variações mensais do campo gravitacional terrestre com resolução de ~300 km, detectando variações de armazenamento total de água (TWS, Total Water Storage) na escala de bacias hidrográficas. A decomposição do sinal GRACE nos compartimentos superficial (reservatórios), subsuperficial (zona vadosa) e subterrâneo (aquíferos) requer informação complementar de modelos hidrológicos ou dados de poços, mas fornece a única medição direta de variações de armazenamento subterrâneo em escala regional, sendo fundamental para monitorar a depleção de aquíferos no Cerrado (onde o aquífero Bambuí e o Urucuiá sustentam a irrigação de grande escala).

**!** ZARC e seguro agrícola

O Zoneamento Agrícola de Risco Climático (ZARC), operado pela Embrapa e pelo MAPA, utiliza modelos de balanço hídrico alimentados por séries históricas de precipitação e

propriedades do solo (CAD, Kc) para definir janelas de plantio com risco climático aceitável para cada cultura e município. O ZARC é pré-requisito para acesso ao Programa de Garantia da Atividade Agropecuária (Proagro) e ao seguro rural, demonstrando a aplicação direta de geotecnologias em políticas públicas de gestão de risco agrícola.

### 9.3. Índices de seca

A caracterização e o monitoramento de secas utilizam índices padronizados que permitem comparações entre regiões e períodos distintos.

#### 9.3.1. SPI (Standardized Precipitation Index)

O SPI, proposto por McKee et al. (1993), padroniza a precipitação acumulada em diferentes escalas temporais (1, 3, 6, 12, 24 meses) em relação à distribuição histórica, ajustando os dados a uma distribuição gama e transformando-os em uma distribuição normal padrão. Valores de SPI  $< -1$  indicam seca moderada,  $< -1,5$  seca severa e  $< -2$  seca extrema. A versatilidade do SPI reside na sua multitemporalidade: SPI-1 detecta secas meteorológicas de curta duração, SPI-3 captura impactos agrícolas, SPI-12 reflete secas hidrológicas e SPI-24 identifica secas estruturais de longo prazo.

#### 9.3.2. SPEI (Standardized Precipitation-Evapotranspiration Index)

O SPEI, proposto por Vicente-Serrano et al. (2010), é uma evolução do SPI que incorpora a evapotranspiração potencial (*ETP*) no cálculo, tornando o índice sensível ao efeito da temperatura sobre a demanda evaporativa. A variável de entrada é a diferença  $D = P - ETP$  (acumulada na escala temporal desejada), ajustada a uma distribuição log-logística e padronizada. Em cenários de aquecimento climático, o SPEI detecta agravamento de secas que o SPI (baseado apenas em precipitação) não identifica, pois o aumento de *ETP* sob temperaturas mais altas reduz a disponibilidade hídrica efetiva mesmo sem mudança na precipitação.

#### 9.3.3. Flash droughts

As secas rápidas (flash droughts) são eventos de intensificação abrupta do déficit hídrico (em escala de semanas, não meses), causados por ondas de calor que elevam a evapotranspiração muito acima da precipitação. Esses eventos escapam dos índices tradicionais baseados em acumulados mensais (SPI, SPEI) e requerem monitoramento sub-semanal integrando temperatura do ar, umidade do solo (SMAP), estresse hídrico da vegetação (índices de temperatura de superfície) e precipitação de alta frequência (GPM/IMERG).

### 9.4. Índices biofísicos de seca agrícola

Os índices de vegetação derivados de sensoriamento remoto fornecem informação espacialmente contínua sobre a condição da vegetação, a uma fração do custo do monitoramento de campo.

O VCI (Vegetation Condition Index), proposto por Kogan (1995), normaliza o NDVI de cada pixel entre os valores históricos mínimo e máximo:

$$VCI = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \times 100$$

onde  $NDVI_{min}$  e  $NDVI_{max}$  são os valores extremos da série histórica (tipicamente 20–30 anos) para aquele pixel e período do ano. Valores de  $VCI < 35\%$  indicam estresse hídrico severo. A normalização histórica permite comparar a condição da vegetação entre regiões com coberturas vegetais distintas (floresta, cerrado, caatinga, cultivo) e detectar anomalias em relação à climatologia local.

A Tabela 9.1 compara os principais índices de seca e suas aplicações.

Tabela 9.1.: Índices de seca, suas fontes de dados e tipos de seca detectados.

Índice	Base de dados	Escala temporal	Tipo de seca detectada
SPI	Precipitação	1–24 meses	Meteorológica, hidrológica
SPEI	Precipitação + ETP	1–24 meses	Meteorológica + efeito térmico
VCI	NDVI (satélite)	8–16 dias	Agrícola
TCI	Temperatura de superfície	8–16 dias	Agrícola (estresse térmico)
ESI	ET real vs. potencial	Semanal	Agrícola (flash droughts)
GRACE-DSI	Anomalia de TWS	Mensal	Hidrológica profunda

## 9.5. Evapotranspiração e seca hidrológica

O algoritmo SEBAL (Surface Energy Balance Algorithm for Land), apresentado no Capítulo 6 no contexto da retroalimentação climática, assume papel central na gestão da seca ao fornecer estimativas espacialmente distribuídas da evapotranspiração real ( $ET_a$ ). A comparação entre  $ET_a$  (o que a superfície realmente evapora) e  $ET_0$  (a demanda evaporativa atmosférica, estimada por Penman-Monteith) produz o coeficiente de estresse hídrico ( $K_s = ET_a/ET_0$ ), que identifica regiões onde a vegetação está sob déficit hídrico mesmo durante a estação chuvosa.

A seca hidrológica, definida como a redução persistente da disponibilidade de água em rios, reservatórios e aquíferos abaixo de limiares operacionais, é monitorada pela combinação de altimetria orbital (nível de reservatórios), gravimetria (GRACE, armazenamento total) e modelagem hidrológica distribuída. Redes LSTM (Long Short-Term Memory), um tipo de rede neural recorrente, têm demonstrado desempenho competitivo com modelos hidrológicos conceituais na previsão de vazões, particularmente em bacias com dados limitados de calibração.

## 9.6. Governança estratégica e resiliência

A gestão proativa da seca requer Sistemas de Suporte à Decisão (DSS) que integrem as múltiplas fontes de informação discutidas neste capítulo em uma plataforma geoespacial unificada. O DSS típico para gestão da seca combina dados meteorológicos em tempo real (estações automáticas, GPM/IMERG), índices de seca derivados de satélite (VCI, SPI, SPEI), modelos de previsão climática sazonal, informações de armazenamento de reservatórios e projeções de demanda hídrica setorial, apresentados em painéis cartográficos (dashboards geoespaciais) acessíveis a gestores e tomadores de decisão.

## 9. Geotecnologias, Recursos Hídricos e Gestão da Seca

O seguro paramétrico, uma inovação financeira em rápida expansão na agricultura, utiliza índices derivados de sensoriamento remoto (precipitação GPM, VCI, SPI) como gatilhos automáticos de indenização, eliminando a necessidade de vistoria de campo e reduzindo drasticamente o tempo entre o evento e o pagamento. A definição do gatilho (por exemplo,  $VCI < 35\%$  por duas semanas consecutivas) e da indenização (proporcional ao déficit) exige calibração cuidadosa contra dados históricos de perda agrícola, configurando um problema de otimização estocástica que integra geotecnologias, ciência atuarial e gestão de risco.

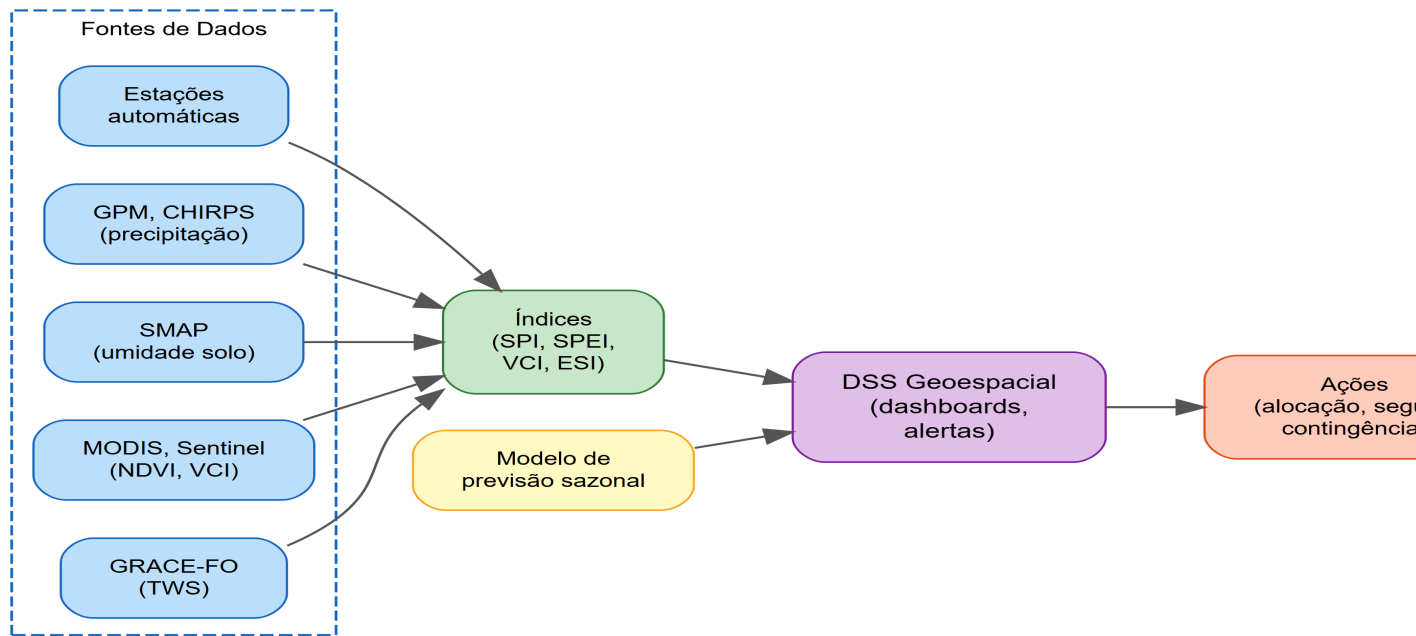


Figura 9.1.: Arquitetura de um Sistema de Suporte à Decisão (DSS) para gestão da seca integrando múltiplas fontes de dados geoespaciais.

**Parte IV.**

**Parte IV — Aplicações Setoriais e  
Modelagem Avançada**

# 10. Geoprocessamento e Monitoramento Ambiental da Mineração

## 10.1. Monitoramento contínuo como paradigma

A atividade minerária impõe transformações profundas e persistentes sobre a paisagem, demandando monitoramento ambiental contínuo que ultrapassa a capacidade de inspeções pontuais de campo. A transição do paradigma de fiscalização periódica para o monitoramento contínuo por sensoriamento remoto foi acelerada no Brasil após os desastres de Mariana (2015) e Brumadinho (2019), que evidenciaram a insuficiência dos mecanismos tradicionais de controle e condicionaram reformas regulatórias (Resolução ANM nº 13/2019) exigindo instrumentação permanente de barragens de rejeito.

O monitoramento remoto de empreendimentos minerários opera em quatro escalas complementares. A escala local cobre a cava, pilhas de estéril e barragens de rejeito com resolução submétrica (LiDAR, drones, GNSS-RTK). A escala de entorno avalia impactos em Áreas de Preservação Permanente (APPs) e unidades de conservação por imagens de média resolução (Sentinel-2, Landsat). A escala regional acompanha plumas de sedimentos e deposição em planícies de inundação. A escala de bacia hidrográfica integra os efeitos cumulativos de múltiplos empreendimentos sobre a qualidade e disponibilidade hídrica.

## 10.2. InSAR e monitoramento de deformação

A interferometria de radar de abertura sintética (InSAR/DInSAR), baseada na comparação de fases entre aquisições sucessivas do Sentinel-1, permite detectar deformações da superfície topográfica com precisão milimétrica (tipicamente 1–5 mm por ciclo de revisita de 6–12 dias). Essa capacidade é particularmente relevante para o monitoramento de estabilidade de taludes de cava, pilhas de estéril e barragens de rejeito, onde deformações precursoras podem anteceder rupturas catastróficas.

A técnica de Persistent Scatterers InSAR (PS-InSAR) identifica alvos coerentes (refletores naturais estáveis, como afloramentos rochosos e estruturas construídas) cuja fase interferométrica é rastreada ao longo de séries temporais de dezenas a centenas de imagens, gerando séries de deslocamento para cada ponto com acurácia submilimétrica. A análise SBAS (Small Baseline Subset) é preferida em áreas vegetadas onde refletores permanentes são escassos, combinando pares interferométricos de curta baseline temporal e espacial para maximizar a coerência.

### Limitações do InSAR em áreas tropicais

A coerência interferométrica é severamente degradada por vegetação densa (dossel fechado), que altera a posição dos espalhadores entre aquisições. Em minas da Amazônia e do Cerrado denso, a cobertura de PS pode ser insuficiente em áreas vegetadas, concentrando-se nas estruturas construídas e superfícies expostas. A integração de InSAR em banda C

(Sentinel-1) com bandas L (ALOS-2 PALSAR-2, futuro NISAR) mitiga parcialmente essa limitação, pois comprimentos de onda maiores penetram parcialmente no dossel.

A Tabela 10.1 resume as capacidades de monitoramento InSAR por tipo de estrutura.

Tabela 10.1.: Técnicas de InSAR e capacidades de monitoramento por tipo de estrutura minerária.

Estrutura	Técnica preferida	Resolução típica	Precisão	Período de detecção
Barragem de rejeito (aterro)	PS-InSAR	20 m	1–2 mm/ano	6–12 dias
Pilha de estéril	SBAS	20 m	3–5 mm/ciclo	12–24 dias
Talude de cava	PS-InSAR + corner reflectors	5 m (TerraSAR-X)	< 1 mm/ano	11 dias
Terreno circundante vegetado	SBAS (banda L)	25 m	5–10 mm/ciclo	14–46 dias

### 10.3. Balanço volumétrico por LiDAR e fotogrametria

O monitoramento topográfico de alta resolução por LiDAR aerotransportado (ALS) ou por fotogrametria Structure from Motion (SfM) a partir de VANT (Veículo Aéreo Não Tripulado) permite a geração de Modelos Digitais de Superfície (MDS) com resolução centimétrica. A diferença entre MDS multitemporais (DoD, DEM of Difference) quantifica com precisão o volume de material removido, depositado ou deslocado em cavas, pilhas e barragens.

O balanço volumétrico é expresso por:

$$\Delta V = \sum_{i=1}^n (Z_{t_2}(i) - Z_{t_1}(i)) \cdot \Delta x \cdot \Delta y$$

onde  $Z_{t_1}(i)$  e  $Z_{t_2}(i)$  são as elevações do pixel  $i$  nos tempos  $t_1$  e  $t_2$ ,  $\Delta x$  e  $\Delta y$  são as dimensões do pixel, e  $n$  é o número total de pixels na área de interesse. A propagação de incerteza considera o erro vertical de cada MDS ( $\sigma_Z$ ), resultando em um limite de detecção mínimo de  $\pm 2\sigma_Z$  para mudanças significativas. Para LiDAR-ALS,  $\sigma_Z$  típico é de 5–15 cm; para SfM-VANT, 2–10 cm dependendo da altitude de voo e da densidade de pontos de controle (GCP).

### 10.4. Recuperação ambiental e histerese vegetacional

A recuperação de áreas degradadas por mineração é monitorada por séries temporais de NDVI, cuja trajetória temporal revela a resiliência do ecossistema. Em regiões de Cerrado e Mata Atlântica, a recuperação do NDVI após revegetação de taludes e pilhas de estéril exibe um padrão de histerese caracterizado por três fases. A fase inicial (0–2 anos) apresenta aumento rápido do NDVI associado ao estabelecimento de espécies herbáceas de cobertura rápida (gramíneas, leguminosas rasteiras), com valores típicos de NDVI de 0,3–0,5. A fase intermediária (2–5 anos) exibe plateau ou ligeira redução do NDVI, associada à competição interespecífica e à transição de comunidade herbácea para arbustiva, com mortalidade diferencial. A fase de maturação (5–15

anos) mostra incremento gradual até valores próximos da vegetação nativa circundante (NDVI 0,6–0,8), representando o estabelecimento de espécies arbóreas e a estratificação do dossel.

**i** Histerese e resiliência

O conceito de histerese implica que a trajetória de recuperação não segue o caminho inverso da degradação. Áreas que foram completamente desmatadas e tiveram o solo removido até o saprolito requerem intervenções de reconstrução de solo (adição de substrato, inoculação microbiana, adubação verde) muito mais intensas do que áreas com solo superficial preservado, e a recuperação até níveis equivalentes de NDVI e biomassa pode exigir prazos significativamente mais longos (10–20 anos vs. 3–5 anos).

## 10.5. Transporte de sedimentos e qualidade da água

A ruptura de barragens de rejeito e a erosão de áreas expostas pela mineração geram plumas de sedimentos cuja dispersão em rios e reservatórios é monitorada por sensoriamento remoto multiespectral. A concentração de sólidos suspensos totais (SST) é estimada por modelos empíricos ou semi-empíricos que relacionam a reflectância da água nas bandas do vermelho e infravermelho próximo com medições in situ de turbidez e SST.

O modelo de advecção-dispersão unidimensional para transporte de poluentes em rios é descrito pela equação:

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} = D_L \frac{\partial^2 C}{\partial x^2} - k \cdot C + S$$

onde  $C$  é a concentração do poluente (mg/L),  $u$  a velocidade média do escoamento (m/s),  $D_L$  o coeficiente de dispersão longitudinal (m<sup>2</sup>/s),  $k$  a taxa de decaimento de primeira ordem (1/s) e  $S$  o termo-fonte (mg/L/s). Para metais pesados associados a rejeitos de mineração de ferro (mangânês, arsênio, cromo), o decaimento é predominantemente por sedimentação (depositional sink) e adsorção a partículas finas, e o coeficiente  $k$  é calibrado com dados de campo de campanhas pós-ruptura.

A Tabela 10.2 apresenta os parâmetros monitorados e as fontes de dados utilizadas.

Tabela 10.2.: Parâmetros de qualidade de água e fontes de dados para monitoramento em áreas de mineração.

Parâmetro	Sensor/Método	Frequência	Resolução espacial
SST (turbidez)	Sentinel-2 (B4/B8A)	5 dias	10–20 m
Clorofila-a	Sentinel-2 (B4/B5)	5 dias	20 m
Metais pesados	Coleta in situ (ICP-OES)	Mensal	Pontual
Nível d'água	Sentinel-3 (altimetria)	27 dias	300 m (ao longo da órbita)
Extensão da pluma	Landsat/Sentinel-2 (RGB+NIR)	5–16 dias	10–30 m

## 10.6. Gestão de risco e governança ambiental

A avaliação integrada de risco ambiental em áreas de mineração utiliza modelos multicritério em ambiente SIG. A lógica fuzzy, apresentada no Capítulo 2 no contexto da álgebra de mapas, é aplicada para transformar variáveis heterogêneas (distância à barragem, declividade, permeabilidade, proximidade de cursos d'água, ocupação do vale a jusante) em graus de pertinência (0–1) e combiná-las pelo operador Fuzzy Gamma:

$$\mu_{gamma} = (\mu_{soma})^{\gamma} \cdot (\mu_{produto})^{1-\gamma}$$

onde  $\gamma$  tipicamente entre 0,7 e 0,9 balanceia características ampliativas e restritivas do modelo, produzindo mapas contínuos de risco que orientam planos de emergência e evacuação.

A governança ambiental moderna da mineração requer interoperabilidade entre gestores (IBAMA, órgãos estaduais, ANM), operadores e comunidades. Plataformas de cadastro ambiental georreferenciado (SIGBM, da ANM) integram informações de instrumentação de barragens (piezômetros, inclinômetros, marcos geodésicos), dados de InSAR, imagens de satélite e resultados de monitoramento de qualidade de água em painéis cartográficos de acesso público, concretizando o princípio de transparência e ampliando a capacidade de fiscalização da sociedade civil.

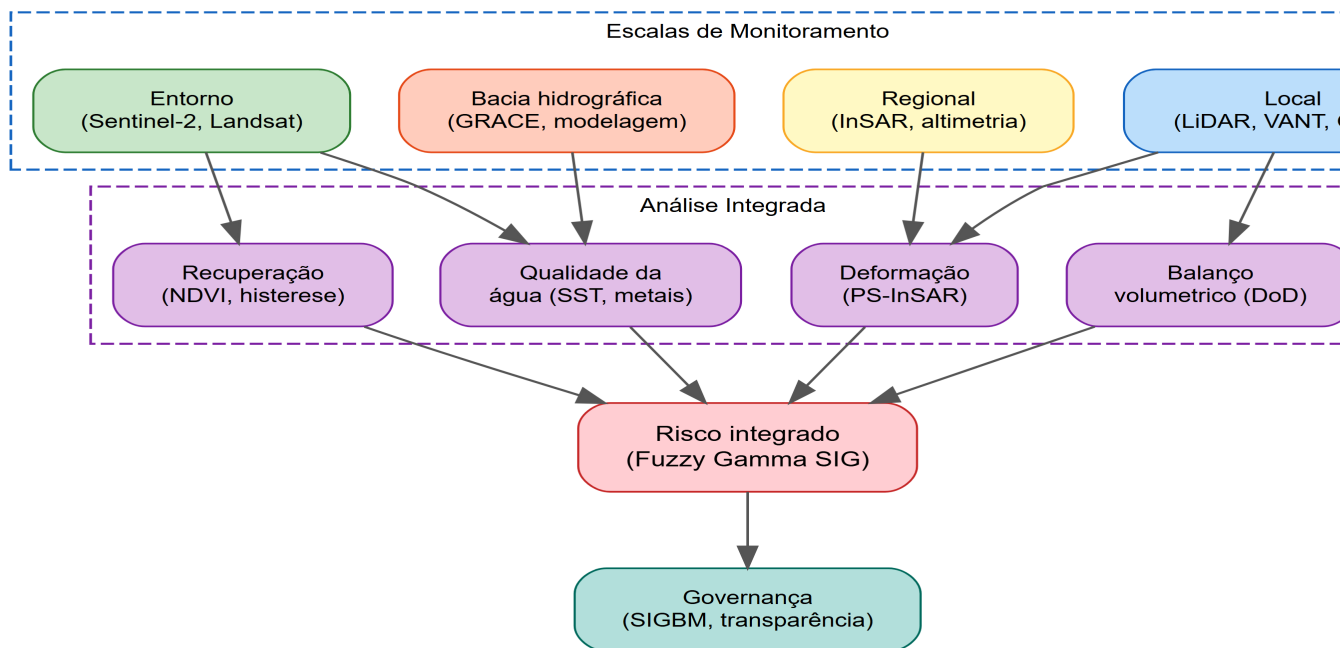


Figura 10.1.: Cadeia de monitoramento ambiental integrado para empreendimentos minerários, combinando múltiplas escalas e fontes de dados.

# 11. SIG Aplicado à Exploração Mineral e Modelagem Preditiva

## 11.1. Sistemas minerais e modelagem prospectiva

A exploração mineral moderna opera sob o conceito de sistemas minerais (mineral systems), proposto por Wyborn et al. (1994), que identifica os processos geológicos essenciais à formação de um depósito mineral: fonte de fluidos e metais, mecanismo de transporte (falhas, zonas de cisalhamento, contatos litológicos), e sítio de deposição (armadilha estrutural ou química). Cada componente do sistema mineral gera assinaturas geológicas, geofísicas e geoquímicas que podem ser mapeadas, e a modelagem prospectiva em SIG consiste em integrar múltiplas camadas de evidência para delimitar áreas com maior probabilidade de conter mineralizações econômicas.

O fluxo de trabalho de exploração mineral baseado em SIG segue uma sequência hierárquica de redução de área: reconhecimento regional (escala 1:1.000.000 a 1:250.000) com dados aerogeofísicos e geologia regional; seleção de alvos (1:100.000 a 1:50.000) com sensoriamento remoto, geoquímica de solo/sedimento e geofísica terrestre; e detalhamento de alvos (1:25.000 a 1:5.000) com mapeamento geológico de detalhe, furos de sondagem e modelagem tridimensional. Em cada escala, o SIG integra camadas de informação por técnicas quantitativas de modelagem espacial.

## 11.2. Dados aerogeofísicos

### 11.2.1. Aeromagnetometria

O levantamento aeromagnético mede as variações do campo magnético terrestre causadas por minerais ferrimagnéticos (magnetita, pirrotita, ilmenita) presentes nas rochas. Os dados são processados para remoção do campo geomagnético internacional de referência (IGRF), produzindo o campo magnético anômalo, que é interpretado em termos de unidades litológicas e estruturas geológicas. Produtos derivados incluem a amplitude do sinal analítico (ASA), que localiza bordas de corpos magnéticos independentemente da direção de magnetização e da inclinação magnética, e as derivadas direcionais (dx, dy, dz), que realçam contatos e falhas segundo sua orientação.

### 11.2.2. Gamaespectrometria

A gamaespectrometria aérea mede a radiação gama natural emitida por isótopos de potássio ( $^{40}\text{K}$ ), urânio ( $^{238}\text{U}$ , medido pelo filho  $^{214}\text{Bi}$ ) e tório ( $^{232}\text{Th}$ , medido pelo filho  $^{208}\text{Tl}$ ) nos 30–45 cm superficiais do solo e rocha. A razão Th/K discrimina unidades geológicas com assinaturas distintas (granitos peraluminosos apresentam alto K e moderado Th, enquanto quartzitos mostram baixo K e alto Th). A razão U/Th identifica anomalias redox associadas a mineralizações uraníferas e hidrotermais. A integração de K, eU, eTh na composição ternária RGB produz mapas de domínios radiométricos que auxiliam no mapeamento litológico e na identificação de zonas de alteração hidrotermal.

 Dados públicos no Brasil

A CPRM (Serviço Geológico do Brasil) disponibiliza dados aerogeofísicos (magnetometria e gamaespectrometria) cobrindo grande parte do território nacional, através do portal GeoSGB (geosgb.cprm.gov.br). Os dados do Programa de Levantamento Aerogeofísico (PLGA) estão em formato GDB e XYZ, prontos para importação em SIG. O CPRM também disponibiliza mapas geológicos no padrão GeoSGB, cartas geoquímicas e dados de recursos minerais.

### 11.3. Sensoriamento remoto hiperespectral

Enquanto sensores multiespectrais (Sentinel-2, Landsat) registram a reflectância em dezenas de bandas relativamente largas, sensores hiperespectrais (AVIRIS, EnMAP, PRISMA, Hyperion, ASTER-SWIR) adquirem centenas de bandas estreitas e contíguas, permitindo a identificação de minerais específicos por suas feições de absorção diagnósticas no espectro eletromagnético. A goethita apresenta feições de absorção em  $\sim 0,48$  e  $\sim 0,90$   $\mu\text{m}$ , a hematita em  $\sim 0,53$  e  $\sim 0,86$   $\mu\text{m}$ , a caulinita em  $\sim 2,16$  e  $\sim 2,21$   $\mu\text{m}$ , e a montmorilonita em  $\sim 2,20$   $\mu\text{m}$ . Esses minerais de alteração são indicadores (pathfinder minerals) de processos hidrotermais associados a mineralizações.

O algoritmo SAM (Spectral Angle Mapper), proposto por Kruse et al. (1993), classifica pixels por comparação angular entre o espectro observado e espectros de referência (bibliotecas espectrais da USGS/JPL). O ângulo espectral  $\alpha$  entre dois espectros  $t$  (alvo) e  $r$  (referência), cada um vetorizado em  $n$  bandas, é calculado por:

$$\alpha = \cos^{-1} \left( \frac{\sum_{i=1}^n t_i \cdot r_i}{\sqrt{\sum_{i=1}^n t_i^2} \cdot \sqrt{\sum_{i=1}^n r_i^2}} \right)$$

Valores de  $\alpha$  próximos de zero indicam alta similaridade espectral. A vantagem do SAM é sua insensibilidade a variações de iluminação (albedo), pois o ângulo depende apenas da forma do espectro e não de sua magnitude, sendo particularmente útil em terrenos com topografia acidentada.

A Tabela 11.1 apresenta minerais indicadores e suas feições diagnósticas.

Tabela 11.1.: Minerais indicadores de processos hidrotermais e suas feições espectrais diagnósticas.

Mineral	Grupo	Feição diagnóstica ( $\mu\text{m}$ )	Processo indicado
Caulinita	Argilomine- ral	2,16; 2,21 (OH)	Alteração argílica
Montmorilo- nita	Argilomine- ral	2,20 (OH + H O)	Alteração argílica
Sericita/Mus- covita	Filossilili- cato	2,20; 2,35 (Al-OH)	Alteração fílica
Clorita	Filossilili- cato	2,25; 2,33 (Mg-OH)	Alteração propilítica
Epidoto	Sorossilili- cato	2,25; 2,33	Alteração propilítica
Goethita	Óxido de Fe	0,48; 0,90	Intemperismo/gossan

Mineral	Grupo	Feição diagnóstica ( m)	Processo indicado
Hematita	Óxido de Fe	0,53; 0,86	Oxidação/gossan
Carbonato (calcita)	Carbo-nato	2,33; 2,53 (CO )	Alteração carbonática

#### 11.4. Modelagem multicritério e AHP

A modelagem de favorabilidade mineral em SIG combina múltiplas camadas de evidência (geologia, geofísica, geoquímica, sensoriamento remoto, estrutural) em um mapa-síntese de potencial mineral. O Processo Analítico Hierárquico (AHP), proposto por Saaty (1980), é amplamente utilizado para ponderar a importância relativa de cada camada de evidência a partir de comparações pareadas entre critérios, avaliadas por especialistas em uma escala de 1 (igual importância) a 9 (importância absoluta). A matriz de comparação pareada  $A = [a_{ij}]$  é normalizada e o autovetor principal fornece os pesos relativos  $w_i$  de cada critério.

A consistência lógica das comparações é verificada pelo Índice de Consistência ( $CI$ ) e pela Razão de Consistência ( $CR$ ):

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

$$CR = \frac{CI}{RI_n}$$

onde  $\lambda_{max}$  é o autovalor principal da matriz,  $n$  o número de critérios e  $RI_n$  o índice de consistência aleatório (tabelado para cada  $n$ ). Matrizes com  $CR > 0,10$  são consideradas inconsistentes e devem ser revisadas. Uma vez obtidos os pesos, os mapas-critério (já reclassificados para uma escala comum, por exemplo, fuzzy 0–1) são combinados pela soma ponderada:

$$FP = \sum_{i=1}^n w_i \cdot \mu_i$$

onde  $FP$  é o índice de favorabilidade e  $\mu_i$  o grau de pertinência fuzzy do  $i$ -ésimo critério.

#### 11.5. Weights of Evidence (WofE)

O método pesos de evidência (Weights of Evidence, WofE), aplicado à exploração mineral por Bonham-Carter (1994), é fundamentado no Teorema de Bayes e utiliza depósitos minerais conhecidos (training points) para calcular a associação espacial entre cada mapa-evidência (binarizado) e a ocorrência mineral. Para cada mapa binário  $B$  (por exemplo, presença/ausência de uma falha em um raio de 500 m), os pesos positivo ( $W^+$ ) e negativo ( $W^-$ ) são calculados como:

$$W^+ = \ln \frac{P(B|D)}{P(B|\bar{D})}$$

$$W^- = \ln \frac{P(\bar{B}|D)}{P(\bar{B}|\bar{D})}$$

onde  $D$  denota a presença de um depósito e  $\bar{D}$  sua ausência. O contraste  $C = W^+ - W^-$  quantifica a força da associação espacial. Valores de  $C > 0$  indicam associação positiva (o padrão favorece a ocorrência), e  $C < 0$  indicam associação negativa. A significância estatística do contraste é avaliada pela razão  $C/s(C)$ , onde  $s(C)$  é o desvio-padrão estimado por propagação de erros.

A probabilidade posterior de ocorrência mineral em cada célula do grid é calculada pela combinação log-linear dos pesos de todas as camadas de evidência:

$$\ln O_{post} = \ln O_{prior} + \sum_j W_j$$

onde  $O$  denota odds (razão de chances) e  $W_j$  é o peso ( $W^+$  ou  $W^-$ ) do  $j$ -ésimo mapa conforme a presença ou ausência do padrão naquela célula.

#### ! Pressuposto de independência condicional

A validade do WofE requer independência condicional entre os mapas-evidência dado o evento de depósito. A violação dessa premissa (por exemplo, correlação entre mapas geológicos e geofísicos derivados das mesmas unidades) infla a probabilidade posterior. O teste da independência condicional é realizado pelo qui-quadrado ( $\chi^2$ ) ou pela razão de Agterberg-Cheng, comparando o número de depósitos previstos pelo modelo com o observado. Violações moderadas são comuns na prática e devem ser reportadas e mitigadas (por exemplo, removendo camadas redundantes ou substituindo por combinações AHP onde a independência não é atendida).

## 11.6. Modelagem por aprendizado de máquina

As abordagens de aprendizado de máquina (Random Forest, SVM, redes neurais, gradient boosting) para modelagem de favorabilidade mineral operam como mapeadores não-lineares entre as camadas de evidência (features) e a presença/ausência de depósitos minerais (variável-alvo), sem pressupor independência condicional entre evidências e sem exigir binarização prévia dos mapas.

A avaliação do desempenho preditivo de qualquer modelo de favorabilidade (AHP, WofE, ML) utiliza a curva ROC (Receiver Operating Characteristic) e a área sob a curva (AUC, Area Under the Curve). A AUC mede a capacidade do modelo de discriminar entre células com e sem depósitos, variando de 0,5 (discriminação aleatória) a 1,0 (discriminação perfeita). Valores de AUC  $> 0,80$  indicam desempenho aceitável, enquanto AUC  $> 0,90$  sugere excelente capacidade preditiva. Para modelos de ML, a validação deve utilizar conjuntos de teste espacialmente independentes (spatial cross-validation, como Leave-Location-Out, discutido no Capítulo 12) para evitar inflação de AUC por autocorrelação espacial.

A Tabela 11.2 compara os métodos discutidos.

Tabela 11.2.: Comparação entre métodos de modelagem prospectiva mineral em SIG.

Método	Fundamentação	Pesos	Independência requerida	Supervisão
AHP + Fuzzy	Decisão multicritério	Especialista	Não	Não
WofE	Bayesiano	Dados (depósitos)	Sim (condicional)	Sim
Random Forest	Ensemble de árvores	Dados	Não	Sim
SVM	Máxima margem	Dados	Não	Sim

### 11.7. Restrições ambientais e planejamento integrado

A modelagem de favorabilidade mineral não pode ser dissociada das restrições ambientais e legais que condicionam a viabilidade de exploração. Em ambiente SIG, as áreas de exclusão (unidades de conservação de proteção integral, terras indígenas, quilombos, APPs) são representadas como camadas de restrição booleana que eliminam porções do mapa de favorabilidade, produzindo o mapa de áreas prospectáveis (favorabilidade mineral menos restrições legais). Restrições parciais (APA, RPPN, zonas de amortecimento) podem ser incorporadas como fatores de ponderação negativa em modelos fuzzy, reduzindo a favorabilidade sem anulá-la completamente.

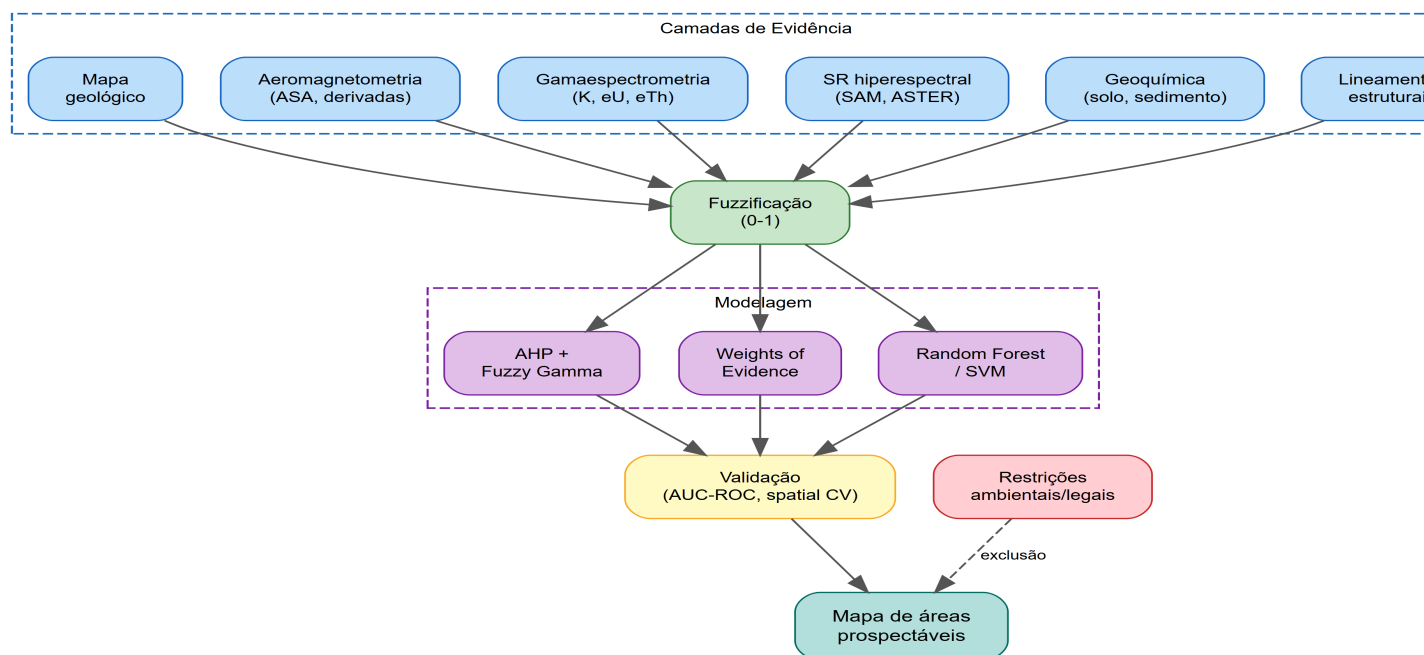


Figura 11.1.: Fluxo de trabalho da modelagem de favorabilidade mineral em SIG, desde a aquisição de dados até a integração com restrições ambientais.

# 12. Inteligência Artificial e Qualidade da Pesquisa Ambiental

## 12.1. O quarto paradigma da ciência

A evolução do método científico é descrita por quatro paradigmas sequenciais, conforme proposto por Hey et al. (2009). O primeiro paradigma (empírico) baseava-se na observação e descrição de fenômenos naturais. O segundo (teórico) introduziu modelos matemáticos e leis fundamentais (Newton, Maxwell, Navier-Stokes). O terceiro (computacional) emergiu com a simulação numérica de processos complexos demais para solução analítica. O quarto paradigma (data-driven ou eScience) é definido pela exploração de grandes volumes de dados heterogêneos por algoritmos de aprendizado de máquina, onde padrões e relações emergem dos dados sem a imposição a priori de equações governantes. Nas geociências ambientais, esse paradigma se concretiza na integração massiva de dados de sensoriamento remoto, redes de monitoramento, modelos climáticos e informações de campo em plataformas computacionais de alta performance (Google Earth Engine, PANGAEA, Copernicus Climate Data Store).

## 12.2. Modelagem híbrida

A tensão entre modelos baseados em processos (process-based) e modelos orientados a dados (data-driven) motivou o desenvolvimento da modelagem híbrida (theory-guided data science), formalizada por Karpatne et al. (2017). A premissa é que o conhecimento do domínio (equações de conservação, leis termodinâmicas, relações empíricas calibradas) deve restringir o espaço de soluções do modelo de aprendizado de máquina, evitando previsões fisicamente impossíveis e melhorando a extrapolação para condições não observadas no treinamento.

Três estratégias de hibridização são amplamente utilizadas. A abordagem de pré-processamento incorpora o conhecimento físico na engenharia de features: em vez de alimentar o modelo com reflectância bruta, calcula-se NDVI, SAVI, índices de umidade, derivações do MDT (declividade, curvatura, TWI), fator LS da RUSLE. A abordagem de restrição incorpora a física como termo de penalização na função-objetivo do modelo: por exemplo, penalizar previsões de erosão líquida negativa (criação de solo ex nihilo) ou previsões de evapotranspiração que violem o balanço de energia ( $R_n - G = H + \lambda ET$ ). A abordagem de pós-processamento ajusta as previsões do modelo de ML para satisfazer restrições de conservação (massa, energia, momento), assegurando que os fluxos previstos fechem o balanço na bacia.

### **i** Exemplo de hibridização em hidrologia

O modelo LSTM (Long Short-Term Memory) treinado para previsão de vazão em bacias hidrográficas pode ser hibridizado incorporando o balanço hídrico como restrição. A perda de treinamento  $\mathcal{L}$  é composta por um termo de erro de previsão (ex. MSE entre vazão observada e prevista) acrescido de um termo de penalização física  $\lambda_{phys} \cdot |P - ET - Q - \Delta S|^2$ , onde  $P$  é precipitação,  $ET$  evapotranspiração,  $Q$  vazão e  $\Delta S$  variação de armazenamento.

O hiperparâmetro  $\lambda_{phys}$  controla o peso da restrição física na função-objetivo.

### 12.3. Validação espacial

A validação de modelos preditivos em geociências requer cuidados específicos relacionados à autocorrelação espacial dos dados, conforme demonstrado por Roberts et al. (2017). A validação cruzada convencional ( $k$ -fold aleatória) distribui observações de treinamento e teste de forma espacialmente intercalada, permitindo que o modelo aprenda a autocorrelação espacial (dados próximos são semelhantes) em vez de relações causais genuínas. O resultado é uma inflação sistemática das métricas de desempenho (acurácia, AUC,  $R^2$ ) que não reflete a capacidade real de generalização do modelo para novas áreas.

A validação cruzada espacial (spatial cross-validation) resolve esse problema agrupando as observações por blocos espaciais (spatial blocking) ou excluindo toda a vizinhança de cada observação de teste (Leave-Location-Out, LLO). O tamanho dos blocos deve ser proporcional ao range do semivariograma dos resíduos do modelo (apresentado no Capítulo 2), assegurando independência efetiva entre conjuntos de treinamento e teste. Em estudos de mapeamento de uso do solo, Roberts et al. (2017) reportaram reduções de AUC de 0,10 a 0,25 quando spatial CV foi utilizada em substituição a  $k$ -fold aleatória, demonstrando que a inflação de desempenho é substancial e pode conduzir a decisões de gestão baseadas em modelos com capacidade preditiva real significativamente inferior à reportada.

A Tabela 12.1 resume os protocolos de validação e seus vieses.

Tabela 12.1.: Protocolos de validação e seus vieses associados à autocorrelação espacial.

Protocolo	Descrição	Viés por autocorrelação	Uso recomendado
$k$ -fold aleatória	Partições aleatórias sem estrutura espacial	Alto (otimista)	Dados independentes (lab)
Spatial blocking	Blocos geográficos como folds	Baixo	Modelos espaciais
Leave-Location-Out (LLO)	Exclui vizinhança do ponto-teste	Muito baixo	Mapeamento preditivo
Leave-Time-Out (LTO)	Treina no passado, testa no futuro	Baixo (temporal)	Séries temporais
Leave-Location-and-Time-Out	Exclui espaço e tempo simultaneamente	Mínimo	Modelos espaço-temporais

### 12.4. Explicabilidade e confiança

O crescente uso de modelos de aprendizado de máquina de alta complexidade (deep learning, gradient boosting, ensemble stacking) nas geociências levantou preocupações legítimas sobre

interpretabilidade e confiança nas previsões. Modelos caixa-preta (black box) produzem previsões acuradas sem revelar quais variáveis e mecanismos os conduzem, dificultando a identificação de erros, a comunicação com gestores e a integração com o conhecimento do domínio.

A IA Explicável (Explainable AI, XAI) aborda essa lacuna por duas famílias de métodos. Métodos intrínsecos utilizam arquiteturas inerentemente interpretáveis (regressão linear, árvore de decisão, modelos aditivos generalizados — GAM). Métodos post-hoc explicam modelos complexos já treinados, atribuindo importância a cada variável preditora para cada previsão individual.

O SHAP (SHapley Additive exPlanations), proposto por Lundberg e Lee (2017) e fundamentado na teoria de jogos cooperativos de Shapley, decompõe a previsão  $f(x)$  de cada instância  $x$  em contribuições individuais de cada feature  $j$ :

$$f(x) = \phi_0 + \sum_{j=1}^M \phi_j$$

onde  $\phi_0$  é o valor base (previsão média sobre o conjunto de treinamento) e  $\phi_j$  é o valor de Shapley da  $j$ -ésima feature. Os valores de Shapley possuem propriedades axiomáticas desejáveis (eficiência, simetria, monotonia, linearidade) e fornecem importância local (por previsão) e global (por agregação). Em aplicações geoespaciais, SHAP maps — mapas onde cada pixel recebe o valor de Shapley de cada feature — revelam a contribuição espacialmente variável de cada variável preditora, conectando previsões do modelo com mecanismos geomorfológicos ou hidrológicos interpretativos.

O LIME (Local Interpretable Model-agnostic Explanations), proposto por Ribeiro et al. (2016), gera explicações locais ajustando um modelo linear interpretável na vizinhança de cada instância de interesse, ponderando as observações vizinhas por distância. LIME é computacionalmente mais barato que SHAP mas carece de suas propriedades axiomáticas.

### ! XAI como requisito ético

Em contextos de gestão ambiental onde decisões baseadas em modelos afetam comunidades (por exemplo, zoneamento de risco de inundação, priorização de áreas para restauração, definição de zonas de exclusão mineral), a explicabilidade não é um luxo técnico, mas um requisito ético e de governança. Modelos cujas decisões não podem ser explicadas não podem ser contestados, auditados ou melhorados com base em conhecimento local, violando princípios de transparência e participação pública na gestão ambiental.

## 12.5. Princípios FAIR e reprodutibilidade

Os princípios FAIR (Findable, Accessible, Interoperable, Reusable), formalizados por Wilkinson et al. (2016), estabelecem diretrizes para a gestão de dados científicos que maximizam sua utilidade e longevidade. Dados Findable possuem identificadores persistentes (DOI) e metadados ricos. Dados Accessible estão disponíveis por protocolos abertos e padronizados (HTTP, OGC WMS/WFS). Dados Interoperable utilizam vocabulários controlados e formatos padronizados (GeoTIFF, GeoPackage, CF-NetCDF). Dados Reusable possuem licenças claras (Creative Commons, ODbL) e proveniência documentada.

Nas geociências brasileiras, a adesão aos princípios FAIR é facilitada por repositórios institucionais (PANGAEA, Zenodo, Figshare) e por infraestruturas de dados espaciais (INDE/IBGE, IDE-Sisema, MapBiomias). A publicação de dados e código em repositórios abertos (GitHub,

GitLab) com documentação adequada e licenças permissivas constitui prática essencial para a reprodutibilidade da pesquisa ambiental.

## 12.6. Green AI e eficiência computacional

O conceito de Green AI, proposto por Schwartz et al. (2020), questiona a tendência de buscar melhorias incrementais de desempenho por aumento exponencial de computação (Red AI), advogando por uma métrica de eficiência que considere o resultado obtido por unidade de custo computacional (resultado/kWh). Em geociências ambientais, essa perspectiva é particularmente relevante porque muitos dados de monitoramento no semiárido e em regiões remotas são escassos (small data), invalidando a premissa de que mais dados e modelos maiores sempre produzem melhores resultados.

Estratégias de Green AI para geociências incluem transfer learning (pré-treinar em grandes conjuntos de dados globais e ajustar com poucos dados locais), modelos eficientes em parâmetros (GAM, XGBoost em vez de deep learning quando o ganho em acurácia é marginal), e otimização de hiperparâmetros por métodos bayesianos (Optuna, Hyperopt) em vez de grid search exaustivo. A meta é obter modelos com desempenho adequado à decisão de gestão (não necessariamente o máximo teórico) com o mínimo de custo computacional e ambiental.

A Tabela 12.2 apresenta diretrizes práticas.

Tabela 12.2.: Princípios de Green AI aplicados a geociências ambientais.

Princípio	Prática recomendada	Indicador
Eficiência	Preferir modelos simples quando $\Delta AUC < 0,02$	Resultado/kWh
Transfer learning	Pré-treinar em dados globais, fine-tune local	Redução de amostras necessárias
Parcimônia	Feature selection rigorosa antes de ML	Nº features vs. desempenho
Reprodutibilidade	Código, dados e ambiente em repositório	FAIR compliance
Transparência	XAI obrigatória em aplicações de gestão	SHAP/LIME reportados

## 12.7. Human-in-the-loop e IA responsável

O paradigma human-in-the-loop reconhece que, especialmente em contextos de dados escassos e alta consequência decisória (gestão ambiental), a supervisão humana contínua do pipeline de IA é essencial. Especialistas do domínio (hidrólogos, pedólogos, geomorfólogos) revisam previsões do modelo, identificam artefatos, validam padrões emergentes contra o conhecimento de campo e fornecem rótulos corrigidos que alimentam ciclos de aprendizado ativo (active learning), onde o modelo solicita a anotação de instâncias que maximizam a redução de incerteza.

A incorporação de incerteza epistêmica (do modelo) e aleatória (dos dados) nas previsões de IA é um requisito para decisões de gestão ambiental informadas por risco. Métodos como ensembles bayesianos, Monte Carlo dropout e previsão conformal fornecem intervalos de confiança pixelwise em mapas preditivos, permitindo que gestores identifiquem regiões onde a previsão é confiável e regiões onde dados adicionais são necessários antes de decidir. A comunicação honesta dessas

incertezas aos tomadores de decisão é um imperativo ético e científico que distingue a aplicação responsável de IA da tecnosolucionismo acrítico.

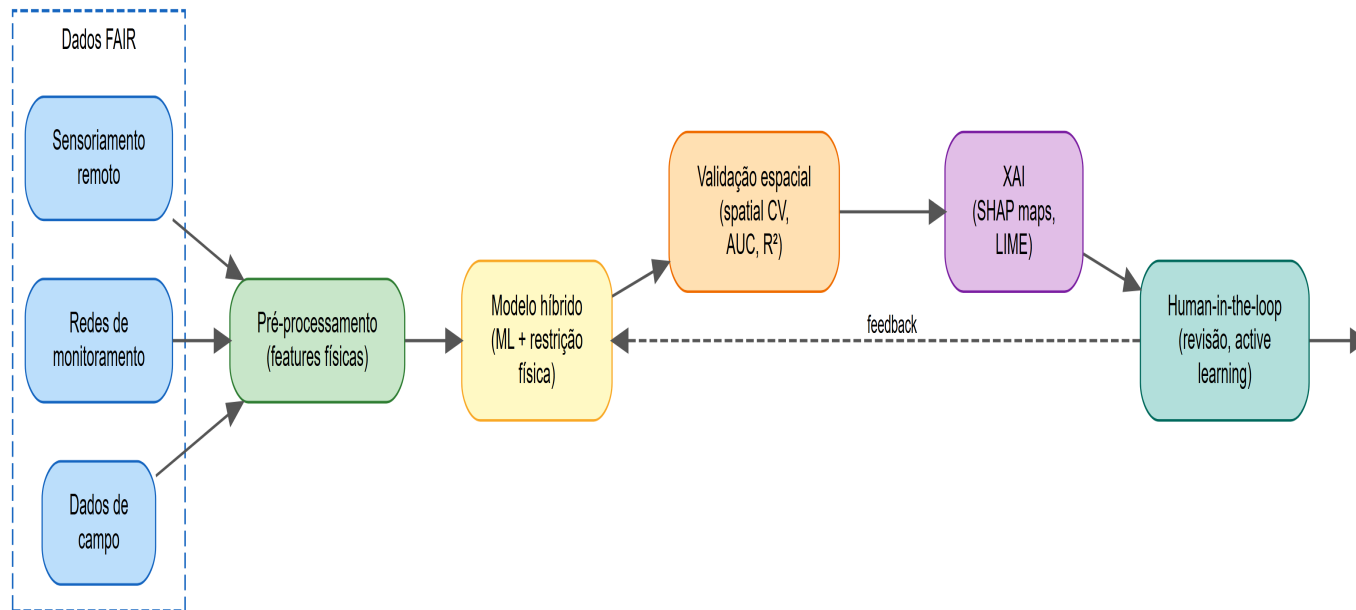


Figura 12.1.: Pipeline de IA responsável para geociências ambientais, integrando princípios FAIR, XAI, validação espacial e supervisão humana.

## Referências

- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of Hydrology*, 212–213, 198–212. [https://doi.org/10.1016/S0022-1694\(98\)00253-4](https://doi.org/10.1016/S0022-1694(98)00253-4)
- Bonham-Carter, G. F. (1994). *Geographic Information Systems for Geoscientists: Modelling with GIS*. Pergamon.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brimhall, G. H., Chadwick, O. A., Lewis, C. J., Compston, W., Williams, I. S., Danti, K. J., Dietrich, W. E., Power, M. E., Hendricks, D., & Bratt, J. (1992). Deformational mass transport and invasive processes in soil evolution. *Science*, 255(5045), 695–702. <https://doi.org/10.1126/science.255.5045.695>
- Charney, J. G. (1975). Dynamics of deserts and drought in the Sahel. *Quarterly Journal of the Royal Meteorological Society*, 101(428), 193–202. <https://doi.org/10.1002/qj.49710142802>
- Congalton, R. G., & Green, K. (2019). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices* (3<sup>o</sup> ed). CRC Press.
- Goldich, S. S. (1938). A Study in Rock-Weathering. *The Journal of Geology*, 46(1), 17–58. <https://doi.org/10.1086/624619>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Hey, T., Tansley, S., & Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research.
- Jenny, H. (1941). Factors of Soil Formation: A System of Quantitative Pedology. *McGraw-Hill*.
- Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., & Kumar, V. (2017). Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. *IEEE Transactions on Knowledge and Data Engineering*, 29(10), 2318–2331. <https://doi.org/10.1109/TKDE.2017.2720168>
- Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12), 2897–2910. <https://doi.org/10.1016/j.rse.2010.07.008>
- Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, 15(11), 91–100. [https://doi.org/10.1016/0273-1177\(95\)00079-T](https://doi.org/10.1016/0273-1177(95)00079-T)
- Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. F. H. (1993). The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44, 145–163. [https://doi.org/10.1016/0034-4257\(93\)90013-N](https://doi.org/10.1016/0034-4257(93)90013-N)
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, 179–184.

- Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. <https://doi.org/10.1016/j.rse.2014.02.015>
- Pontius Jr, 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. <https://doi.org/10.1080/01431161.2011.552923>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- Richards, L. A. (1954). *Diagnosis and Improvement of Saline and Alkali Soils*. USDA.
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, *1*, 309–317.
- Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. McGraw-Hill.
- Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54–63. <https://doi.org/10.1145/3381831>
- Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. *Eos, Transactions American Geophysical Union*, *38*(6), 913–920. <https://doi.org/10.1029/TR038i006p00913>
- Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, *33*(2), 309–319. <https://doi.org/10.1029/96WR03137>
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*(sup1), 234–240. <https://doi.org/10.2307/143141>
- Tomlin, C. D. (1990). *Geographic Information Systems and Cartographic Modeling*. Prentice Hall.
- Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, *114*(1), 106–115. <https://doi.org/10.1016/j.rse.2009.08.014>
- Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Silva Santos, L. B. da, Bourne, P. E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wyborn, L. A. I., Heinrich, C. A., & Jaques, A. L. (1994). Australian Proterozoic mineral systems: essential ingredients and mappable criteria. Em *The AusIMM Annual Conference* (p. 109–115).