

Revista Brasileira de Cartografia (2017), N° 69/3: 541-559  
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto  
ISSN: 1808-0936

## **INTEGRIDADE ECOSISTÊMICA A PARTIR DE DADOS DE SENSORIAMENTO REMOTO E REDES BAYESIANAS**

*Ecosystem Integrity based on Remote Sensing data and Bayesian Belief Network*

**Margareth Simões<sup>1,2</sup>, Rodrigo Ferraz<sup>1</sup> & Andrei Alves<sup>2</sup>**

**<sup>1</sup> Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA  
Embrapa Solos**

R. Jardim Botânico, 1024 - Jardim Botânico, Rio de Janeiro - RJ, 22460-000, Brasil  
{margareth.simoese, rodrigo.demonte}@embrapa.br

**<sup>2</sup> Universidade do Estado do Rio de Janeiro -UERJ**

**Faculdade de Engenharia – FEN / Depto. de Eng. de Sistemas e Computação – DESC / Programa de Pós  
Graduação em Meio Ambiente - PPGMA**

São Francisco Xavier 524, Bloco F sala 12.005. Maracanã - Rio de Janeiro - RJ - 20.550-900, Brasil  
{margareth.labex, kaageo}@gmail.com

*Recebido em 5 de Janeiro, 2016/ Aceito em 11 de Janeiro, 2017  
Received on January 5, 2016/ Accepted on January 11, 2017*

### **RESUMO**

A Integridade Ecosistêmica (IE) pode ser definida como sendo um estado de equilíbrio de um dado sistema natural que é capaz de se autorregular por meio de diversos processos funcionais. A biodiversidade, enquanto uma característica qualitativa dos ecossistemas, constitui um indicador de estado referente ao afastamento das condições iniciais de um dado sistema natural. O mapeamento da Integridade Ecosistêmica tem se apresentado como um importante indicador para avaliar a relação entre a perda de biodiversidade e os impactos nos serviços ecosistêmicos em florestas tropicais, uma vez que a IE representa a conexão da biodiversidade com a capacidade dos ecossistemas em manterem os processos de auto-organização. Assim, o objetivo do presente trabalho é apresentar uma abordagem metodológica desenvolvida para geração de um índice de Integridade Ecosistêmica em escala regional, para diferentes padrões fitofisionômicos de paisagem, utilizando um modelo probabilístico baseado em Redes Bayesianas (BBN), dados de Sensoriamento Remoto (SR) e GIS. O procedimento experimental foi realizado sobre região da Amazônia Legal Brasileira. Os resultados preliminares foram promissores, possibilitando quantificar as áreas da Amazônia Legal com maior ou menor Integridade Ecosistêmica. Utilizando-se a mesma rede bayesiana, com dados de satélite atualizados, torna-se possível monitorar a IE ao longo do tempo, podendo, até mesmo, servir para estabelecer um protocolo de monitoramento da mudança da IE de uma dada região de interesse.

**Palavras chaves:** Biodiversidade, Amazônia, MODIS, Indicadores Ambientais, Serviços Ecosistêmicos.

### **ABSTRACT**

Ecosystem Integrity (EI) may be defined as an equilibrium state of a given natural system which is able to regulate itself throughout many functional processes. Biodiversity, as a qualitative characteristic of ecosystems, is a status indicator regarding the exclusion of initial conditions of a given natural system. Mapping Ecosystem Integrity has emerged as an important indicator to assess the relationship between biodiversity loss and the impacts on ecosystem services in

tropical forests, once the EI represents a connection between biodiversity and the ability of ecosystems to maintain the self-organization process. The objective of this paper is to present a methodological approach developed for generating an Ecosystem Integrity index at a regional scale, for different phytophysiognomies patterns of landscapes, using a probabilistic model based on Bayesian Belief Networks (BBN), Remote Sensing data and GIS. The experimental procedure was performed in Brazil's Legal Amazon region. Preliminary results were promising, making it possible to quantify areas of the Amazon with high or low Ecosystem Integrity. Using the same bayesian network, with updated satellite data, becomes possible to monitor the EI over time, and may even serve to establish a monitoring protocol and monitor IE changes in a given region of interest.

**Keywords:** Ecosystem Integrity, Bayesian Belief Networks, Remote Sensing, Biodiversity, Amazon.

## 1. INTRODUÇÃO

A biodiversidade desempenha um importante papel no bem-estar humano, em particular, a partir da provisão dos serviços ecossistêmicos, que podem ser definidos como benefícios que a população obtém dos ecossistemas. Os serviços ecossistêmicos podem ser muitos: provisionamento (alimentos, água, madeira, etc.), regulação (regulação do clima, enchentes, doenças, etc.), apoio (como formação do solo, a fotossíntese e ciclagem de nutrientes), dentre outros.

A capacidade dos ecossistemas em fornecer bens e serviços depende da relação entre processos ecossistêmicos e dos componentes provenientes deles, em que os limites de uso sustentável são determinados por critérios ecológicos, tais como integridade, resiliência e resistência. O valor ecológico ou importância de um dado ecossistema é, portanto, determinado tanto pela integridade das funções de regulação e de habitat do ecossistema e por parâmetros ambientais, tais como a complexidade, diversidade e raridade (DE GROOT, 2005).

Um dos grandes desafios em estudos de biodiversidade é encontrar uma relação de medida entre a integridade da biodiversidade e os serviços ecossistêmicos num contexto sócio ecológico que considera, por exemplo, o impacto de várias ações relacionadas à atividade humana.

Como o conceito de biodiversidade é bastante diverso, abrangendo distintas visões que focalizam diferentes níveis de observação dos sistemas biológicos, diferentes indicadores podem ser usados para se caracterizar e/ou quantificar a biodiversidade de um dado sistema natural. Entretanto, em termos práticos, a biodiversidade constitui, enquanto uma característica qualitativa dos ecossistemas, um

indicador de estado referente ao afastamento das condições primiciais de um dado sistema natural. Assim, o conceito de biodiversidade pode ser aplicado para orientar a definição de indicadores que caracterizem o estado de um dado sistema natural no tocante à sua capacidade de autorregulação ou manutenção de processos ecossistêmicos (SIMÕES *et al.*, 2014).

A Integridade Ecosistêmica (IE) pode ser definida como sendo um estado de equilíbrio de um dado sistema natural que é capaz de se autorregular por meio de diversos processos funcionais. Desta forma, a IE constitui uma propriedade latente dos ecossistemas que não é diretamente observável (EQUIHUA *et al.*, 2013).

Diversos indicadores podem ser utilizados para quantificar a IE, podendo ser variáveis relacionadas à estrutura e/ou a funcionalidade dos ecossistemas, como por exemplo: (i) estrutura da vegetação (características dendrométricas); (ii) diversidade biótica (número de espécies ou de grupos funcionais); (iii) presença de espécies chaves ou indicadoras de estágios clímax; (iv) entrada de energia solar (balanço de radiação, produção de biomassa, sequestro de carbono); (v) saída de matéria e energia (respiração, transpiração, evapotranspiração); (vi) fluxos de nutrientes (mineralização), etc. Porém, devido à complexidade dos ecossistemas e à compreensão ainda incompleta dos processos ecossistêmicos em respostas as perturbações antrópicas, deve-se considerar que a causalidade das relações entre os variáveis escolhidas como indicadores e o estado de integridade ecossistêmica nem sempre é tão evidente, guardando certo grau de incerteza.

Uma forma de lidar com as incertezas nas modelagens de fenômenos latentes, ou seja, que não são diretamente observáveis, é a utilização

de modelos probabilísticos ou estocásticos. Estes modelos permitem trabalhar com fenômenos cujo estado é indeterminado, e tem origem em eventos aleatórios, como a abordagem das Redes Bayesianas (SIMÕES *et al.*, 2015).

Em trabalhos anteriores relacionados à presente pesquisa (SIMÕES *et al.*, 2015; VERWEIJ *et al.*, 2014, SIMÕES *et al.*, 2014) foram comparadas diferentes abordagens para o mapeamento da IE na área da Amazônia Legal Brasileira, dentre elas: Rede Bayesiana (Bayesian Belief Network - BBN) com algoritmo de Máxima Probabilidade; Rede Bayesiana (BBN) com algoritmo de Distribuição de Probabilidade (data-driven) e modelo baseado em regras de conhecimento (knowledge driven). Apesar de os resultados apresentarem semelhança de distribuição de área por categoria de IE nas três abordagens, houve uma diferença entre a distribuição do padrão espacial. Na abordagem baseada em regra de conhecimento observa-se uma tendência de se agrupar as categorias (clusters), enquanto na abordagem da modelagem por distribuição de probabilidades apresenta uma imagem mais dispersa. Os resultados mostraram que a utilização do método de BBNs tem se apresentado como uma boa alternativa de metodologia de suporte a decisão no âmbito da gestão dos recursos naturais. A IE e as suas relações com os diversos serviços ecossistêmicos, como por exemplo, os fatores de mitigação de mudanças climáticas, podem fornecer informações importantes sobre os diferentes ecossistemas, balizando a proposição de políticas públicas e alternativas de gestão ambiental.

Outro importante fator de análise em trabalhos relacionados à IE é o fator da escala geográfica: local, regional ou global. Muitos dos trabalhos cujo objetivo é medir a integridade de determinado ambiente são locais (OLTHOF *et al.*, 2006; TIERNEY *et al.*, 2009; PEREIRA *et al.*, 2012; ALBERTI, 2010), possibilitando a realização de medidas físicas, químicas e biológicas sobre amostras coletadas, permitindo avaliação do estado de integridade de um ecossistema local. Para este tipo de pesquisa, dois problemas fundamentais já estão, aparentemente, solucionados: a) os dados de entrada do modelo, que são as amostras físicas, químicas e biológicas do ambiente; e b) a avaliação das amostras,

que pode ser física, química ou biológica, quantitativa ou qualitativa.

No entanto, a aquisição de dados de entrada e a avaliação de resultados são problemáticas que aumentam proporcionalmente à escala de trabalho, sejam elas regionais ou globais. Neste sentido, os dados de Sensoriamento Remoto (SR) tem desempenhado importante papel para a elaboração de pesquisas em escala regional. Dados e produtos provenientes do satélite Terra, sensor MODIS, por exemplo, permitem obter informações importantes sobre o ambiente e serviços ecossistêmicos, e tem se mostrado suficientes para a formação de evidências para o mapeamento da IE.

Na literatura são encontrados muitos trabalhos que propõem metodologias para geração e uso de indicadores de biodiversidade a partir de dados de sensoriamento remoto e outras variáveis (PETROU & PETROU, 2011; EEA, 2010; CHAMBERS *et al.*, 2007; TURNER *et al.*, 2003); no entanto, são escassas as propostas metodológicas destinadas a fazer uso destes indicadores para geração de um índice de IE em escala regional.

Assim, entendendo a inovação e complexidade do tema, o objetivo do presente trabalho foi apresentar de modo detalhado uma abordagem metodológica desenvolvida para geração de um Índice de Integridade Ecosistêmica em escala regional, para diferentes padrões fitofisionômicos de paisagem, utilizando um modelo probabilístico baseado em Redes Bayesianas (BBN) e dados de sensoriamento remoto.

Neste trabalho, considerou-se a IE como um representante da biodiversidade. Ou seja, como sendo a capacidade de auto-organização dos ecossistemas biodiversos, de manutenção dos processos ecossistêmicos relacionados aos fluxos de matéria e de energia, essenciais para os ciclos biogeoquímicos terrestres, notadamente, o ciclo do carbono e da água. Integrando um modelo de Integridade Ecosistêmica (BBN) em um Sistema de Informações Geográficas (SIG), tornou-se possível mapear um grande número de entidades espaciais e realizar múltiplas interações para gerar diferentes mapas de Integridade Ecosistêmica da Amazônia Legal Brasileira.

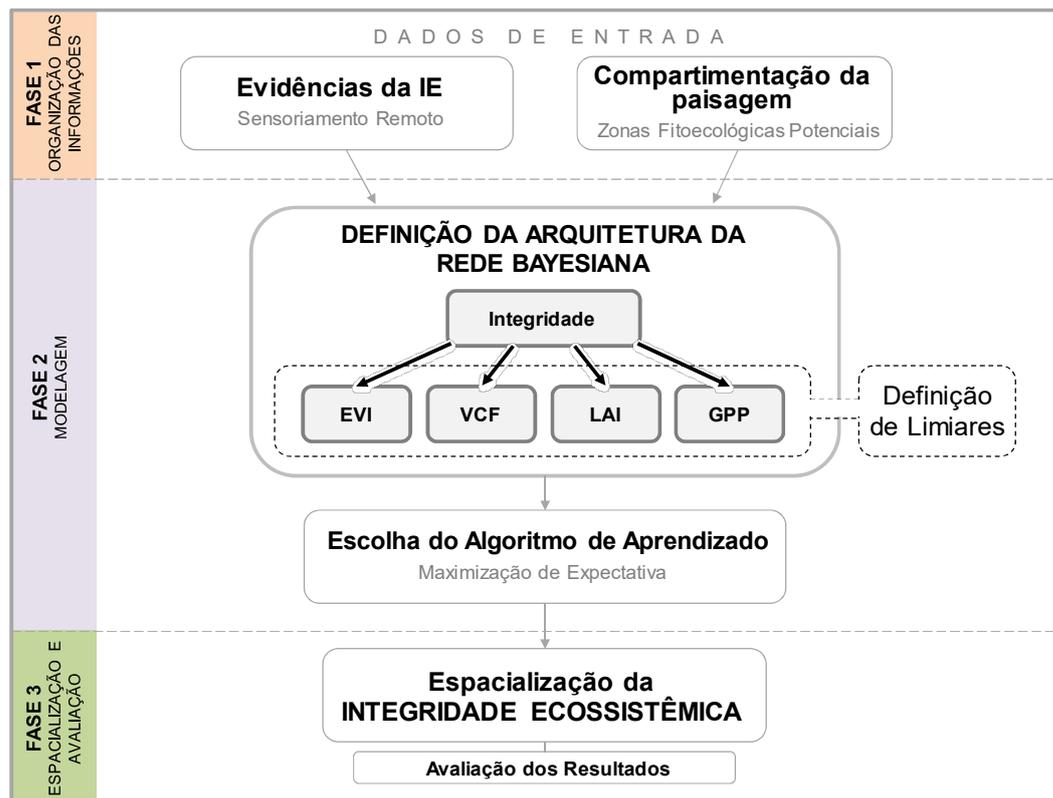


Fig. 1 – Metodologia para geração do Índice de Integridade Ecosistêmica.

## 2. METODOLOGIA PARA GERAÇÃO DO ÍNDICE DE INTEGRIDADE ECOSSISTÊMICA

A abordagem metodológica criada consiste na geração de um modelo espacial baseado na distribuição de probabilidades de variáveis consideradas como evidências da IE, utilizando a metodologia das Redes Bayesianas (Bayesiana Belief Network - BBN) e utilizando o algoritmo de aprendizado Maximização de Expectativa (EM - Expectation Maximization).

A Figura 1 apresenta de modo simplificado a metodologia desenvolvida para a geração do índice de Integridade Ecosistêmica.

Na “FASE 1 – ORGANIZAÇÃO DAS INFORMAÇÕES” são coletadas, sistematizadas e processadas todas as informações que serão utilizadas como dados de entrada do modelo. Ainda, é definida e organizada toda infraestrutura computacional necessária para a aplicação da abordagem metodológica, como os softwares utilizados no processo de modelagem da rede, na etapa de espacialização e na validação dos resultados.

Na “FASE 2 – MODELAGEM” é definida a topologia da rede, ou seja, é criado um modelo padrão da rede que possa ser aplicado em padrões distintos de paisagem. Ainda nesta etapa são definidos, para cada tipo de paisagem, limiares que serão utilizados no processo de discretização das evidências da rede bayesiana, com o objetivo de identificar quais elementos da paisagem possuem integridade mais alta ou mais baixa. Em seguida, estes dados são submetidos a um algoritmo de aprendizado que calcula, dada uma condicionante de entrada, a probabilidade de a IE ser alta ou baixa para cada elemento (pixel).

Na “FASE 3 – ESPACIALIZAÇÃO E AVALIAÇÃO” é realizada a espacialização da IE, onde os dados gerados na saída do processo de modelagem, representados por um valor normalizado (valor entre 0 e 1) para cada elemento (pixel), são especializados para se obter um mapa. Posteriormente, os dados são avaliados a partir de dados de referência e procedimentos estatísticos, necessários para a validação da abordagem metodológica.

Alguns fatores essenciais para o sucesso desta abordagem serão detalhados nas subseções seguintes, são eles:

- Escolha das evidências da Integridade Ecológica;
- Compartimentação da paisagem;
- Método das Redes Bayesianas;
- Algoritmo Maximização de Expectativa (EM);
- Modelagem da rede;
- Definição dos limites das evidências que compõem a rede.

## 2.1 Evidências da Integridade Ecológica

As evidências da IE são os dados de entrada do modelo e que serão utilizados pelo algoritmo de aprendizado para geração do índice de integridade. A definição das evidências é um fator fundamental no processo de geração da IE e está diretamente relacionada com o objetivo do trabalho.

As evidências podem ter origem em dados de sensoriamento remoto, de mapeamentos sistemáticos, coletados em campo, dentre outros. No entanto, é importante ressaltar a necessidade da obtenção de dados de fontes confiáveis, com alto padrão de qualidade, que abranjam grande parte da área analisada e que possibilitem a continuidade do mapeamento em tempos futuros. Ainda, as evidências precisam ser compatíveis entre si no tempo (dia, mês, ou ano), em escala e, obrigatoriamente, ser de significativa relevância para a geração da IE.

A proposta desta abordagem é que, para estudos em regiões que carecem de dados de mapeamento sistemático atualizado ou de dados coletados em campo, sejam utilizados dados de sensoriamento remoto.

No presente trabalho foram utilizados como evidências dados do sensor MODIS - Moderate Resolution Imaging Spectroradiometer a bordo do satélite Terra. Produtos advindos do sensor MODIS tem se mostrado suficientes para a formação de evidências para o mapeamento da IE.

O sensor MODIS é um instrumento a bordo dos satélites Terra e Aqua, ambos concebidos e gerenciados pelo EOS – Earth Observing System da NASA - National Aeronautics and Space Administration. Tanto o sensor Terra/MODIS quando Aqua/MODIS estão visitando toda a superfície da Terra de 1 a 2 dias, adquirindo dados em 36 bandas espectrais. Estes dados

têm auxiliado pesquisadores na compreensão da dinâmica global e em processos que ocorrem na Terra, nos oceanos e na atmosfera mais baixa (NASA/USGS, 2014).

Os seguintes produtos foram definidos como evidências advindas de dados de sensoriamento remoto:

- EVI - Enhanced Vegetation Index. O “Índice de Vegetação Realçado” permite destacar as áreas com maior quantidade de vegetação em relação a área de menor quantidade de vegetação. Em comparação com o NDVI, o EVI minimiza as variações de fundo do dossel e mantém a sensibilidade sobre as condições de vegetação densa. Também, usa a faixa azul do espectro para eliminar a contaminação residual atmosfera causada pela fumaça e sub-pixels fina nuvens (NASA/USGS, 2014).
- LAI - Leaf Area Index. O “Índice de Área Foliar” é um produto gerado a partir de dados do sensor MODIS (a partir do ano 2000) pelo Center for Global Change Data Processing and Analysis of Beijing Normal University. Basicamente, o índice representa a unidade de área de folhas por unidade de área de terreno (LIANG & XIAO, 2012).
- VCF – Vegetation Continuous Fields: percent tree cover. O índice “Campos Contínuos de Vegetação: percentual de cobertura de árvores” é um produto derivado de dados MODIS e representa, por pixel, o percentual de área coberta por vegetação (DIMICELI *et al.*, 2011);
- GPP - Gross Primary Productivity. O índice “Produtividade Primária Bruta” corresponde à quantidade de matéria orgânica produzida pelo ecossistema em determinada área, durante determinado intervalo de tempo (NASA/USGS, 2014).

A Figura 2 apresenta um exemplo da diferença de resposta espectral dos produtos de sensoriamento remoto utilizados como evidências da IE para a área da Floresta Nacional do Tapajós-PA. As imagens apresentam resolução espacial de 1km e datam de janeiro de 2010.

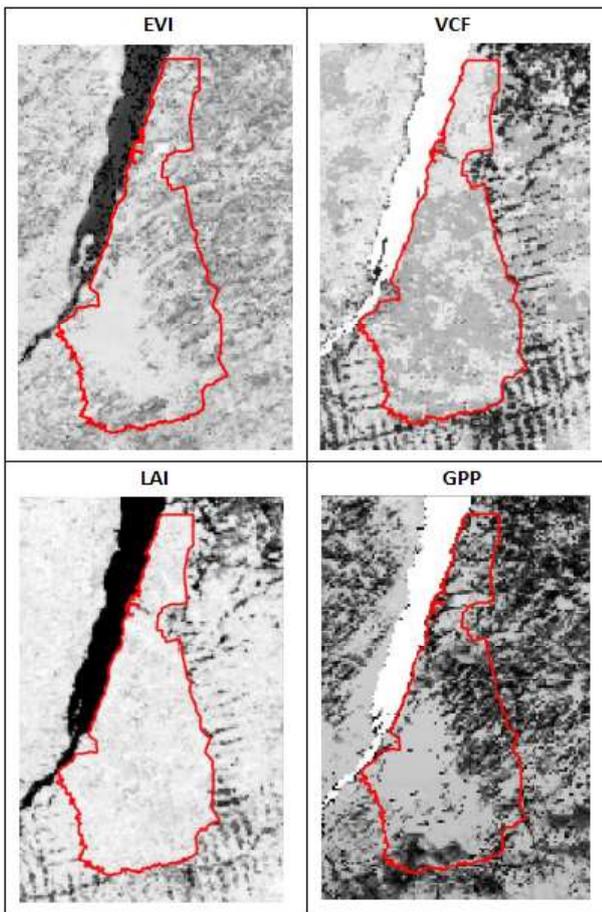


Fig. 2 – Exemplo da diferença de resposta espectral para as evidências, na Floresta Nacional do Tapajós-PA.

## 2.2 Compartimentação (ou fragmentação) da paisagem

Além das evidências advindas de produtos de sensoriamento remoto, foram utilizadas técnicas de compartimentação (ou fragmentação) da paisagem. O objetivo de compartimentar a paisagem é inserir limites espaciais no modelo probabilístico para separar áreas com características fitofisionômicas distintas. Com isso, torna-se possível modelar uma rede para cada tipo de fitofisionomia, possibilitando ajustar os valores das evidências de acordo com a característica de cada região.

Nesta abordagem, como fragmentador da paisagem foram utilizados limites de Zonas Fitoecológicas Potenciais (ZFP). Diferentemente das Zonas Fitoecológicas Atuais (presente) as Zonas Fitoecológicas Potenciais remetem ao passado com um ambiente original em que todas as fitofisionomias apresentavam-se preservadas e íntegras. Os dados de Zonas Fitoecológicas

Potenciais utilizados no presente trabalho foram gerados por conhecimento especialista a partir de uma metodologia de agregação de classes fitofisionômicas, utilizando como base o atributo “Vegetação Pretérita – Vegetação existente antes de a área ser antropizada” dos dados vetoriais de vegetação do IBGE, do ano de 2008.

Dependendo da região em estudo, outros delimitadores de áreas, que permitam identificar diferentes fitofisionomias ou limites de vegetação potenciais, podem ser utilizados. Como exemplo, pode ser citada a metodologia das Zonas de Vida de Holdridge (HOLDRIDGE, 1977).

No contexto da escolha das evidências da IE, considerando as evidências citadas, parte-se do pressuposto que, para determinada fitofisionomia, áreas com índice de vegetação elevado, maior índice de área foliar, maior percentual de área coberta por vegetação e com maior quantidade de matéria orgânica produzida, apresentam-se como regiões onde a probabilidade de o ecossistema local se manter íntegro é maior do que em regiões onde os valores das evidências são baixos. Vale ressaltar que, o que define se o valor de uma evidência é alto ou baixo (no contexto da IE) é um limiar de corte previamente analisado e estabelecido por um especialista. Portanto, devido às diferenças fitofisionômicas em cada região, fragmentar a paisagem por si só não garante a qualidade do resultado. A modelagem da rede e a etapa de definição dos limiares são de extrema importância no processo.

## 2.3. Método das Redes Bayesianas (Bayesiana Belief Network - BBN)

O método bayesiano foi escolhido no processo de modelagem da IE, em detrimento a outros métodos, porque está fundamentado no conceito de probabilidade a priori e a posteriori e consiste em determinar a probabilidade de ocorrer um determinado evento ou fenômeno de interesse dado a presença de um condicionante ou evidência. Transformando uma série de dados determinísticos em probabilidades o método faz uso da inserção de conjuntos para gerar funções de probabilidades condicionais, baseados nos princípios introduzidos por Bayes. A modelagem matemática do teorema de Bayes liga a inferência racional (probabilidade a posteriori) a subjetividade (probabilidade a

priori) e a experiência empírica (probabilidades condicionais). As Redes Bayesianas são modelos probabilísticos que representam relacionamentos correlativos e causais entre variáveis. Ou seja, descreve a distribuição de probabilidades entre as variáveis tidas como evidências e o fenômeno de interesse. Maiores informações sobre o método bayesiano pode ser obtido em Lindley (1972) e em Lauria e Duchessi (2006).

### 2.4 Maximização de Expectativa (EM)

Pode-se dizer que o algoritmo de Maximização de Expectativas é uma generalização do algoritmo de estimativa de Máxima Verossimilhança para o caso dos dados incompletos. O algoritmo EM busca encontrar os parâmetros  $\hat{\theta}$  que maximizam a probabilidade logarítmica  $\log P(x; \theta)$  dos dados observados. Basicamente, o problema de otimização abordado pelo EM é mais complexo que a otimização realizada pelo algoritmo de Máxima Verossimilhança (DO & BATZOGLOU, 2008).

Para o caso onde os dados estão completos, a função objetivo  $\log P(x; z; \theta)$  tem um único ótimo global, e que na maioria das vezes pode ser encontrado de forma fechada. Em contraste, no caso dos dados incompletos, a função  $\log P(x; \theta)$  tem múltiplos máximos locais e nenhuma solução de forma fechada. Para lidar com isso, o EM reduz a difícil tarefa de otimizar  $\log P(x; \theta)$  em uma sequência de subproblemas de otimização mais simples, cujas funções objetivo tem um único máximo global que, dessa vez, podem ser encontrados de forma fechada. Esses subproblemas são escolhidos de uma maneira que garante que suas soluções correspondentes  $\hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \dots$  e irão convergir para um ótimo local de  $\log P(x; \theta)$  (DO & BATZOGLOU, 2008).

Mais especificamente, o algoritmo Maximização de Expectativa alterna entre duas fases, E (Expectativa) e M (Maximização). Durante a fase E, o EM escolhe a função  $g_t$  que limita  $\log P(x; \theta)$  inferiormente totalmente, e para a qual  $g_t(\hat{\theta}^{(t)}) = \log P(x; \hat{\theta}^{(t)})$ . Durante a fase M, o algoritmo passa para o novo conjunto de parâmetros  $\hat{\theta}^{(t+1)}$  que maximiza  $g_t$ . A medida que o limiar inferior  $g_t$  se aproxima da função objetivo em  $\hat{\theta}^{(t)}$ , temos:

$$\log P(x; \hat{\theta}^{(t)}) = g_t(\hat{\theta}^{(t)}) \leq g_t(\hat{\theta}^{(t+1)}) = \log P(x; \hat{\theta}^{(t+1)}) \quad (1)$$

Desta maneira, a função objetivo aumenta monotonicamente a cada iteração do EM.

A Figura 3 apresenta um exemplo do funcionamento do algoritmo EM. Na etapa 1, o EM começa com uma estimativa inicial dos parâmetros. Na etapa 2, fase E, uma distribuição de probabilidade sobre possíveis conclusões é calculada usando os parâmetros atuais. As contagens mostradas na tabela são os números esperados de cabeças e caudas de acordo com esta distribuição. Na etapa 3, fase M, novos parâmetros são determinados usando as conclusões atuais. Na etapa 4, depois de várias repetições das fases E e M, o algoritmo converge.

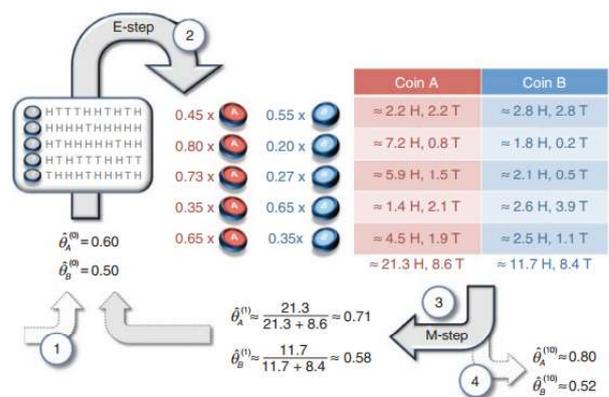


Fig. 3 – Exemplo do funcionamento do algoritmo EM. Fonte: Adaptado de Do e Batzoglou (2008).

### 2.5 Modelagem da Rede Bayesiana

As redes bayesianas são representadas por grafos que denotam uma relação de causa e efeito entre variáveis aleatórias interdependentes em modelos probabilísticos.

Para a modelagem da rede foram realizados diferentes testes e análises, com diferentes dados de evidências e modelos de rede. A rede utilizada nesta abordagem foi estabelecida a partir da concepção de um modelo conceitual - Integridade Ecológica e Evidências - que integrou o conhecimento de diversos especialistas (ecologistas, biólogos, geógrafos, engenheiros florestais e especialistas em GIS e Sensoriamento Remoto, etc.) da Embrapa Solos e do projeto ROBIN - Role Of Biodiversity In climate change mitigatioN (<http://robinproject.info/>). Em um processo interativo, por meio de videoconferências, os especialistas foram ajustando e aprimorando o modelo conceitual da rede, considerando os pressupostos teóricos e a

disponibilidade de dados.

A partir do aprimoramento do modelo, percebeu-se que para obter a melhor resposta às diferentes características fitofisionômicas da região, seria necessário rodar o modelo de modo independente para cada Zona Fitoecológica Potencial, permitindo definir limiares de modo independente, para cada tipo de evidência.

A Figura 4 apresenta um exemplo da rede bayesiana criada, de modo independente, para cada Zona Fitoecológica Potencial da Amazônia Legal Brasileira. A estrutura da rede foi elaborada de modo que o nó “Integridade” representasse uma relação de causa direta das evidências, que são os atributos ecossistêmicos (EVI, VCF, LAI, e GPP). Assim, a alteração do estado de cada um dos atributos ecossistêmicos implica diretamente na IE.

A escolha de utilização de uma Rede Bayesiana para cada tipo de Zona Fitoecológica Potencial permitiu estressar e adequar os limiares de corte das evidências da rede. Para esta etapa foi criada uma metodologia específica, detalhada na seção seguinte.

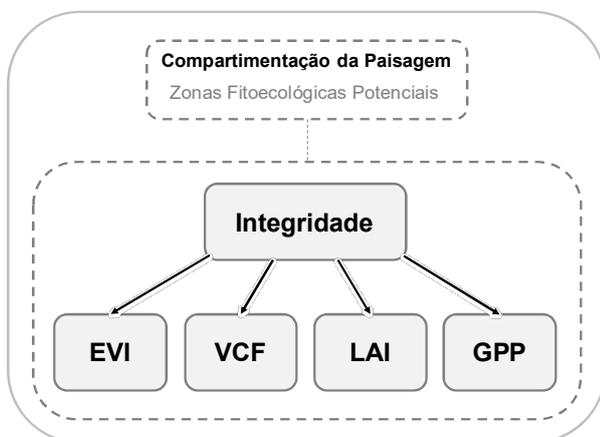


Fig. 4 – Exemplo de Rede Bayesiana criada para cada Zona Fitoecológica Potencial.

## 2.5 Metodologia para definição dos limiares das evidências

A metodologia para definição dos limiares foi baseada em análise estatística e análise por fotointerpretação sobre os atributos (pixel) das evidências, em áreas de Unidades de Conservação (UC), para cada tipo de unidade de paisagem (Zonas Fitoecológicas Potenciais).

A etapa de ajuste dos limiares das evidências de uma rede bayesiana é etapa fundamental da

abordagem metodológica, pois permite definir qual será o valor de corte para considerar se o atributo analisado (pixel da imagem) tem maior ou menor grau de integridade.

Como os dados de entrada do modelo denotam uma distribuição espacial contínua, um limiar de corte é utilizado para discretizar a informação em classes. Nesta abordagem, os dados foram discretizados em apenas duas classes: Baixa Integridade e Alta Integridade. Desta forma, um conjunto de amostras sempre estará distribuído dentro de uma destas classes. Por exemplo, para uma fitofisionomia composta predominantemente por floresta ombrófila, dada uma evidência “EVI” (em que os valores variam entre 0 e 1), o limiar de corte definido pelo especialista poderá ser o valor 0,7. Isso significa que, no modelo, os valores abaixo deste limiar entrarão na classe Baixa Integridade, enquanto os valores acima deste limiar entrarão na classe Alta Integridade.

Como o método utilizado está fundamentado no conceito de probabilidade a priori e posteriori, o modelo é consideravelmente sensível aos ajustes dos limiares. Obviamente, uma definição equivocada de um limiar influenciará diretamente no resultado final do processo.

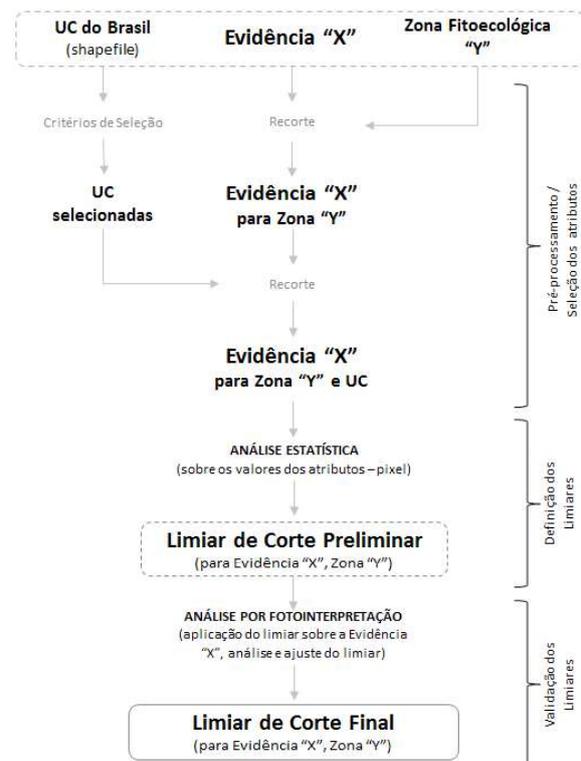


Fig. 5 – Metodologia para definição do limiar de corte das evidências.

O ajuste dos limiares foi realizado de modo independente, para cada tipo de evidência, para cada Zona Fitoecológica Potencial. Ainda, como existe o pressuposto de que as informações de entrada do modelo são evidências de atributos mais ou menos íntegros, as análises para definição dos limiares foram realizadas principalmente em áreas de Unidades de Conservação (UC), pois são áreas que, em teoria, deveriam apresentar maior índice de IE, por estarem sujeitas à menor interferência antrópica e por estarem sob uma legislação que “garante” a preservação de suas características originais.

A Figura 5 apresenta um exemplo das etapas do procedimento metodológico adotado para definição do limiar de corte para cada evidência.

Para seleção das UC foi utilizado o arquivo vetorial (shapefile) “Unidades de Conservação do Brasil”, do Ministério do Meio Ambiente, ano 2014, escala 1:5.000 a 1:100.000. Com o objetivo de utilizar as UC que apresentassem maior grau de integridade (dada sua categoria de uso), foram selecionadas unidades dentro dos seguintes critérios:

- Tipos de unidades com menor influência antrópica;
- Data de criação anterior a 2010 (ano referência das evidências);
- Mantidos os limites com padrão de qualidade “Correto” e “Aproximado” e descartados os limites com padrão de qualidade “Esquemático”.
- Após a aplicação dos critérios de seleção, foram utilizadas UC do tipo:
- Unidade de Proteção Integral: Estação Ecológica e Reserva Biológica;
- Unidades de Uso Sustentável: Floresta Nacional e Área de Proteção Ambiental.

Conforme apresenta a Figura 5, para uma determinada evidência foram selecionados apenas os atributos que se localizavam dentro do limite de uma única Zona Fitoecológica Potencial. Em seguida, estes atributos passam por uma segunda etapa de seleção, baseada nos limites das UC.

Após esta etapa inicia-se a análise estatística. A partir de uma análise teórica da resposta espectral da evidência em estudo, foi estudado o padrão de distribuição estatística dos

dados com base nas seguintes informações: valor mínimo, valor máximo, 1º quartil, 3º quartil, mediana (1º e 3º quartil) e moda. Baseado nos dados estatísticos e no estudo da evidência foi definido um primeiro limiar de corte, que serviu como base inicial para o processo de validação.

Posteriormente, iniciou-se uma etapa de validação dos limiares de corte, onde o limiar preliminar foi aplicado sobre a imagem de evidência e, utilizando técnicas de fotointerpretação, o valor foi ajustado, chegando ao limiar de corte final. Diferentemente da etapa de análise estatística, nesta etapa de análise por fotointerpretação, foi analisado o limiar de corte não apenas para a área das unidades de conservação, mas para toda a área da Zona Fitoecológica Potencial, com o objetivo de identificar se existiam discrepâncias entre os valores definidos como indicativos de Alta Integridade e as áreas para as quais este limiar seria aplicado.

É importante ressaltar que, nesta etapa da abordagem metodológica, a utilização de um limiar de discretização tem o objetivo principal de indicar, pelo especialista, os elementos mais íntegros, dada à evidência em análise. Portanto, a discretização em duas classes nesta etapa não deve ser confundida com a divisão de classes criada no mapa de integridade e utilizada para avaliação dos resultados.

## **2.6 Metodologia para avaliação dos resultados**

A metodologia para avaliar os resultados foi baseada na comparação da classificação da IE com uma base de dados de referência (gabarito).

A partir da elaboração de uma matriz de confusão foi possível calcular, para cada elemento (pixel), a exatidão global e as taxas de classificação por classe (COHEN, 1960), os erros de omissão, comissão e erros totais da classificação em análise (MATHER, 2005) e o índice Kappa.

O coeficiente de concordância Kappa informa o acerto da variação intra e inter-observador, sendo utilizado quando dois observadores classificam separadamente uma amostra de objetos empregando a mesma escala de classes. Normalmente, um dos observadores é a referência real ou verdade. (COHEN, 1960; TSO & MATHER, 2009).

### 3. ESTUDO DE CASO APLICADO À AMAZÔNIA LEGAL

Devido à importância global do ecossistema proveniente da floresta Amazônica e com o objetivo de realizar uma análise em escala regional, a abordagem metodológica foi aplicada para a área da Amazônia Legal Brasileira (AML).

#### 3.1 Área de estudo

A área de estudo engloba nove estados brasileiros: Acre, Amazonas, Amapá, Maranhão (parcialmente), Mato Grosso, Pará, Rondônia, Roraima e Tocantins. É importante ressaltar também, que a escolha da área de estudo levou em consideração às diferentes características fitofisionômicas apresentadas na região, permitindo avaliar a abordagem em diferentes regiões.

#### 3.2 Infraestrutura computacional

Para o desenvolvimento para pesquisa foram utilizados os softwares ArcGIS 10.3 (<http://www.esri.com/>) e QGIS 2.8.3 (<http://www.qgis.org/>) para o pré-processamento e análise dos dados geográficos. Para criar e rodar a rede BBN a partir do modelo conceitual estabelecido com base no conhecimento dos especialistas foi utilizado o software NETICA 5.1.8 (<https://www.norsys.com/>). Ainda, foi utilizado o software Minitab Statistical Software 17 (<http://www.minitab.com/>) para análise dos dados estatísticos dos limiares e avaliação dos resultados.

#### 3.3 Base de dados de entrada do modelo

Os dados de sensoriamento remoto utilizados como evidências de entrada do modelo foram:

- EVI – Índice de Vegetação Realçado (Enhanced Vegetation Index)  
Produto: MODIS – MOD13A3.005 (Monthly L3 Global 1km);  
Periodicidade: Mensal;  
Data: 01/01/2010;  
Resolução espacial: 1 km;  
Formato do arquivo: HDF;  
Subdataset: Subdataset 1 | Variable: “1\_km\_monthly\_EVI”;  
Fonte: NASA/USGS, 2014;  
Informações gerais: “<https://lpdaac.usgs.gov/>

[dataset\\_discovery/modis/modis\\_products\\_table/mod13a3](dataset_discovery/modis/modis_products_table/mod13a3)”.

- LAI – Índice de Área Foliar (Leaf Area Index)  
Produto: MODIS - MOD09A1;  
Periodicidade: 8 dias;  
Data: 01/01/2010 a 08/01/2010;  
Resolução espacial: 1 km;  
Formato do arquivo: HDF;  
Subdataset: Subdataset 0 | Variable: “LAI”;  
Fonte: LIANG & XIAO, 2012; NASA/USGS, 2014;  
Informações gerais: “<http://glcf.umd.edu/data/lai/>”.
  - VCF – Percentual de cobertura de árvores (Vegetation Continuous Fields - Percent tree cover)  
Produto: MODIS - MOD44B\_V5;  
Periodicidade: Anual;  
Data: 2010;  
Resolução espacial: 250m (reamostrado para 1km);  
Formato do arquivo: GEOTIF;  
Subdataset: TRE – Percent tree cover;  
Fonte: DIMICELI *et al.*, 2011; NASA/USGS, 2014;  
Informações gerais: “<http://glcf.umd.edu/data/vcf/>”.
  - GPP – Produtividade Primária Bruta (Gross Primary Productivity)  
Produto: MODIS - MOD17A2.005 (8-Day L4 Global 1km);  
Periodicidade: 8 dias;  
Data: 01/01/2010 a 08/01/2010;  
Resolução espacial: 1 km;  
Formato do arquivo: HDF;  
Subdataset: Subdataset 0 | Variable: “Gpp\_1km”;  
Fonte: NASA/USGS, 2014;  
Informações gerais: “[https://lpdaac.usgs.gov/dataset\\_discovery/modis/modis\\_products\\_table/mod17a2](https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod17a2)”.
- Os dados vetoriais (shapefiles) utilizados na abordagem foram:
- Zonas Fitoecológicas Potenciais  
Ano: 2008;  
Escala: 1:250.000  
Fonte primária: IBGE.  
Fonte secundária: Adaptado de PROBIO, 2008.
  - Unidades de Conservação do Brasil  
Ano: 2014;  
Escala: 1:5.000 a 1:100.000;  
Fonte: Ministério do Meio Ambiente.

- Limite da Amazônia Legal  
Ano: 2008;  
Escala: 1:250.000  
Fonte: IBGE.
- Unidades de Federação do Brasil (UF)  
Ano: 2008;  
Escala: 1:250.000;  
Fonte: IBGE.

As Zonas Fitoecológicas Potenciais foram utilizadas para delimitar áreas com características fitofisionômicas distintas. Os dados de Unidades de Conservação do Brasil foram utilizados na etapa de definição de limiares das evidências, conforme já mencionado em seção anterior. O Limite da Amazônia Legal foi utilizado como máscara para recortar os dados na imagem. A camada UF foi utilizada apenas como dado de referência política para análise.

A Figura 6 apresenta as etapas de pré-processamento dos dados geográficos desde sua aquisição até a elaboração da rede. Para os dados de SR, foram realizadas as seguintes atividades: compatibilização da Projeção/Datum; mosaicos (para dados que possuíam diversos quadrantes), reamostragem do tamanho dos pixels para resolução espacial de 1 km (algoritmo vizinho mais próximo) e recorte das imagens para o limite da Amazônia Legal. Para os dados vetoriais (shapefile), também foi necessário realizar o ajustes de Projeção/Datum, bem como a conversão dos dados para o formato raster e o recorte para o limite da AML. Após esta etapa, todos os dados passaram a ser exibidos em formato raster, com tamanhos de pixel e projeções cartográficas compatíveis.

Após a etapa de compatibilização das informações foi realizado um processamento para converter os dados da imagem para o formato tabular. O algoritmo utilizado extrai, a partir de uma determinada localização (centroide do pixel, coordenadas X, Y), os valores dos pixels de todas as camadas selecionadas e insere em uma tabela, onde cada linha representa um único pixel e cada coluna faz referência às camadas utilizadas no processo.

Uma vez criada a tabela com todas as informações, iniciou-se a etapa de remoção dos Outliers (valores atípicos, discrepantes), informações discrepantes que não poderiam ser utilizadas na análise estatística para definição

dos limiares nem da etapa de processamento das redes bayesianas. A remoção dos valores discrepantes incluiu a remoção dos dados correspondentes às classes “massa d’água”, “água” e/ou valores nulos e, também, à reclassificação de pixels classificados como “não classificados”, “cobertura urbana”, “áreas esparsas e de vegetação escassa”.

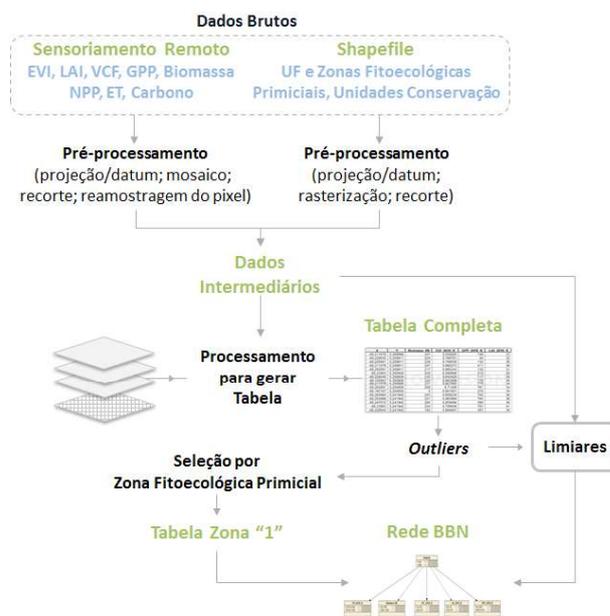


Fig. 6 – Etapas de pré-processamento dos dados geográficos utilizados no procedimento experimental.

A remoção dos valores atípicos representou 3,12% do total dos dados (pixels), passando de 5.075.097 para 4.916.821.

A etapa seguinte foi à realização de uma seleção dos dados da tabela por Zona Fitoecológica Potencial. O resultado foi a obtenção de uma tabela para cada ZFP, utilizada para definição dos limiares e como dados de entrada do modelo.

### 3.4 Zonas Fitoecológicas Potenciais e Limiares das Evidências

Conforme mencionado em seção anterior, foram utilizadas Zonas Fitoecológicas Potenciais para delimitar áreas com características fitofisionômicas distintas e que remetem a um ambiente passado preservado e íntegro.

A Tabela 1 apresenta as Zonas Fitoecológicas Potenciais (ZFP) utilizadas para a área da Amazônia Legal (AML).

Tabela 1: Zonas Fitoecológicas Potenciais para a área da AML

ZFP	VEGETAÇÃO DE ORIGEM
HOFd	Floresta Ombrófila Densa Terras Baixas
AOFd	Floresta Ombrófila Densa Aluvial
MOFd	Floresta Ombrófila Densa Montana
STF	Floresta Estacional Semidecidual Submontana
OTF/STF	Floresta Ombrófila/Floresta Estacional
LOFs	Vegetação Ombrófila Aberta
HOFs	Vegetação Ombrófila Aberta Submontana
AOFs	Vegetação Ombrófila Aberta Aluvial
PVF	Áreas das Formações Pioneiras Vegetação
WSF/OTF	Campinarana/Floresta Ombrófila
WSFd	Campinarana Florestada
WSFs	Campinarana Gramineo-Lenhosa
SF/OTF	Savana/Floresta Ombrófila
SF/STF	Savana/Floresta Estacional
DTF	Savana Estépica/Floresta Estacional
SWFd	Savana-Estépica Florestada; Savana-Estépica Gramineo-Lenhosa
SWFs	Savana Arborizada; Savana-Estépica Arborizada
SFF	Savana Gramineo-Lenhosa; Savana Parque
HMV	Refúgios Vegetacionais Montano

Para cada ZFP foram estabelecidos os limiares de corte das evidências, para a área da Amazônia Legal. A Tabela 2 apresenta os limiares definidos.

Tabela 2: Limiares das evidências por ZFP para a área da AML

ZFP	CLASSE	EVI	VCF	LAI	GPP
HOFd	Baixa	0,0 - 0,84	0 - 78	0 - 50	0 - 480
	Alta	0,84 - 1,0	78 - 88	50 - 68	480 - 943
LOFd	Baixa	0,0 - 0,85	0 - 72	0 - 51	0 - 582
	Alta	0,85 - 1,00	72 - 88	51 - 68	582 - 943

AOFd	Baixa	0,0 - 0,84	0 - 72	0 - 50	0 - 560
	Alta	0,84 - 1,00	72 - 88	50 - 68	560 - 943
MOFd	Baixa	0,0 - 0,87	0 - 72	0 - 48	0 - 750
	Alta	0,87 - 1,00	72 - 88	48 - 68	750-943
STF	Baixa	0 - 0,81	0 - 57	0 - 45	0 - 560
	Alta	0,81 - 1,00	57 - 88	45 - 68	560 - 943
OTF/STF	Baixa	0,0 - 0,88	0 - 76	0 - 50	0 - 665
	Alta	0,88 - 1,00	76 - 88	50 - 68	665 - 943
LOFs	Baixa	0 - 0,84	0 - 77	0 - 50	0 - 576
	Alta	0,84 - 1,00	77 - 88	50 - 68	576 - 943
HOFs	Baixa	0 - 0,80	0 - 77	0 - 52	0 - 600
	Alta	0,80 - 1,00	77 - 88	52 - 68	600 - 943
AOFs	Baixa	0 - 0,84	0 - 76	0 - 49	0 - 594
	Alta	0,84 - 1,00	76 - 88	49 - 68	594 - 943
PVF	Baixa	0,0 - 0,71	0 - 36	0 - 36	0 - 440
	Alta	0,71 - 1,00	36 - 88	36 - 68	440 - 943
WSF/OTF	Baixa	0,0 - 0,83	0 - 77	0 - 48	0 - 628
	Alta	0,83 - 1,00	77 - 88	48 - 68	628 - 943
WSFd	Baixa	0 - 0,83	0 - 75	0 - 46	0 - 700
	Alta	0,83 - 1,00	75 - 86	46 - 59	700 - 821
WSFs	Baixa	0,0 - 0,70	0 - 56	0 - 35	0 - 690
	Alta	0,70 - 1,00	56 - 88	35 - 68	690 - 943
SF/OTF	Baixa	0 - 0,79	0 - 76	0 - 51	0 - 537
	Alta	0,79 - 1,00	76 - 88	51 - 68	537 - 943
SF/STF	Baixa	0 - 0,78	0 - 72	0 - 48	0 - 480
	Alta	0,78 - 1,00	72 - 88	48 - 68	480 - 943
DTF	Baixa	0,0 - 0,76	0 - 30	0 - 30	0 - 300
	Alta	0,76 - 1,00	30 - 88	30 - 68	300 - 943
SWFd	Baixa	0 - 0,78	0 - 26	0 - 32	0 - 300
	Alta	0,78 - 1,00	26 - 88	32 - 68	300 - 943
SWFs	Baixa	0 - 0,75	0 - 15	0 - 26	0 - 260
	Alta	0,75 - 1,00	15 - 88	26 - 68	260 - 943
SFF	Baixa	0,00 - 0,72	0 - 12	0 - 23	0 - 242
	Alta	0,72 - 1,00	dez/88	23 - 68	242 - 943
HMV	Baixa	0 - 0,85	0 - 60	0 - 45	0 - 700
	Alta	0,85 - 1,00	60 - 88	45 - 68	700 - 943

### 3.5 Modelagem da rede

Na etapa de modelagem da rede, os dados de cada evidência foram discretizados em duas classes: Alta IE e Outros (equivalente aos valores baixos e intermediários), utilizando como referência limiares previamente definidos. A rede foi modelada de modo individual, para cada ZFP, conforme modelo apresentado na Figura 4. Após a etapa de modelagem, as redes foram executadas utilizando o algoritmo de aprendizado.

Os parâmetros utilizados no processo de modelagem da rede foram:

- *Software*: NETICA 5.1.8;
- Tipo de nó: *Nature*;
- Formato dos dados: *Continuous*;
- Discretização: 2 classes;
- Intervalos da discretização: definidos pelos limiares;
- Estado do nó:  $0 = Low / 1 = High$ ;
- Algoritmo de aprendizado: *EM*.

A Figura 7 apresenta um exemplo da rede criada no software para a zona “LOFd”.

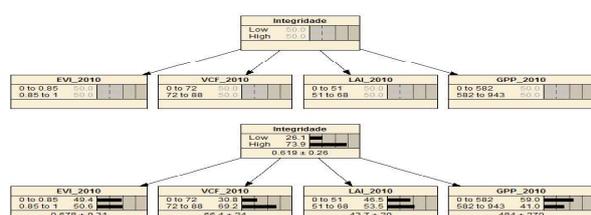


Fig. 7 – Exemplo de rede Bayesiana modelada para a ZFP “LOFd”.

### 3.6 Dados de referência para avaliação dos resultados

Para avaliar as classes “IE Alta” e “IE Baixa” foram utilizadas diferentes bases de dados. Não foi encontrada nenhuma base de referência que permitisse avaliar a classe “IE Intermediária”.

Podem ser citados como fatores que contribuíram para dificuldade em encontrar dados de referência: inovação do tema, escala de trabalho e dimensão da área de estudo.

Para avaliar a classe “IE Baixa” foram utilizados como referência dados de desflorestamentos mapeados até o ano de 2010 pelo

Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite - PRODES, do Instituto Nacional de Pesquisas Espaciais – INPE (<http://www.obt.inpe.br/prodes/index.php>).

Na avaliação dos resultados, partiu-se do pressuposto que uma área desflorestada, em alguma instância de tempo, não pode ter seu ecossistema íntegro, pois o conceito de Integridade Ecológica está relacionado ao seu estado de origem.

Para avaliar os dados de “IE Alta” foram utilizados como referência classes do “Mapa de Cobertura e Uso da Terra do Brasil 2010” (IBGE, 2010), que possuem características de apresentarem maiores índices de integridade. São elas:

Vegetação Florestal: mais de 75% dos polígonos são ocupados por vegetação florestal. Inclui as formações: Floresta Densa, Floresta Aberta, Floresta Estacional, áreas de mangues, e inclui áreas remanescentes primárias e estágios evoluídos de recomposição florestal (capoeirões/capoeiras) das diversas regiões fitogeográficas consideradas como florestais.

Vegetação campestre: incluem-se nessa categoria as Savanas, Estepes, Savanas-Estépicas, Formações Pioneiras e Refúgios Ecológicos. Encontram-se disseminadas por diferentes regiões fitogeográficas, compreendendo diferentes tipologias primárias: estepes planaltinas, campos rupestres das serras costeiras e campos hidro-arenosos litorâneos (restinga).

Vegetação campestre alagada: área ocupada por vegetação natural herbácea (cobertura de 10% ou mais), permanentemente ou periodicamente inundada por água doce ou salobra (estuários, pântanos, etc.). Inclui as áreas de charcos, pântanos, campos úmidos, etc. O período de inundação deve ser no mínimo 2 meses por ano.

Ambos os dados de referência foram processados para serem compatíveis em formato e resolução espacial com os resultados da IE.

## 4. RESULTADOS

A partir de análise preliminar dos resultados da IE para áreas de Unidades de Conservação e para áreas de desmatamento já conhecidas, foram definidas as classes temáticas para elaboração do

mapa de Integridade Ecosistêmica:

- 0,00 a 0,35 – IE Baixa;
- 0,35 a 0,60 – IE Intermediária;
- 0,60 a 1,00 – IE Alta.

A Figura 8 apresenta o mapa da IE gerado utilizando a abordagem metodológica proposta. É importante ressaltar que não fez parte dos objetivos do presente trabalho analisar quais são os fatores e/ou agentes desencadeadores da Integridade Ecosistêmica baixa, intermediária ou alta, sendo este tema objeto para trabalhos futuros

#### 4.1 Integridade Ecosistêmica x Dados de Referência

Utilizando como base de referência os dados apresentados na seção “3.6” foi possível avaliar os resultados para uma área de 3.216.156 km<sup>2</sup>, o que equivale a 65,46% da área total da Amazônia Legal, que é de 4.912.814 km<sup>2</sup> (área calculada com base no limite utilizado da Amazônia Legal, a partir de uma base de dados raster com pixels de 1 km<sup>2</sup>). É importante ressaltar que dos 34,35% de área não validada, 15,21% equivalem a classe IE Intermediária.

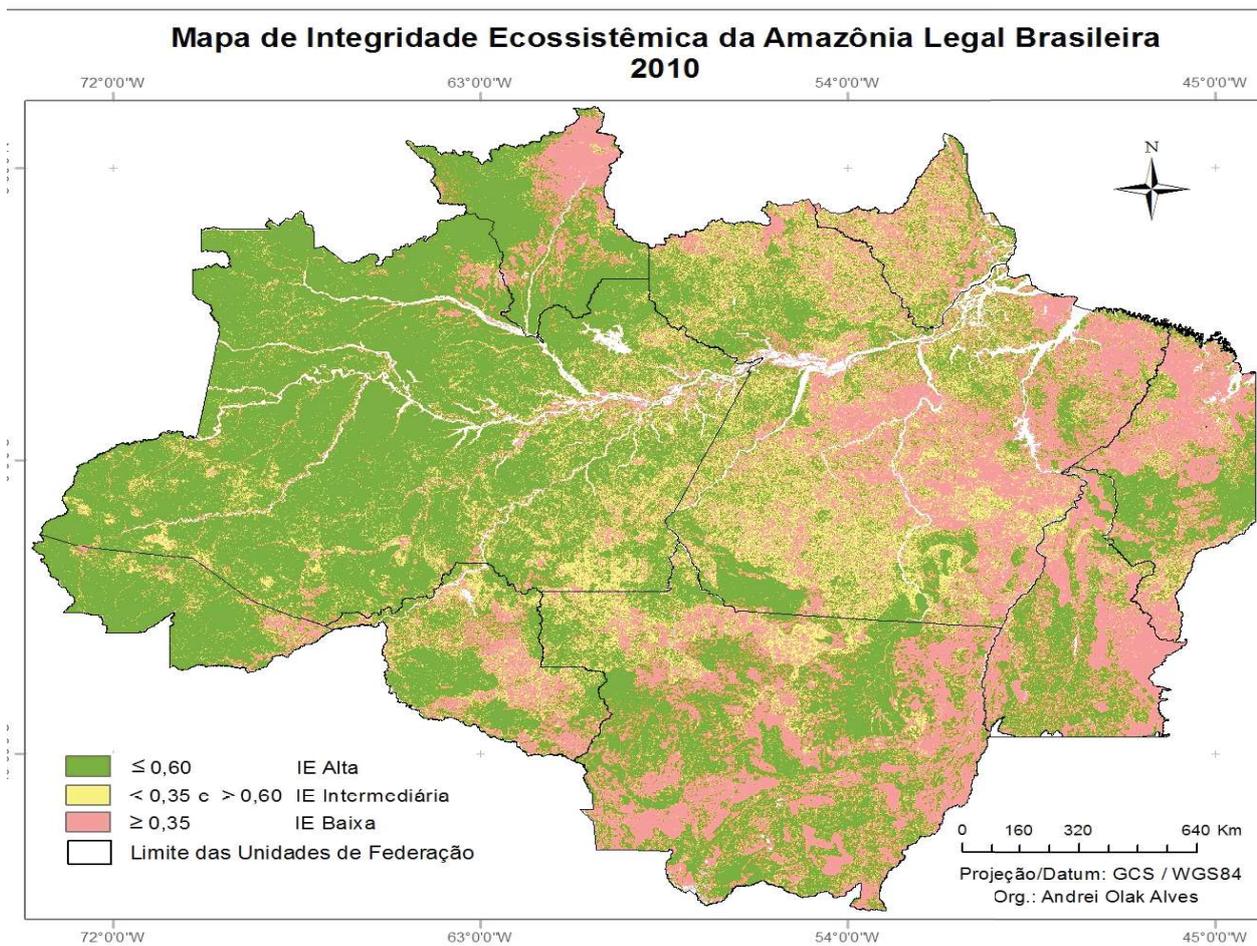


Fig. 8 – Mapa de Integridade Ecosistêmica da Amazônia Legal.

Dos 65,46% da área total avaliada, 12,87% correspondem aos dados de referência de desflorestamento e 52,60% se referem aos dados de referência de vegetação.

A Tabela 3 apresenta uma síntese dos resultados de validação da IE em comparação aos dados de referência para a área da Amazônia Legal. Já a Figura 9 apresenta um mapa contendo a espacialização destes resultados.

Pelos dados apresentados é possível identificar que os acertos foram de 80,49% para IE Alta e de 81,96% para IE Baixa, resultando em uma taxa de acerto global de 80,78%. Já os erros para IE Alta e IE Baixa foram, respectivamente, de 19,51% e 18,04%, resultando em uma taxa de erro global de 19,22%.

O índice Kappa, que mede a concordância entre as classes, foi de 0,51, considerado

moderado. Isso se deu por conta do elevado Erro de Comissão para a classe IE Baixa, 49,32%. Este erro indica que muitos dos elementos que deveriam pertencer à classe IE Baixa foram classificados erroneamente como IE Alta.

Tabela 3: Resultados da avaliação da IE para área da AML

Matriz de Confusão				Erros de Comissão	Kappa
Classes IE Alta		Referência			
		IE Alta	IE Baixa		
Avaliado	IE Alta	2079835	114017	5,20%	0,51
	IE Baixa	504240	518064	49,32%	
Erros de Omissão		19,51%	18,04%	Erro Global 19,22%	
Taxa de classificação por classe		80,49%	81,96%	Exatidão Global 80,78%	

que os principais erros ocorreram nas regiões centro-sudoeste do Estado do Pará, norte do Pará e em grande parte do Estado do Amapá. Estas regiões apresentam como principais Zonas Fitoecológicas Potenciais: HOFd - Floresta Ombrófila Densa Terras Baixas; e HOFs - Vegetação Ombrófila Aberta Submontana.

Na ZPF “HOFd” muitos elementos foram classificados erroneamente como sendo da classe IE Alta. Do total da área avaliada (70,29%), a taxa de exatidão global foi de 66,91%, a taxa de erro global foi de 33,09% e o Kappa foi de 0,34. Analisando os dados de entrada do modelo para esta ZFP pode-se perceber que a região apresentava alto índice de área foliar (LAI), percentual de cobertura de árvores (VCF) e índice de vegetação realçado (EVI) com valores intermediários e baixos valores de produtividade primária bruta (GPP). Neste caso, o modelo priorizou os valores altos e intermediários de LAI, VCF e EVI para considerar a área como IE Alta. No entanto, com base nos dados de referência, a área tenderia mais às informações de GPP.

Pela Figura 9 é possível identificar a espacialização da taxa de acerto e, também,

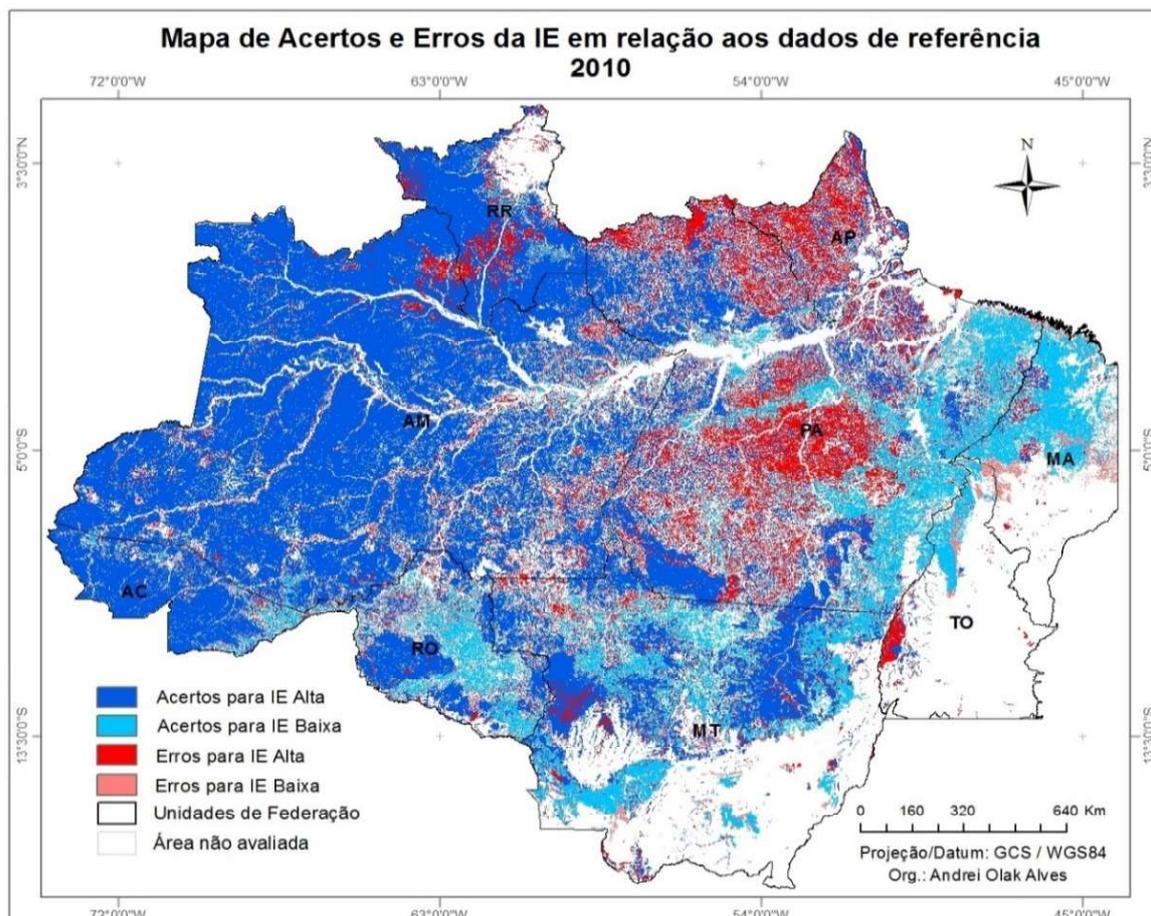


Fig. 9 – Mapa de Acertos e Erros da IE em relação aos dados de referência.

De modo semelhante, na ZPF “HOFs” muitos elementos também foram classificados como sendo da classe IE Alta. Do total da área avaliada (60,58%), a taxa de exatidão global foi de 63,9%, a taxa de erro global foi de 36,1% e o Kappa foi de 0,31. Com base nos dados de entrada do modelo, pode-se perceber que a região apresentava alto índice percentual de cobertura de árvores (VCF), valores intermediários em relação aos Índices de Vegetação Realçados (EVI) e de área foliar (LAI) e baixos valores de produtividade primária bruta (GPP). Neste caso, o modelo também priorizou os valores altos e intermediários de VCF, EVI e LAI para considerar a área como IE Alta.

Estes resultados mostram que o método é sensível aos ajustes realizados no processo modelagem da rede. Nas zonas que apresentaram maiores erros, um novo ajuste do limiar de discretização das evidências poderia ser suficiente para melhorar os resultados. Isso significa que, para uma determinada evidência, o limiar de corte pode ter sido muito ou pouco restritivo para a definição da IE Alta. Portanto, a utilização do conhecimento de especialistas nesta etapa é fundamental para o sucesso da abordagem metodológica.

De modo geral, os resultados foram considerados satisfatórios, apresentando uma consistência de acertos na maior parte da área analisada, com uma taxa de acerto global acima de 80%.

#### 4.2 Resultados da IE para a AML

A resultados da espacialização da Integridade Ecológica para a área da Amazônia Legal mostraram que 53,22% apresentam IE Alta, 15,21% apresentam IE Intermediária e 31,57% apresentam IE Baixa.

A Tabela 4 apresenta os percentuais da IE por UF, levando em consideração a área total da Amazônia Legal. Já a Tabela 4 apresenta os percentuais da IE por UF considerando a área total de cada UF.

Quando considerada a área total da Amazônia Legal, os Estados com maior área (AM, PA, MT) apresentam maior percentual de IE Alta. Quando analisado o resultado levando em consideração a área de cada UF, os Estados Acre, Amazonas e Rondônia apresentam percentual de mais de 50% de IE Alta, seguidos

dos Estados Rondônia e Mato Grosso.

Tabela 4: Percentual de IE por UF em relação à área total da AML

UF	% IE Alta	% IE Intermediária	% IE Baixa
AM	45,24	26,55	7,76
MT	16,50	12,92	25,59
PA	14,23	40,87	33,45
RR	5,32	1,23	4,53
AC	5,18	1,84	1,08
RO	4,49	5,29	5,27
TO	4,08	2,81	9,37
MA	3,56	3,77	8,77
AP	1,40	4,73	4,18

A Tabela 5 apresenta, também, a densidade demográfica (hab/km<sup>2</sup>) de cada Estado. Apesar de os Estados do Amazonas e Roraima apresentarem a menor densidade demográfica da região, os dados mostram que, para a área da Amazônia Legal, o número de habitantes por km<sup>2</sup> não é um indicativo que permite fazer uma relação direta com a Integridade Ecológica.

Tabela 5: Percentual de IE por UF em relação à área da própria UF

UF	% IE Alta	% IE Intermediária	% IE Baixa	Área (km <sup>2</sup> )	Densidade Demogr. (hab/km <sup>2</sup> )
AC	81,64	8,29	10,08	165.762	4,47
AM	78,77	13,21	8,02	1.501.522	2,23
RR	63,65	4,20	32,16	218.474	2,01
RO	49,20	16,57	34,23	238.620	6,58
MT	46,65	10,43	42,92	924.896	3,36
TO	39,09	7,68	53,23	273.154	4,98
MA	36,21	10,95	52,84	257.395	19,81
PA	31,10	25,53	43,37	1.196.135	6,07
AP	26,79	25,80	47,40	136.856	4,69

\* Fonte da Densidade Demográfica: IBGE, Censo 2010.

A Tabela 6 apresenta os percentuais da IE por Zona Fitoecológica Potencial, levando em consideração a área total da Amazônia Legal.

Os resultados mostram que as regiões que apresentam maior índice de Integridade Ecológica são as de Floresta Ombrófila Densa de Terras Baixas (LOFd), Vegetação Ombrófila Aberta (LOFs), e as regiões de transição entre

Campinarana e Floresta Ombrófila (WSF/OTF), quando analisada a proporção entre IE Alta e IE Baixa. Para as áreas de Floresta Ombrófila Densa e de Submontana, apesar do segundo lugar da lista, a porcentagem de IE Baixa representa quase o dobro da IE Alta e, também, apresenta grande quantidade de IE Intermediária. As regiões de Savana Gramíneo-Lenhosa / Savana Parque (SFF) apresentam a mesma proporção de IE Alta e IE Baixa. As áreas com menor IE são as áreas de Floresta Estacional Semidecidual Montana (STF), as áreas de transição entre Savana e Floresta Estacional (SF/STF), Vegetação Ombrófila Aberta Submontana (HOFs) e Savana Arborizada /Savana Estépica Arborizada (SWFs), quando analisada a proporção entre IE Baixa e IE Alta.

Tabela 6: Percentual de IE por ZFP em relação à área total da AML

Zona	% IE Alta	% IE Intermediária	% IE Baixa	Área (km <sup>2</sup> )
LOFd	25,11	17,59	11,27	962.757
HOFd	12,79	31,19	23,57	932.942
SFF	11,68	4,44	11,69	519.791
LOFs	11,12	6,38	3,65	395.212
WSF/OTF	6,71	1,24	0,82	197.395
HOFs	5,62	23,36	11,77	504.037
OTF/STF	4,77	2,08	6,03	233.802
SWFs	3,53	2,53	6,02	204.536
AOFd	3,49	3,02	2,81	157.302
SF/OTF	3,42	3,46	3,21	165.099
SF/STF	2,74	0,07	9,72	222.777
WSFd	2,29	0,71	0,80	77.445
PVF	1,56	1,27	2,07	82.296
AOFs	1,44	1,15	0,83	59.095
MOFd	1,07	0,21	0,46	36.751
STF	0,87	0,35	3,19	74.794
DTF	0,66	0,47	0,57	29.532
WSFs	0,63	0,18	0,63	27.738
SWFd	0,44	0,27	0,82	26.283
HMV	0,07	0,04	0,07	3.230

## 5. CONCLUSÕES

Este trabalho apresentou uma abordagem metodológica e resultados inovadores à comunidade científica, tendo como principal produto um retrato da Integridade Ecológica na Amazônia Legal Brasileira.

Os resultados permitiram concluir que:

- É possível fazer uso de produtos de sensoriamento remoto e modelagem probabilística para mapear a IE em escala regional;
- A abordagem metodológica utilizada se mostrou suficiente para mapear a IE na área da Amazônia Legal Brasileira;
- Os produtos EVI, LAI, VCF e GPP se constituíram como boas evidências do estado da IE;
- A fragmentação da paisagem por Zonas Fitoecológicas Potenciais foi fundamental no processo de modelagem da rede, pois permitiu definir os limiares das evidências para cada tipo de fitossionomia, melhorando os resultados.

Como principais dificuldades apresentadas podem ser citadas a definição e seleção de evidências, a definição dos limiares das evidências no processo de modelagem da rede e a aquisição de dados de referência.

Como principais benefícios desta abordagem podem ser citados:

- Possibilidade de ser aplicada em regiões que necessitam de levantamentos em escala regional e que carecem de dados de mapeamento sistemático atualizados ou de dados coletados em campo, como os inventários florestais.
- Possibilidade de monitorar a IE ao longo do tempo;
- Possibilidade de observar tendências e avaliar a eficiência das políticas públicas na mitigação de impactos ambientais;
- Pode servir para estabelecer a relação entre a perda de biodiversidade e os impactos sobre os serviços dos ecossistemas de florestas tropicais.

Como sugestões para trabalhos futuros podem ser citados a análise dos fatores e/ou agentes desencadeadores da baixa, intermediária ou alta IE e o cruzamento dos resultados da IE com produtos evidências de Serviços Ecológicos, com o objetivo de avaliar a relação entre eles.

## AGRADECIMENTOS

Os autores agradecem à Embrapa Solos, à CAPES (por concessão de bolsa de estudo em nível de Doutorado), ao Programa de Pós-

Graduação em Meio Ambiente – PPGMA da Universidade do Estado do Rio de Janeiro (UERJ), ao projeto EU FP7 ROBIN.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERTI, M. Maintaining ecological integrity and sustaining ecosystem function in urban areas. **Current Opinion in Environmental Sustainability**, v. 2, n. 3, p. 178-184, 2010.
- CHAMBERS, J. Q.; ASNER, G.P.; MORTON, D.C.; ANDERSON, L.O.; SAATCHI, S.S.; ESPÍRITO-SANTO, F.D.B.; PALACE, M. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. **Trends in Ecology & Evolution**, v. 22(8), p. 414–423, 2007.
- COHEN, J. A. Coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, v. XX, n. 1, p. 37–46, 1960.
- DE GROOT, R. Function-analysis and valuation as a tool to assess land use conflicts in planning for sustainable, multi-functional landscapes. **Landscape and Urban Planning**, v.75, p. 175–186, 2005.
- DIMICELI, C.M.; CARROL, M.L.; SOHLBERG, R.A.; HUANG, C.; HANSEN, M.C.; TOWNSHEND, J.R.G. **Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000 – 2010**. Collection 5 Percent Tree Cover, University of Maryland, College Park, MD, USA, 2011. Disponível em: <<http://glcf.umd.edu/data/vcf/>> Acesso em: 15 out 2014.
- DO, C. B.; BATZOGLOU, S. What is the expectation maximization algorithm? **Nature Biotechnology**, v. 26, n. 8, p. 897-899, Aug, 2008.
- EEA - European Environment Agency. **Assessing biodiversity in Europe - the 2010 report**. Copenhagen, n. 5/2010, 2010. 64 p.
- EQUIHUA M.; MAEDA P.D.; EQUIHUA J.; ALANIZ N. G.; KOLB M.; SCHMIDT M. **ROBIN - Role Of Biodiversity In climate change mitigation** - 1st Report 15th of March 2013. 79 p.
- HOLDRIDGE, L.R. Life Zone Ecology. In: II CA. Tropical Science Center, San José, Costa Rica, 1977. **Anais**. p.118, 1977.
- IBGE - Instituto Brasileiro de Geografia e Estatística. **Mapa de uso e cobertura da terra do Brasil**. 2010. Disponível em: <[ftp://geoftp.ibge.gov.br/mapas\\_tematicos/mapas\\_murais/shapes/cobertura\\_e\\_uso\\_da\\_terra/](ftp://geoftp.ibge.gov.br/mapas_tematicos/mapas_murais/shapes/cobertura_e_uso_da_terra/)> Acesso em: 23 set. 2015.
- LAURIA E. J. M.; DUCHESSI, P. J. A methodology for developing bayesian networks: an application to information technology (IT) implementation. **European Journal of Operational Research**, v.179, p. 24-252, 2006.
- LIANG, S.; XIAO, Z. **Global Land Surface Products: Leaf Area Index Product Data Collection (1985-2010)**, Beijing Normal University, 2012. Disponível em: <<http://glcf.umd.edu/data/lai/>> Acesso em: 15 mai 2014.
- LINDLEY D. V. **Bayesian Statistics, a Review**. Philadelphia, PA: SIAM, 1972. 83 p.
- MATHER, P. M. **Computer processing of remotely-sensed images: an introduction**. 3 ed. Nottingham: John Wiley & Sons, 2005. 460 p.
- NASA. **MODIS Moderate resolution imaging spectroradiometer: about**. 2013. Disponível em: <<https://modis.gsfc.nasa.gov/about/>>. Acesso em: 10 set. 2014.
- NASA/USGS. LP DAAC - **Land processes distributed active archive center**. MODIS products table. 2014. Disponível em: <[https://lpdaac.usgs.gov/dataset\\_discovery/modis/modis\\_products\\_table/](https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/)>. Acesso em: 05 de Maio 2014.
- OLTHOF, I., POULIOT, D., FRASER, R., CLOUSTON, A., WANG, S., CHEN, W., & SAWADA, M. Using Satellite Remote Sensing to Assess and Monitor Ecosystem Integrity and Climate Change in Canada's National Parks. In: **Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on**. IEEE, 2006. p. 686-689.
- DA SILVA PEREIRA, P., CHAVES FERNANDES, L. A., LOPES DA MOTA OLIVEIRA, J., & FERNANDES BAPTISTA, D. Avaliação da integridade ecológica de rios em áreas do zoneamento ecológico econômico do complexo hidrográfico Guapiaçu-Macacu, RJ,

- Brasil. **Ambiente & Água-An Interdisciplinary Journal of Applied Science**, v. 7, n. 1, 2012.
- PETROU, Z.; PETROU, M. 2011. A review of remote sensing methods for biodiversity assessment and bioindicator extraction. In: **Space Technology (ICST), 2nd International Conference on, 2011. Anais**. p. 1- 5, Sept, 2011.
- SIMÕES, M.; FERRAZ, R. P. D.; VERWEIJ, P.; EQUIHUAL, M.; MAQUEO, O.; ALVES, A. O. Modelo Bayesiano de dados espaciais aplicado ao mapeamento da integridade ecológica da Amazônia brasileira. **Enviroment et Geomatique: approches comparées France-Brésil, 2014. Anais. São Paulo / Rennes: Universidade de São Paulo / Université Rennes**, p. 371-378, 2014.
- SIMÕES, M.; FERRAZ, R.; VERWEIJ, P. ; ALVES, A.O. ; EQUIHUAL, M. ; MAQUEO, O. . Integridade ecológica da Amazônia brasileira através do uso de redes bayesianas e suas relações com os serviços ecológicos. In: **XVII Simpósio Brasileiro de Sensoriamento Remoto, 2015. Anais. São José dos Campos: INPE**, p. 5240-5247, 2015.
- TIERNEY, G. L., FABER-LANGENDOEN, D., MITCHELL, B. R., SHRIVER, W. G., & GIBBS, J. P. Monitoring and evaluating the ecological integrity of forest ecosystems. **Frontiers in Ecology and the Environment**, v. 7, n. 6, p. 308-316, 2009.
- TSO, B.; MATHER, P.M. **Classification methods for remotely sensed data**. 2 ed. New York: CRC Press, 2009. 347 p.
- TURNER, W.; SPECTOR, S.; GARDINER, E. P.; FLADELAND, M.; STERLING, E. STEININGER, M. Remote sensing for biodiversity science and conservation, **Trends in Ecology & Evolution**, v. 18(6), p. 306-314, 2003.
- VERWEIJ P.; SIMOES, M.; ALVES, A.; FERRAZ, R.; CORMONT, A. 2014. Linking Bayesian Belief Networks and GIS to assess the Ecosystem Integrity in the Brazilian Amazon. International Environmental Modelling and Software Society (iEMSs), 2014. **7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA, 2014. Anais**. Disponível em: <<http://www.iemss.org/society/index.php/iemss-2014-proceedings>>. Acesso em: 05 abr 2015.