

Análise multivariada aplicada a zoneamento para a predição de desmatamento: a validação do método AMAZON-PD

Darcton Policarpo Damiano¹
Eduardo Henrique Geraldi Araújo²

¹Usinas Siderúrgicas de Minas Gerais S/A– USIMINAS
R. Prof. José V. Mendonça, 3011 – 31.310-260 – Belo Horizonte – MG, Brasil
darcton@ieav.cta.br

²DCT – Consultoria e Treinamento Ltda
Avenida Tívoli, 443/10 – 12.245-230 – São José dos Campos – SP, Brasil
eduaraujo@gmail.com

Abstract. This is a validation paper of its predecessors – “Análise multivariada aplicada a zoneamento: o método AMAZON”, and “Análise multivariada aplicada a zoneamento para a predição de desmatamento: o método AMAZON-PD”. By shortening the period of study and simultaneously increasing the observation frequency, it was possible to verify that not only the method designed is capable to predict deforestation sites but its performance can be improved by a great deal as well. The results show better values for the parameter defined as DM (meaning model performance) and a much “leaner” buffer of predicted deforestation sites. This is a direct indication that a smaller area is needed to be controlled by taking AMAZON-PD into account in order to obtain the same amount of efficiency, *ceteris paribus*. Once more, the attained results, which accuracy was verified upon classification tables, allowed inferring the future location of deforestation sites with striking precision.

Palavras-chave: deforestation, remote sensing, geographical information systems, multivariate analysis, logistic regression, land cover change, desmatamento, sensoriamento remoto, sistemas de informação geográfica, análise multivariada, regressão logística, mudanças na cobertura do solo.

1. Introdução

O Brasil possui cerca de 3,6 milhões de km² de florestas e ocupa o terceiro lugar na lista dos países com maior área florestal densa do mundo, fato que o deixa atrás apenas da Rússia e do Canadá. Em termos de florestas tropicais, no entanto, o Brasil é o país com maior cobertura, possuindo uma área três vezes superior à existente na República Democrática do Congo, segundo colocado no ranking mundial para esses tipos de florestas (FAO, 1995).

O aumento dos preços da madeira nos mercados nacional e internacional tornou a exploração madeireira uma atividade atraente. Isso aumentou a quantidade de espécies exploradas. Somem-se a isso as previsões e projeções de exaustão dos recursos florestais do sudeste da Ásia, que hoje ainda suprem em elevada proporção os mercados mundiais de madeiras tropicais, e pode-se ter uma boa idéia da importância que a região amazônica detém, somente nesse contexto.

Outros fatores, além da atração econômica dos preços alcançados pela madeira de lei, provocaram o recrudescimento da exploração madeireira na Amazônia. O governo brasileiro, por exemplo, desempenhou um papel preponderante nesse panorama, tanto em função de iniciativas tomadas, como pela falta delas.

Ainda que as exportações de madeira da Amazônia representem uma parcela modesta do mercado internacional de madeiras tropicais, há uma tendência natural de crescimento desse mercado, em vista da expectativa de escassez de madeiras nobres no futuro próximo.

A se confirmar essa tendência, a exploração madeireira será a principal atividade econômica de uso do solo em toda a Amazônia. Portanto, adequar práticas de manejo à dinâmica dos mercados nacional e internacional será o próximo grande desafio para a Amazônia. Além disso, há que se considerar que a expansão da atividade madeireira já atinge várias áreas de uso restrito, como terras indígenas e unidades de conservação.

Tão logo perdem seu interesse vinculado ao valor intrínseco das madeiras de lei, as terras amazônicas assumem o interesse provocado pela atividade pecuária extensiva, o que provoca o efeito mais visível de impacto ambiental, que consiste do avanço da frente de desmatamento sobre a floresta.

O controle sobre o meio ambiente em uma região extensa e de difícil acesso como a Amazônia ainda está muito longe de um padrão minimamente aceitável. Ainda que haja denúncias e processos, a falta de meios e de métodos, por parte dos órgãos competentes, induz as indústrias a permanecer na ilegalidade, dilapidando os patrimônios natural e nacional existentes na região.

Existem preocupações por parte da sociedade brasileira e isso se verifica na elaboração cuidadosa de leis ambientais. No entanto, a legislação pertinente não passa de uma carta de intenções, com pouco ou nenhum resultado de ordem prática, dada a limitação do estado brasileiro para implantá-la a contento.

Nesse sentido, foi idealizado um modelo capaz de prover uma ferramenta voltada a políticas ambientais de amplo alcance. Em especial, busca-se uma orientação ao emprego de meios de controle de impactos ambientais com base na predição estatística desses impactos.

2. O método AMAZON-PD

Conforme demonstrado no estudo de tese que deu origem a este artigo, o método AMAZON-PD é definido pela equação a seguir:

$$\text{Logito}\left[P(IPT_{pd})\right] = \ln\left[\frac{P(IPT_{pd})}{1 - P(IPT_{pd})}\right] = \beta_{0_{reg}} + \beta_{1_{reg}} PAD_{ab} + \beta_{2_{reg}} ROD_{ab} + \beta_{3_{reg}} HID_{ab} + \beta_{4_{reg}} EPA_{ab} + \beta_{5_{reg}} DEC_{ab} + \beta_{6_{reg}} ALT_{ab} + \beta_{7_{reg}} IFR_{ab} + \beta_{8_{reg}} APE_{ab} + \beta_{9_{reg}} W_{1ab} + \beta_{10_{reg}} W_{2ab} \quad (1)$$

onde $\beta_{i_{reg}}$ são os coeficientes das variáveis independentes, obtidos via regressão a partir dos coeficientes dessas variáveis até a data anterior ao ano-base.

A opção pelo estudo da cobertura do solo (desmatamento) e não de seu uso, que pode ter diferentes vertentes (extração madeireira, pecuária extensiva e agricultura, por exemplo) evita uma forçosa adoção de variáveis de cunho socioeconômico, o que foge ao escopo do estudo. Não obstante essa busca por uma maior simplicidade em termos de modelagem, deve-se ressaltar que a compreensão do efeito (cobertura do solo = desmatamento), a partir de uma análise de fatores intervenientes (variáveis independentes) que possibilite sua predição em datas futuras, permite lançar um renovado olhar sobre suas causas (uso do solo = extração madeireira ou pecuária ou agricultura).

Em outras palavras, a compreensão dessas causas (mudanças no uso do solo), adquirida pelo vasto conhecimento já acumulado a partir de fatores de ordem socioeconômica e extensamente tratados na literatura, pode ser bastante enriquecida pelo aporte da abordagem aqui proposta.

O uso de um modelo restrito a “apenas” uma cena Landsat (aproximadamente 34000 km²) se deve ao simples fato de que o custo computacional de estendê-lo a toda a região amazônica, que requer um total de 229 imagens para ser coberta pelo satélite Landsat, tornaria inviável o estudo de tese, no qual se baseia este trabalho (DAMIÃO, 2007), a se considerar a concepção pixel-a-pixel adotada.

Assim, foi necessária a reamostragem da cena em células de 4 x 4 pixels, de maneira a permitir o processamento dos diversos modelos com mais de 2,5 milhões de amostras cada. Essa condição ressalta ainda mais a importância do estudo em função da impossibilidade de se realizar um estudo com tantos dados num pacote estatístico convencional. Não menos trabalhosa seria a necessidade de formatação dos dados para a inserção nesses aplicativos.

No estudo de tese, não foi possível analisar a evolução anual do desmatamento, por conta do volume de dados que seria necessário processar para testar a hipótese nele formulada. Na continuação daquele estudo, porém, optou-se por analisar um novo período, dessa feita de 2000 a 2007, com uma frequência anual de observação. A escolha das variáveis foi feita de tal maneira que todas elas fossem incluídas, à exceção de ‘HID’ (não há na região estudada rede hidrográfica que justifique seu uso).

Com isso, é possível verificar a robustez do AMAZON em geral e do AMAZON-PD em particular. Como era de se esperar, o aumento do número de cenas analisadas por período (8 cenas em 8 anos, contra 6 cenas em 20 anos) melhorou sensivelmente a aderência dos coeficientes gerados nas regressões.

O resultado foi a geração de mapas de tendência de desmatamento mais precisos, com parâmetros de desempenho que são corroborados tanto pelos valores constantes das tabelas de classificação (DM, AM, VP e FP), quanto pela visualização mais “enxuta” para a predição de desmatamento, conforme se pode verificar na Tabela 1 e na Figura 1, que estabelece uma comparação de desempenho entre os modelos AMAZON₀₀₋₀₄ (com as variáveis ‘PAD’, ‘ROD’, ‘EPA’, ‘DEC’ e ‘W1’) e AMAZON₀₀₋₀₁ (com as variáveis ‘PAD’, ‘ROD’, ‘EPA’, ‘DEC’, ‘ALT’, ‘W1’ e ‘W2’).

Sua eficiência é medida por meio de tabelas de classificação, nas quais os valores esperados para ‘floresta’ e ‘desmatamento’ são confrontados com os valores observados para essas mesmas classes. Com isso, são gerados os termos de ajuste do modelo¹ (AM), de verdadeiros positivos² (VP) e de falsos positivos³ (FP), que permitem escolher o modelo preservado para o ano-base em questão, com base no desempenho do modelo⁴ (DM), cujos critérios são tais que: AM ≥ 70%; VP ≥ 70%; e FP < 30%.

O impacto imediato desta nova formatação do método AMAZON-PD (frequência anual) pode ser dimensionado pela área a ser coberta em uma eventual ação de controle sobre as áreas a proteger, por parte do Poder Público, 45% menor do que aquela a ser delimitada pelo método desenvolvido anteriormente (DTO esp). As implicações operacionais e logísticas decorrentes desta evolução são evidentes, naturalmente.

TABELA 1 – Comparação de desempenho das tabelas de classificação dos modelos AMAZON₀₀₋₀₄ e AMAZON₀₀₋₀₁.

AMAZON ₀₀₋₀₄	DM = 0,378		AMAZON ₀₀₋₀₁	DM = 0,478	
2387747	DTO obs	FLT obs	2298715	DTO obs	FLT obs
DTO esp	196788	688244	DTO esp	68049	419981
FLT esp	32164	1470551	FLT esp	19053	1791632
69,83%	85,95%	31,88%	80,90%	78,13%	18,99%
(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)	(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)

¹ - Percentual resultante do somatório das predições corretas para as classes ‘floresta’ e ‘desmatamento’, dividido pelo total de células da máscara analisada.

² - Percentual resultante da divisão do número de células corretamente preditas para a classe ‘desmatamento’ pelo número total de células dessa classe.

³ - Percentual resultante da divisão do número de células erroneamente preditas para a classe ‘desmatamento’ pelo número total de células da classe ‘floresta’.

⁴ - DM = AM x (VP – FP)

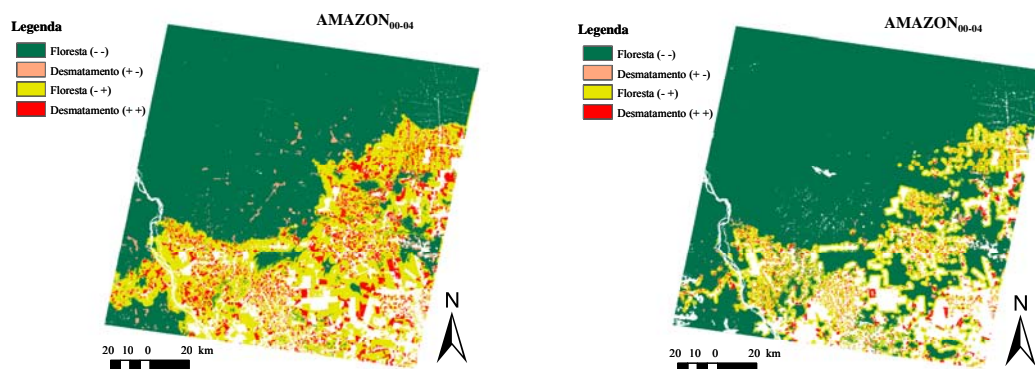


Figura 1. Comparação de desempenho entre os modelos $AMAZON_{00-04}$ e $AMAZON_{00-01}$.

3. Resultados e Discussão

3.1 Desempenho do AMAZON

A espacialização das operações de regressão logística feita para os anos-base de 2001 a 2006 é apresentada a seguir, nas tabelas de 2 a 4, acompanhadas pelas figuras de 2 a 4. Comparativamente aos resultados obtidos no estudo de tese, o AMAZON apresentou desempenho sensivelmente superior, tanto em termos quantitativos quanto em termos qualitativos.

TABELA 2 – Tabelas de classificação dos modelos $AMAZON_{01-02}$ e $AMAZON_{02-03}$.

$AMAZON_{01-02}$	DM = 0,421		$AMAZON_{02-03}$	DM = 0,343	
2207994	DTO obs	FLT obs	2176711	DTO obs	FLT obs
DTO esp	24327	540191	DTO esp	50495	590593
FLT esp	5819	1637657	FLT esp	16324	1519299
75,27%	80,70%	24,80%	72,12%	75,57%	27,99%
(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)	(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)

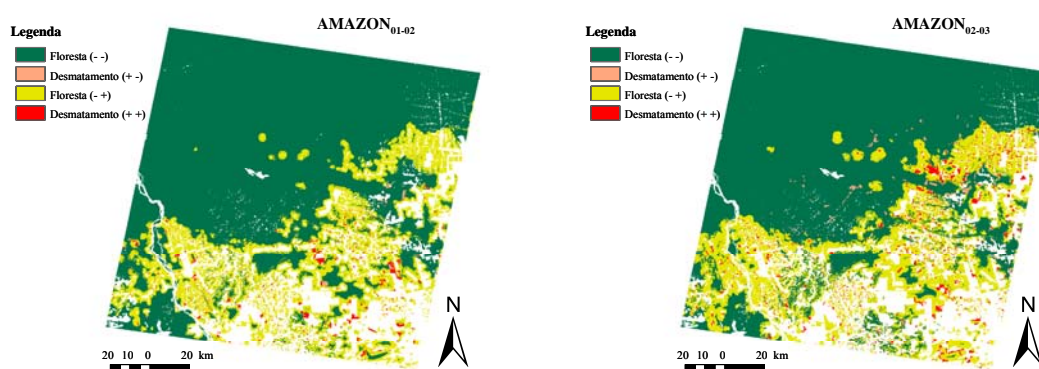


Figura 2. Espacialização das regressões logísticas executadas para os modelos $AMAZON_{01-02}$ e $AMAZON_{02-03}$.

TABELA 3 – Tabelas de classificação dos modelos AMAZON₀₃₋₀₄ e AMAZON₀₄₋₀₅.

AMAZON ₀₃₋₀₄	DM = 0,375		AMAZON ₀₄₋₀₅	DM = 0,372	
2107338	DTO obs	FLT obs	2052379	DTO obs	FLT obs
DTO esp	40494	541707	DTO esp	30072	314288
FLT esp	11970	1513167	FLT esp	19896	1688123
73,73%	77,18%	26,36%	83,72%	60,18%	15,70%
(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)	(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)

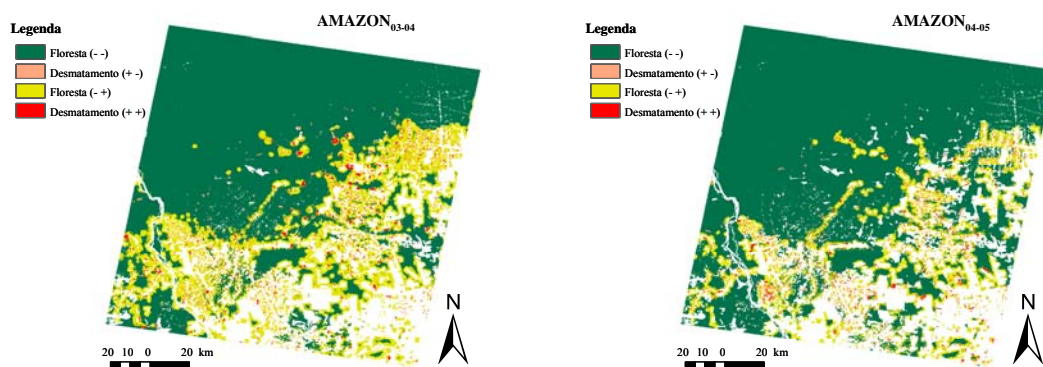


Figura 3. Espacialização das regressões logísticas executadas para os modelos AMAZON₀₃₋₀₄ e AMAZON₀₄₋₀₅.

TABELA 4 – Tabelas de classificação dos modelos AMAZON₀₅₋₀₆ e AMAZON₀₆₋₀₇.

AMAZON ₀₅₋₀₆	DM = 0,397		AMAZON ₀₆₋₀₇	DM = 0,392	
2002411	DTO obs	FLT obs	1962550	DTO obs	FLT obs
DTO esp	33109	550814	DTO esp	24413	555390
FLT esp	6752	1411736	FLT esp	4792	1377955
72,16%	83,06%	28,07%	71,46%	83,59%	28,73%
(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)	(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)

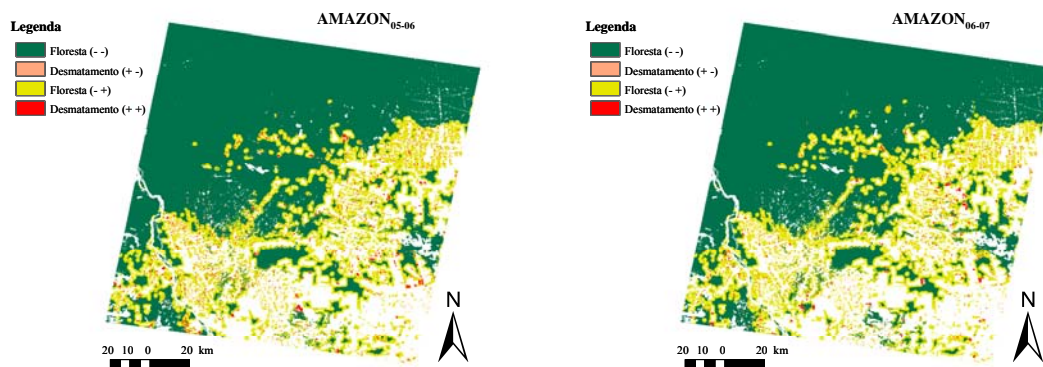


Figura 4. Espacialização das regressões logísticas executadas para os modelos AMAZON₀₅₋₀₆ e AMAZON₀₆₋₀₇.

Qualitativamente, pode-se observar uma predição bem mais “enxuta” do que aquelas verificadas anteriormente, como a espacialização ilustrada pela Figura 1 (AMAZON₀₀₋₀₄). Além disso, os valores alcançados para caracterizar o desempenho dos modelos são, via de regra, mais elevados e mais consistentes, uma vez que atendem os critérios críticos pré-estabelecidos. A única exceção ocorre no modelo AMAZON₀₄₋₀₅, cujos verdadeiros positivos alcançaram apenas 60,18%, quando deveriam ultrapassar os 70%.

A razão para isso está no processo de eliminação de variáveis, que deve ser feito caso-a-caso, com vistas a se obter o modelo mais aderente possível. Como o volume de dados é muito grande, este aprimoramento deve ser buscado para a próxima versão do AMAZON, que busca executar toda a operação de forma automatizada. Com isso, espera-se um ganho simultâneo de desempenho e de precisão.

3.2 Desempenho do AMAZON-PD

A espacialização das operações de regressão logística feita para os anos-base de 2001 a 2006 é apresentada a seguir, nas tabelas de 2 a 4, acompanhadas pelas figuras de 2 a 4.

TABELA 5 – Tabela de classificação do modelo AMAZON-PD₀₀₋₀₇.

AMAZON ₀₅₋₀₆		DM = 0,397	
2002411	DTO obs	FLT obs	
DTO esp	33109	550814	
FLT esp	6752	1411736	
72,16%	83,06%	28,07%	
(ajuste do modelo)	(verdadeiros positivos)	(falsos positivos)	

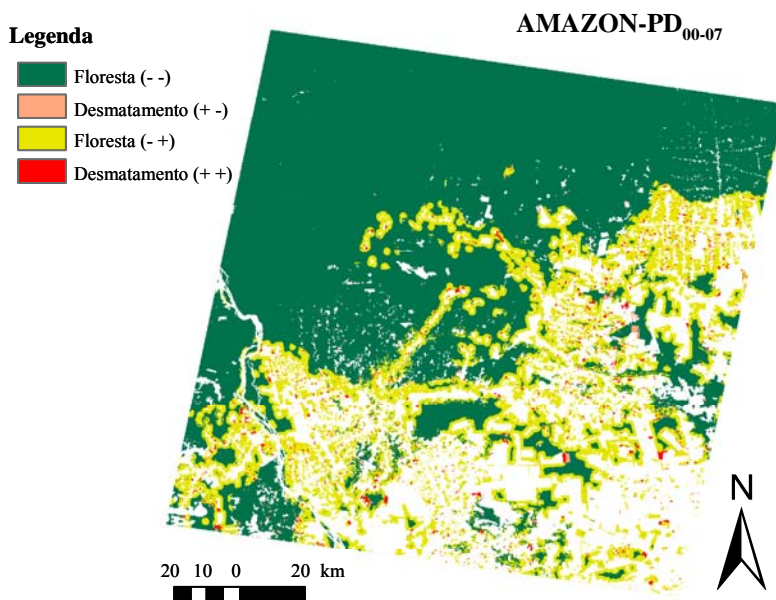


Figura 5. Desempenho do AMAZON-PD no ano-base de 2006 (projeção para 2007), levando-se em conta as variáveis ‘PAD’, ‘ROD’, ‘EPA’, ‘DEC’, ‘ALT’, ‘W1’ e ‘W2’ e baseando-se nos coeficientes ‘ β_n ’ dos anos-base anteriores a 2006 (2000, 2001, 2002, 2003, 2004 e 2005).

Comparativamente aos resultados obtidos com o AMAZON-PD do estudo de tese, também se verifica a melhora nos aspectos qualitativo e quantitativo. Além disso, é importante ressaltar a grande vantagem inerente ao AMAZON-PD, que é a sua capacidade em estabelecer os parâmetros dos coeficientes de predição sem que seja necessário conhecer a variável-resposta cujo comportamento se quer antecipar.

4 Conclusões

O aprimoramento contínuo do AMAZON-PD o torna uma ferramenta operacional muito eficiente para o processo de auxílio à decisão no combate ao desmatamento ilícito da Amazônia, especialmente quando estiver automatizado o processo de seleção de variáveis, a escolha e aplicação do método de análise multivariada (até agora, restringiu-se à regressão logística, por questões acadêmicas) e a regressão dos coeficientes.

A agilidade no manuseio dos dados, a compatibilidade das extensões dos arquivos – que tornou desnecessárias transformações nos dados para adequá-los a outros aplicativos – e a redução do tempo de processamento – com menos tempo perdido na formatação de informações – tornam o AMAZON-PD uma “solução elegante”, na medida em que ele permite usar um único software em sua construção.

Por outro lado, são diversas as razões que conferem ao AMAZON-PD uma abordagem inédita na modelagem de cobertura do solo. Uma delas está no fato de que, ao invés de extrapolar o conhecimento obtido de certo número de amostras para todo o universo amostral, por meio de inferências estatísticas, o método é descritivo estatisticamente, ou seja, reflete o comportamento de todas as amostras, sem exceção.

A consistência dos resultados alcançados com o AMAZON-PD se fundamenta na observação sistemática e de longa duração de uma mesma área. Essa característica, aliada ao uso conjunto de variáveis categorizadas de exposição, de controle e de interação (multiplicativas), permitiu a obtenção de resultados concretos sobre a dinâmica do desmatamento no município de São Félix do Xingu.

É bastante razoável presumir um comportamento similar para essas variáveis em outros pontos da Amazônia. Mesmo que não venha a ser extrapolável nos parâmetros, ainda assim o AMAZON-PD certamente o será nos procedimentos metodológicos de modelagem.

Outras duas contribuições científicas do AMAZON-PD merecem destaque. Em primeiro lugar, ao abstrair questões socioeconômicas – importantes, mas de difícil caracterização em uma escala regional – este método provê contribuições científicas à área de política e gestão ambiental. Como se baseiam em variáveis biofísicas, os mapas de tendência gerados a partir do AMAZON-PD permitem uma alocação adequada dos recursos sempre finitos em relação à área virtualmente infinita que constitui a Amazônia brasileira e cuja preservação se almeja.

Referências Bibliográficas

Damião, D. P. **Uso de técnicas de análise multivariada para a predição de desmatamento na Amazônia: O modelo ‘AMAZON-PD’**. 2007. 125 f. Doutorado (Tese em Desenvolvimento Sustentável) – Centro de Desenvolvimento Sustentável – CDS, Universidade de Brasília – UnB, Brasília – DF

Damião, D. P.; Souza, N. M. **Análise multivariada aplicada a zoneamento: o método AMAZON**. No prelo.

Damião, D. P.; Souza, N. M. **Análise multivariada aplicada a zoneamento para a predição de desmatamento: o método AMAZON-PD**. No prelo.

Food and Agriculture Organization (FAO). **Planning for Sustainable Use of Land Resources**. FAO Land and Water Bulletin 2. Rome: Food and Agriculture Organization of the United Nations. 1995.