

VERIFICAÇÃO DA EXATIDÃO EM CLASSIFICAÇÃO DIGITAL DE POVOAMENTOS FLORESTAIS EM IMAGEM ORBITAL MEDIANTE TRÊS ÍNDICES

ÉDSON LUIS BOLFE¹

ELIANA LIMA DA FONSECA¹

RUDINEY SOARES PEREIRA²

PEDRO ROBERTO DE AZAMBUJA MADRUGA²

¹EMBRAPA - Embrapa Tabuleiros Costeiros
Caixa Postal 44 - 49001-970 - Aracaju - SE, Brasil
{bolfe, eliana}@cpatc.embrapa.br

²UFSM - Universidade Federal de Santa Maria
Caixa Postal 15 - 97119-900 - Santa Maria - RS, Brasil
{rudiney, madruga}@ufsm.br

Abstract. The use of products of forest origin has increased significantly in the last decades. The search for alternatives presupposing high productivity, brought the introduction of exotic species, mainly *Eucalyptus* sp. and *Pinus* sp. This work evaluated the precision of the gotten digital classification in the survey of forest implanted stands and natural in the region of the of the systematic map area of Cachoeira do Sul - RS, using geoprocessing techniques, remote sensing, SIG (system of geographic information) and GPS (global positioning system). It was verified that the area is occupied by natural vegetation (35,54 %), *Pinus* sp. (1,89 %) and *Eucalyptus* sp. (0,77 %), whose precision in the supervised digital classification was the following one: Global precision (85,23 %), Kappa (84,90 %) and Tau (77,74 %). It was concluded that the three accuracy indexes can be used, although the indexes Kappa and Tau werw more consistent.

Keywords: geoprocessing, accuracy indexes, forest surveys.

1. Introdução

A demanda por produtos de origem florestal aumentou sensivelmente nas últimas décadas, levando a silvicultura a buscar alternativas que pressupõe alta produtividade, concretizando-se com a introdução de espécies exóticas, principalmente *Eucalyptus* sp. e *Pinus* sp.

Tradicionalmente, o levantamento do potencial de recursos florestais não é efetivo, pois demanda elevado fluxo de informações, as quais são difíceis de alocar, interpretar e armazenar de tal forma que garantam confiável banco de dados. Neste sentido, buscou-se no presente trabalho, sistematizar informações de povoamentos florestais através da utilização de técnicas de geoprocessamento e sensoriamento remoto.

Conforme Silva et al. (1998), geoprocessamento é um conjunto de procedimentos computacionais que, operando sobre bases de dados geocodificados, executa análises, reformulações e síntese sobre os dados ambientais, tornando-se utilizáveis em um sistema de processamento automático. Novo (1992), define sensoriamento remoto como sendo a tecnologia da utilização conjunta de sensores remotos, equipamentos de processamento e transmissão de dados, com o objetivo de estudar o ambiente terrestre através de registro e análise das interações eletromagnéticas com elementos componentes do planeta Terra, em suas mais diferentes manifestações.

Dessa forma, levantou-se os povoamentos florestais a partir de técnicas de classificação digital supervisionada em seu método de máxima verossimilhança, onde segundo Pereira et al. (1995), o algoritmo de classificação consiste num princípio estatístico paramétrico, considerando as classes envolvidas numa função densidade de probabilidade gaussiana.

Ao trabalhar com mapas e informações provenientes de imagens de satélite, definiu-se avaliar a precisão da classificação digital dos temas abordados. Essa análise é de fundamental

importância sob o ponto de vista técnico e de subsídio a instrumentos de fomento governamental. Conforme Story & Congalton (1986), a maneira mais comum para expressar a precisão, tanto de imagens quanto de mapas, está na declaração da porcentagem da área de mapa que foi corretamente classificada quando comparada com dados de referência ou “verdade de campo”, denominada de Exatidão global. Esta declaração normalmente é derivada de uma contra parte da classificação correta gerada por amostragem dos dados classificados, e expressa na forma de matriz de erro, algumas vezes denominada de matriz de confusão ou tabela de contingência. Jensen (1986), sugere que a análise comparativa da precisão específica local, através da matriz de erro pode fornecer métodos melhores de comparar a precisão de mapeamentos de uso da terra, que uma simples comparação de precisão através de uma estimação global.

Entretanto, Gong & Howarth (1990) utilizaram o índice Kappa (K) como uma medida de precisão importante a ser associada à matriz de erro, por representá-la inteiramente, isto é, considerando todos os elementos da matriz e não apenas aqueles que se situam na diagonal principal, como acontece com o índice de Exatidão global.

Foody (1992), ressaltou que o grau de concordância por chance poderia estar sendo superestimado, pelo fato de se incluir também a concordância real, e que por conta disso a magnitude de Kappa não refletiria a concordância presente na classificação, apenas descontada a casualidade. Isso gerou um outro índice para a medição da precisão de classificação, definido por Ma & Redmond (1995), como índice Tau (T).

Dessa forma, o presente trabalho objetivou avaliar os índices de acurácia: Exatidão Global, Kappa e Tau nos dados obtidos sob a classificação digital supervisionada de povoamentos florestais implantados e nativos da carta sistemática DSG/IBGE SH. 22-Y-A / Cachoeira do Sul, RS.

2. Material e Métodos

A região de estudo situa-se na parte central do Estado do Rio Grande do Sul, correspondendo à carta sistemática DSG/IBGE - SH. 22-Y-A, escala 1:250.000, referente à Cachoeira do Sul. Esta região abrange os municípios de Cachoeira do Sul, São Sepé, Caçapava do Sul, Lavras do Sul, Encruzilhada do Sul, Santana da Boa Vista e Vila Nova do Sul. Possui dimensão de 16.124,20 km², representando 5,60 % da área total do Estado. Localiza-se na transição das regiões fisiográficas do Escudo Riograndense e Depressão Central, fazendo parte da Bacia Hidrográfica do Rio Jacuí. Conforme RadamBrasil (1986), os solos predominantes na região são os Podzólicos e secundariamente Brunizén, além dos Planossolos nas planícies aluviais. A topografia apresenta-se de duas formas, onde ao norte região encontra-se altitudes em torno de 50 á 150 metros com declividades não superior a 15°, já ao sul encontra-se altitude variando entre 150 e 300 metros, ocorrendo declividades de até 25°. A economia da região é baseada no setor primário, onde destacam-se a agricultura irrigada (arroz) e de sequeiro, (soja e milho), as áreas reflorestamento, pecuária extensiva e a extração mineral (calcário).

Os materiais cartográficos utilizados no trabalho foram imagens TM / Landsat 5 de 1995, carta topográfica SH 22-Y-A na escala 1: 250.000 e carta de solos na escala 1: 750.000. O processamento e análise efetuou-se em microcomputadores, mesa digitalizadora, empregando o sistema de informações geográficas Idrisi for Windows. Nos trabalhos de campo utilizou-se GPS (sistema de posicionamento global por satélite).

A metodologia empregada foi subdividida nas seguintes etapas: (a) processamento digital das imagens, (b) georreferenciamento, (c) classificação temática, (d) quantificação do uso da terra e (e) aferição da acurácia temática.

No processamento digital das imagens TM, utilizou-se as bandas espectrais: 3, 4 e 5 de intervalo espectral 0,63-0,69 μm , 0,76-0,90 μm e 1,55-1,75 μm . Estas foram processadas de forma a elaborar-se a composição colorida, associando-se as bandas 3, 4 e 5 aos canais azul, verde e vermelho respectivamente.

Na etapa de georreferenciamento das imagens, objetivou-se referenciá-las geograficamente ao terreno por meio de sua localização espacial, utilizando-se o sistema de coordenadas UTM (*Universal Transversa de Mercator*). Foram alocadas coordenadas de pontos visíveis na imagem e na carta sistemática, de forma a estabelecer-se 22 pontos de controle. Conforme Eastman (1994), o erro admitido no processo de georreferenciamento para a elaboração de mapeamentos digitais é de 1/2 da resolução da imagem reamostrada.

Na etapa de classificação temática de uso da terra foi aplicada a classificação digital supervisionada, desenvolvida em função do conhecimento prévio de áreas amostrais obtidas no trabalho de campo. Isso permitiu a seleção de áreas de treinamento confiáveis, definindo-se que o algoritmo classificador pelo método de máxima verossimilhança operasse com base na distribuição de probabilidade de cada classe. As classes temáticas foram estabelecidas em função dos temas de interesse: florestas implantadas de *Pinus* sp., *Eucalyptus* sp., florestas nativas (florestas primárias, florestas secundárias, matas de galeria e capoeiras).

A quantificação do uso da terra constituiu em determinar as áreas das diferentes classes obtidas através da função de cálculo de áreas, sendo este efetuado a partir da contagem do número de pixels classificados em cada um dos temas (classes de uso da terra), considerando a resolução espacial de 30 metros.

Como última etapa aferiu-se a acurácia temática dos resultados obtidos na classificação digital dos povoamentos florestais, submetidos a um sistema de grade, proporcionando uma amostragem sistemática. Determinou-se um arquivo raster de 2.500 m x 2.500 m para os povoamentos de florestas implantadas e de 10.000 m x 10.000 m nos povoamentos de florestas nativas, no qual foram alocadas áreas controles em cada um dos pontos de cruzamento das linhas raster. Primeiramente identificou-se através do SIG as coordenadas referenciadas desses pontos, posteriormente inseridas no GPS, que serviram de orientação para sua localização.

As informações de "veracidade de campo" foram inseridas e analisadas dentro dos índices de acurácia descritos por Brites et al. (1996): Exatidão global ($G = [\sum_{i=1}^M n_{ii}] / N$), Kappa ($K = [Po - Pc] / [1 - Pc]$) e Tau ($T = [Po - 1/M] / [1 - 1/M]$), sendo: $\sum_{i=1}^M n_{ii}$ = somatório dos pontos corretamente classificados dentro de cada classe; N = número total de pontos contemplados na matriz; Po = proporção de unidades que concordam plenamente; Pc = proporção de unidades que concordam por casualidade e M = número de categorias na classificação.

3. Resultados e Discussão

No processamento digital das imagens observou-se a faixa de diferenciação da resposta espectral dos temas abordados através do número digital (ND), calculado a partir da redução das componentes RGB para valores de 256 padrões. A classe de florestas nativa com ND entre 150 e 162, de dimensões relativamente reduzidas, está distribuída principalmente ao longo da rede de drenagem e encostas. As classes de florestas implantadas de *Eucalyptus* sp. (ND 158 á 168) e *Pinus* sp. (ND 155 á 167), com dimensões variáveis, com formato regular estão relativamente bem distribuídas na área de estudo, pois as empresas florestais existentes formaram-se principalmente a partir de recursos oriundos dos incentivos fiscais da década de 70 e 80, estabelecendo uma estrutura fundiária dispersa na região. Verificou-se que a apesar de existir diferenças referentes à reflectância das florestas implantadas, essa diferenciação ficou prejudicada pelo ND, pois a região apresenta-se com áreas heterogêneas de

reflorestamento, onde encontra-se talhões de *Pinus* sp. e *Eucalyptus* sp. alternados por vegetação de mata ciliar.

No processo de georreferência, obteve-se 22 pontos de controle, sendo o cálculo do número de linhas e colunas da imagem respectivamente 6541 e 5125, com o erro médio quadrático de 0,002419 unidades de imagem, ou seja, 13,70 metros, cujo valor está dentro do padrão admitido (1/2 da resolução do pixel) de 30 m estabelecido por Eastman (1994).

Na quantificação do uso da terra dos povoamentos florestais, verifica-se (Figura 1) que o maior percentual encontrado foi de florestas nativas, com 5.300,94 km² perfazendo 32,88% da área de estudo. Os povoamentos de floresta de *Pinus* sp. e *Eucalyptus* sp. apresentaram 305,34 Km² e 124,16 Km², representando 1,89% e 0,77%, respectivamente.

A cobertura florestal na área de estudo totalizou 35,54 % e está demonstrado na figura 2. Os demais uso do solo (agricultura irrigada, agricultura de sequeiro, solo exposto, pastagem, corpos d'água e áreas urbanizadas), totalizaram 64,46 % de cobertura da área de estudo.

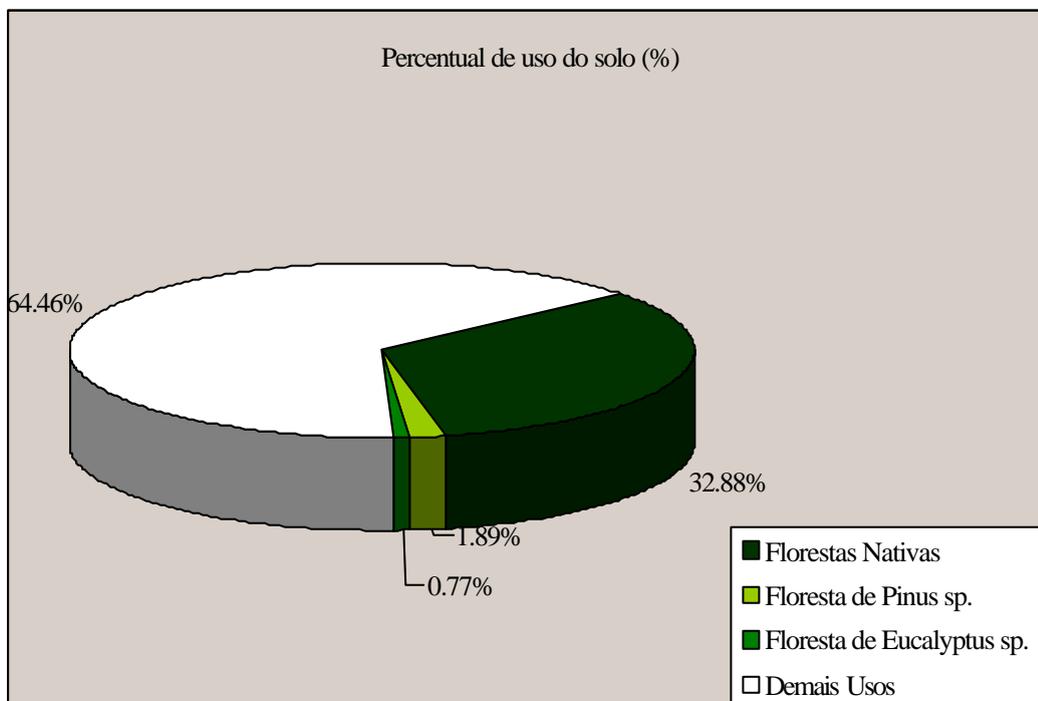


Figura 1: Percentual de uso do solo na área de estudo.

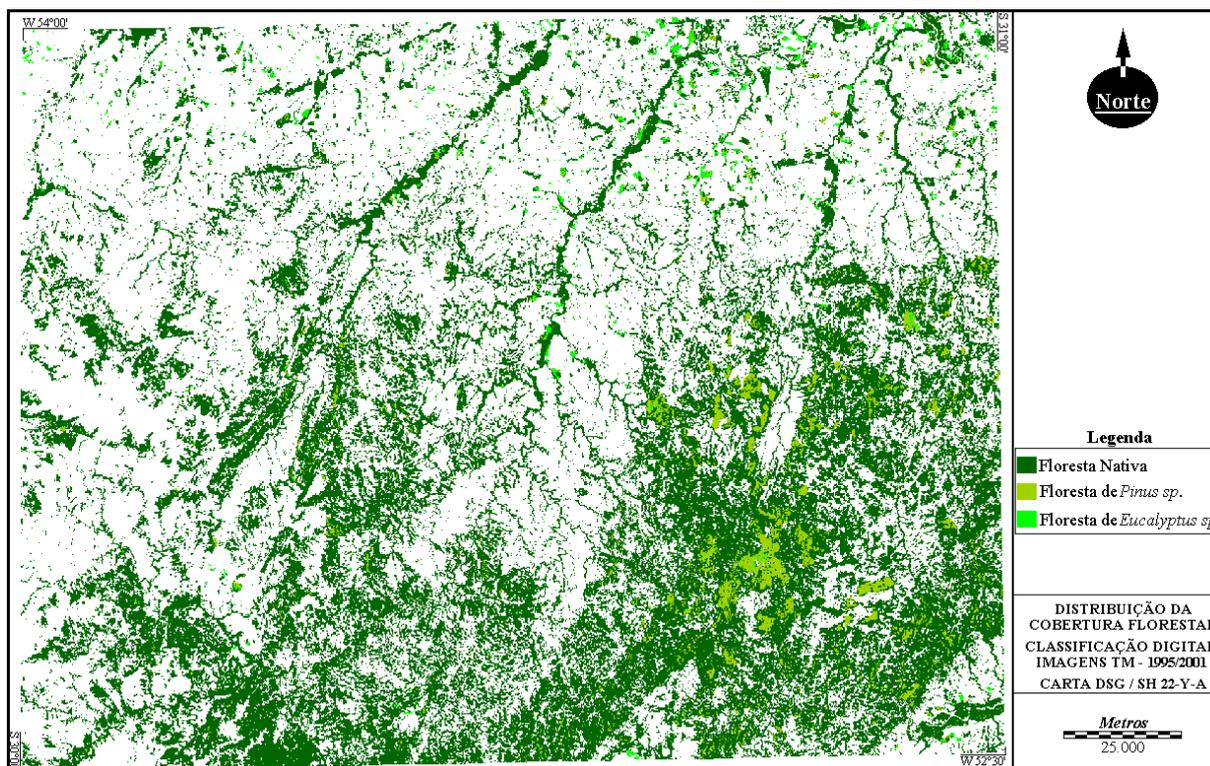


Figura 2. Distribuição da cobertura florestal na área de estudo.

No cruzamento das linhas "raster" obteve-se 45 pontos de controle para as florestas nativas, 27 pontos para a floresta de *Pinus* sp. e 16 pontos para a floresta de *Eucalyptus* sp., totalizando 88 pontos de controle. O resultado do cruzamento dos pontos de controle da imagem referência com a imagem classificada digitalmente está expressa na matriz de erros para a área de estudo, conforme quadro 1.

Quadro 1 : Matriz de erros para análise da acurácia obtida na classificação digital.

Área de controle	Floresta Nativa	Floresta de <i>Pinus</i> sp.	Floresta de <i>Eucalyptus</i> sp.	\sum_{x+i}
Floresta Nativa	43	0	2	45
Floresta de <i>Pinus</i> sp.	6	18	3	27
Floresta de <i>Eucalyptus</i> sp.	0	2	14	16
\sum_{xi+}	49	20	19	75

Pode-se observar na tendência da distribuição dos erros ou confusões na classificação digital das classes temáticas propostas, que a maior confusão ocorreu quando amostrada a classe floresta de *Pinus* sp., onde classificou-se 6 pontos como sendo de floresta nativa e 3 pontos como sendo de floresta de *Eucalyptus* sp.. Ocasionado provavelmente por dois fatores: o primeiro pela presença de mata galeria no interior de povoamentos implantados de *Pinus* sp.; o segundo em função dos limiares da classificação digital, onde as respostas espectrais da floresta nativa (ND 150 a 162) e da floresta de *Eucalyptus* sp. (ND 158 a 168), atingem uma faixa de valores semelhantes à resposta espectral dos povoamentos de *Pinus* sp. (ND 155 a 167).

As classes que não sofreram confusão foram as de floresta de *Pinus* sp. e floresta de *Eucalyptus* sp., quando amostrada a classe floresta nativa e as classes floresta nativa e floresta

de *Pinus* sp. quando amostrada a classe floresta de *Eucalyptus* sp. Obteve-se assim 75 pontos de controle corretamente classificados dos 88 pontos avaliados.

Os valores obtidos da acurácia para três diferentes índices estão apresentados na Figura 3.

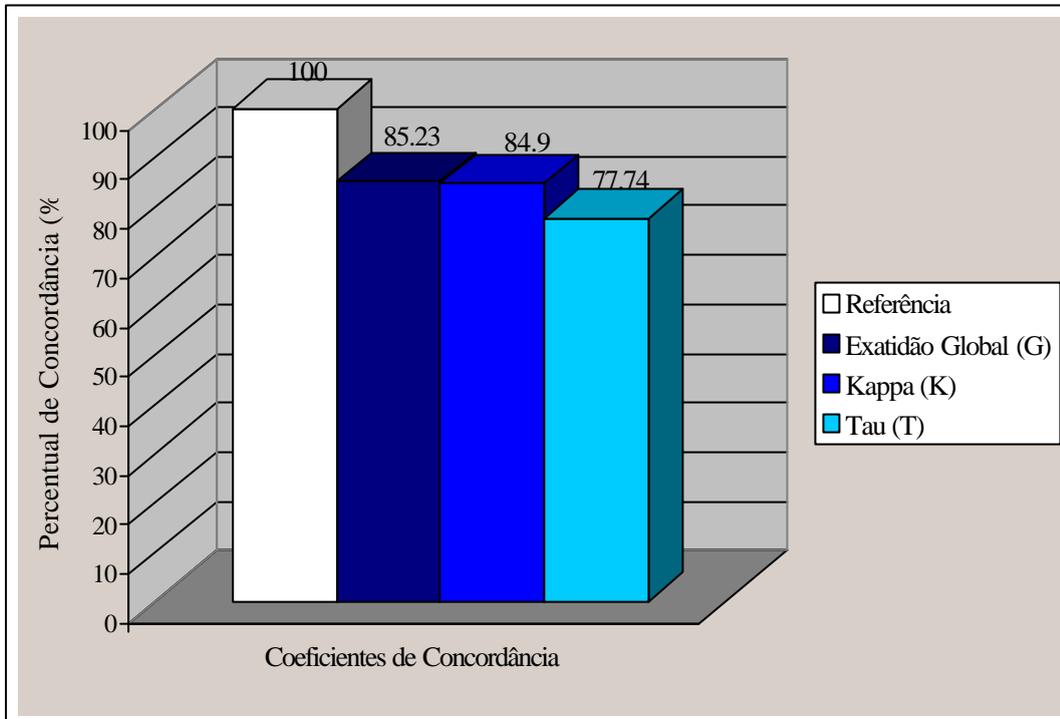


Figura 3 . Índices de acurácia na classificação digital em %.

A Exatidão global da classificação digital, expressa pela razão entre os pontos corretamente classificados e o total de pontos de referência, calculado em 85,23% está dentro do patamar estabelecido por Jensen (1986), de 85%.

O valor do coeficiente Kappa obtido de 84,90 %, menor do que a Exatidão global, se justifica pelo fato do índice Kappa considerar todas as células da matriz de erros; o mesmo ocorre com o índice Tau, com valor de 77,74 %.

4. Conclusão

A partir dos resultados obtidos, chegou-se às seguintes conclusões:

O processo de georreferenciamento utilizando, polinômio cúbico, apresentou precisão dentro da escala trabalhada, sendo de 13,70 m.

O local de estudo possui expressivo índice de cobertura florestal, atingindo 35,54% de cobertura da área.

Os índices de Exatidão global (85,23 %), Kappa (84,90%) e Tau (77,74%), podem ser utilizados na aferição de acurácia da classificação digital.

Embora a Exatidão global apresente um valor mais alto, os índices Kappa e Tau possuem mais consistência por envolver no valor final todas as células da matriz de erros.

A determinação da matriz de erros permite verificar e analisar as confusões ocorridas entre as classes temáticas levantadas.

Referências

- Brites, R.S.; Soares, V.P.; Ribeiro, C.A.A.S. Verificação da exatidão em classificação de uma imagem orbital mediante a utilização de três índices. *Revista Árvore*. Viçosa, v. 20, n. 3, p. 415-424, 1996.
- Eastman, J.R. *Idrisi: exercícios tutorais*. Porto Alegre: UFRGS, 1994. 109 p.
- Foody, G.M. On the compensation for chance agreement in image classification accuracy assessment. *Photogrametric Engineering and Remote Sensing*. Bethesda, v. 58, n. 10, p. 1459-1460, 1992.
- Gong, P.; Howarth, P.J. An assessment of some factors influencing multispectral land-cover classification. *Photogrametric Engineering and Remote Sensing*. Bethesda, v. 56, n. 5, p. 597-603, 1990.
- Jensen, J.R. *Introductory digital image processing*. Englewood Cliffs: Prentice - Hall, 1986. 51p
- Ma, Z.; Redmond, R.L. Tau coefficient for accuracy assessment of classification of remote sensing data. *Photogrametric Engineering and Remote Sensing*. Bethesda, v.61, n.4, p.435-439, 1995.
- Novo, E.M.L. *Sensoriamento remoto, princípios e aplicações*. São Paulo: Blucher, 1992. 308 p.
- Pereira, R.S.; Madruga, P.R. de A.; Hasenack, H. *Geoprocessamento aplicado ao planejamento de uso de recursos naturais*. Santa Maria: UFSM/DER/FATEC, 1995. 40p.
- RadamBrasil. *Levantamento de recursos naturais*. Rio de Janeiro: FIBGE, 1986. p. 313-581.
- Silva, E. M.; Gonçalves, M. E.; Murolo, Carlos, A. *A pesquisa operacional: programação linear, simulação*. 3. ed. São Paulo: Atlas, 1998. 184p.
- Story, M.; Congalton, R.G. Accuracy assessment: a user's perspective. *Photogrametric Engineering and Remote Sensing*. Bethesda, v. 52, n. 3, p. 397-399. 1986.