

ANÁLISE DA CONFIABILIDADE TEMÁTICA DE RESULTADOS OBTIDOS POR MEIO DE CLASSIFICAÇÃO DIGITAL DE IMAGENS LANDSAT-5 TM

LINGNAU, Christel⁽¹⁾; ANTUNES, Alzir Felipe Buffara⁽²⁾

⁽¹⁾Universidade Federal do Paraná.
R. Des. Otávio do Amaral, 716
80.730-400 Curitiba-Pr

⁽²⁾Universidade Federal do Paraná. Centro Integrado de Estudos em Geoprocessamento-CIEG
Caixa Postal 19.076, Centro Politécnico- 81.530-900 Curitiba-Pr
E. Mail: felippe@pc01.cieg.ufpr.br ou felippe@cce.ufpr.br

ABSTRACT

This paper presents different ways of accuracy assessment of remote sensed data for multipurpose cadastral applications. It is based on two kinds of thematic information : Ombrophyla Dense Forest segmentation and *Pinus spp* wood. The error matrices were used as reference to accuracy assessment procedures. The calculation of different index and its methodology were discussed such as: *Kappa*, *Tau*, *mean accuracy index and map accuracy index*.

Keywords: accuracy, coefficient of agreement, types of errors

RESUMO

O presente trabalho apresenta diferentes formas de avaliação da exatidão de informações oriundas de dados de sensoriamento remoto, aplicado ao cadastro multifinalitário. O estudo se baseia em dois mapas temáticos: segmentação da Floresta Ombrófila e reflorestamento de *Pinus spp*, onde diferentes matrizes de erros foram utilizadas como referência para avaliar a acurácia. Esta foi verificada através da discussão de diferentes coeficientes de concordância, tais como: *Kappa*, *Tau*, *índice médio de acurácia e índice de acurácia do mapeamento*, sendo que cada um deles expressa a acurácia obtida de uma perspectiva metodológica diferente.

Palavras-chave: acurácia, coeficiente de concordância, tipos de erro

1. INTRODUÇÃO

A conservação, preservação, planos de manejo sustentado e planejamentos em florestas plantadas requerem um mapeamento da situação atual da área em questão. Os mapas temáticos então disponíveis são tidos geralmente como corretos. No entanto, sabe-se que erros desde a obtenção dos dados até a geração de um mapa podem ocorrer em diferentes níveis e escalas.

Grande parte do mapeamento digital de cobertura vegetal, a nível regional, vem sendo realizado através de imagens de sensoriamento remoto. Técnicas de classificação supervisionada associadas ao processamento digital de imagens tem dado bons resultados na extração de informação que compõem determinada região. No entanto todo o processo de classificação está sujeito a erros dos mais diversos tipos e fontes. A avaliação da acurácia dos dados extraídos das imagens está diretamente relacionado com a qualidade do mapeamento. Todas as classes que compõem a imagem ou o mapa temático devem ser associadas a um controle de acurácia para que

estas possam ser validadas. Entende-se que um mapeamento temático só deva ser utilizado como fonte fidedigna de informações quando associados a processos de avaliação da acurácia dos dados que o compõem. Observa-se que raramente os mapas provenientes de classificação digital ou mesmo da fotointerpretação apresentam um relatório de precisão das informações, apesar da maioria dos *softwares* de processamento digital de imagem possuírem as ferramentas de cálculo.

2. OBJETIVOS

Este artigo objetiva expor ao usuário as mais diferentes formas de se efetuar um avaliação da classificação digital de uma imagem. A intenção é apresentar um resumo das principais técnicas e os fatores que devam ser levados em conta no processo de verificação da qualidade dos dados temáticos.

3. ÁREA DE ESTUDO

A avaliação da acurácia de mapas temáticos obtidos a partir da classificação digital de imagens de satélite Landsat-5 TM depende do tipo de informação existente que possa ser utilizada como referência. Aqui dois casos podem ser definidos:

- existem fotografias aéreas;
- As categorias são homogêneas, ou seja, um grupo de elementos de imagem pertencem a mesma categoria ou ainda a assinatura espectral representada por um elemento de imagem refere-se a categoria;
- Fácil acesso às áreas de controle;
- Difícil acesso às áreas de controle.

Em função do exposto acima, para abranger a maior variabilidade possível de tipos de problemas que possam surgir na avaliação da acurácia de mapas temáticos, este trabalho apresenta duas áreas de estudo Floresta Ombrófila e Reflorestamentos de *Pinus spp*— com características bem distintas.

A Floresta Ombrófila situa-se na Área de Proteção Ambiental de Guaraqueçaba, Litoral do Paraná, abrangendo uma superfície de aproximadamente 121.131 ha e 143 km de perímetro, incluindo o Parque Nacional do Superagui, as Ilhas do Pinheiro, do Pinheirinho, Rasa e Pavoça, a porção continental que abrange a sede do município de Guaraqueçaba e as Serras do Tromomô e da Fazenda.

Os reflorestamentos de *Pinus* situam-se no município de Bocaina do Sul- SC., entre as coordenadas de 49°00' e 50°00' de longitude e 27°20' a 28°10' de latitude sul. O relevo é plano a levemente ondulado. Os povoamentos florestais de pinus apresentam limites bem definidos, o mesmo tratamento silvicultural e são equiâneos. Trata-se neste caso de um objeto de estudo menos complexo do que a floresta ombrófila.

4. MATERIAL - MAPAS TEMÁTICOS

A avaliação da acurácia será realizada com base em dois mapas temáticos obtidos a partir da classificação supervisionada pelo método da máxima verossimilhança. Para a área de Guaraqueçaba foram definidas 6 classes de vegetação (TAB. 1), sendo que estas classificadas com base nas informações extraídas das bandas 2, 3, 4 e 5 do sensor Thematic Mapper. Uma pós-classificação foi efetuada com base no Modelo Digital do Terreno, onde a imagem foi reclassificada de acordo com as altitudes (ANTUNES 1996). Usaram-se os seguintes intervalos de altitude (RODERJAN & KUNYOSHI 1988; IBGE 1992):

- [0, 50m], Floresta Ombrófila Densa Terras Baixas;
- [50, 600m], Floresta Ombrófila Densa Submontana;
- [600, 733m], Floresta Ombrófila Densa Montana.

TAB. 1: Classes de Vegetação para a Área de Proteção Ambiental de Guaraqueçaba-Pr.

Classes de Vegetação	Descrição
1. F.O.D. Submontana	Áreas de Floresta Ombrófila com altitudes superiores a 50 metros e em fase sucessional avançada;
2. Formação Pioneira de Influência Fluvial	Áreas situadas nos intercordões litorâneos sobre solos hidromórficos;
3. Formação Pioneira de Influência Fluvio-Marinha	Manguezais nas desembocaduras dos rios e interior da baía;
4. Formação Pioneira de Influência Marinha	Restingas, vegetação próxima as praias e dos cordões litorâneos;
5. F.O.D das Terras Baixas	Floresta aluvial situadas nas zonas de transição entre as formações pioneiras e a floresta submontana;
6. Vegetação Alterada	Áreas de capoeira, capoeirinha, pastagens.

Os povoamentos florestais de pinus foram classificados em relação às fases de desenvolvimento. A definição de seis classes de idade (TAB. 2) foi possível a partir da técnica de transformação de dados, a Tasseled Cap, sendo que na classificação foram usados a Brightness, Greenness e a Wetness (LINGNAU, 1995).

Ambas as imagens temáticas geradas apresentaram o efeito de "salt and peppers", ou seja, pixels ou pequenos agrupamentos isolados criando ruídos na imagem. A eliminação desses ruídos foi alcançado utilizando-se um filtro de moda (3x3), o qual possibilitou agregar os pequenos agrupamentos à classe mais próxima do valor majoritário da janela. Este processo generaliza a imagem excluindo informações desnecessárias.

TAB. 2: Fases de desenvolvimento dos povoamentos florestais de *Pinus* spp. na área de estudo Lages-SC. (LINGNAU, 1995).

Classes de Idade	Descrição
1. Corte raso	Áreas com solo exposto e resíduos (galhos e acículas secas e verdes);
2. 1 a 3 anos	Plantio e tratamento silvicultural (coroamento, roçada e replantios);
3. 4 a 6 anos	Fase em que ocorre o fechamento de copa;
4. 7 a 10 anos	Povoamento fechado, seca dos galhos inferiores e atividade de crescimento alta;
5. 11 a 15 anos	Povoamento com pouca ou nenhuma abertura de copa; atividade de crescimento média;
6. 16 a 25 anos	Povoamento com dossel aberto a pouco aberto; atividade de crescimento baixa.

5. PROCEDIMENTOS METODOLÓGICOS

5.1 Tipos de Erros

Qualquer tipo de medição, mesmo apurada, está sujeita a erros. Segundo THAPA et al. (1992) e GEMAEL (1994) os erros podem ser três tipos:

- Erros grosseiros são causados pela falta de atenção do observador, erros na leitura de valores nos equipamentos ou escalas. Estes tipos de erros podem ser detectados e eliminados imediatamente após as medições através de leituras múltiplas;
- Erros sistemáticos ocorrem em inventários de campo e fotogrametria devidos às condições do meio, imperfeições nos equipamentos e limitações humanas. Na análise estatística o erro sistemático introduz o *bias*, ou seja, é observada uma tendência nos valores estimados (medidos);
- Removendo os erros grosseiros e sistemáticos restam ainda os erros que ocorrem ao acaso, os quais são de magnitude pequena. Os erros aleatórios ou acidentais, ao contrário dos anteriores ocorrem ora num ora noutro sentido e não podem ser vinculadas a nenhuma causa específica. Apresentam as seguintes características:
 - erros positivos e negativos ocorrem na mesma frequência;
 - erros menores são mais frequentes do que erros maiores;
 - apresentam uma distribuição normal.

Os três tipos de erros citados acima podem ocorrer a nível amostral ou não. O erro amostral é ocasionado por apenas uma parte da população amostrada e é expresso pelo erro padrão. Se erros amostrais não ocorrem, então, o erro padrão seria a medida adequada para expressar o grau de precisão do inventário.

5.2 Precisão e Acurácia

A discussão sobre a análise da confiabilidade temática da imagem requer que inicialmente sejam definidos os termos acurácia, precisão e exatidão, os quais geralmente são usados de forma confusa. HUSCH et al. (1982) definem precisão como o grau de concordância de uma série de observações ou medidas, enquanto a acurácia esta relacionada a proximidade de uma observação ou medida do seu valor real. Pode-se, então, associar a precisão os erros aleatórios ou acidentais, e acurácia aos efeitos dos erros aleatórios e sistemáticos conjuntamente (GEMAEL, 1994). Uma observação pode ser considerada como exata se não tiver ocorrido o erro sistemático ou *bias*. A falta de exatidão das observações são geralmente oriundas dos erros sistemáticos e tendem a se acumular num mesmo sentido.

Considerando uma amostragem, a acurácia seria expressa pelo desvio padrão da amostra do valor real da população. A precisão poderia ser relacionada ao desvio padrão da média da amostra. É possível que uma amostragem possua uma alta precisão, ou seja em pequeno desvio em relação a média, e ao mesmo tempo, devido a falta de exatidão, um grau acurácia inferior.

5.3 Análise das Possíveis Fontes de Erros

Erros podem ocorrer em diferentes níveis, ou seja, desde a aquisição de dados, processamento até na verificação dos resultados obtidos. Estes diferentes tipos de erros muitas vezes não são distinguíveis mensuráveis. Na obtenção dos dados a partir de imagens de satélite os erros podem ser primários ou secundários. Erros primários são ocasionados devido a sensor, condições do meio (variações de temperatura, umidade, iluminação) e ainda na interpretação da imagem. Existem, também os erros secundários, que são introduzidos através do processamento dos dados e sua reprodução (THAPA et al. (1992).

- No processamento podem ocorrer erros na correção geométrica, sendo estes ainda podem ser devido a má definição dos pontos de controle e/ou no algoritmo de interpolação dos elementos de imagem;
- Áreas de treinamento mal definidas: caso as áreas de treinamento não forem homogêneas e um mínimo de variância entre os níveis de cinza for obtido, na classificação os elementos de imagem poderão ser agrupados erroneamente;
- Elementos de imagem mistos são aqueles que representam mais de uma categoria. Ocorrem geralmente em bordaduras das classes;
- Sobreposição de classes no espaço de atributos, daqueles que foram escolhidos de forma representativa;
- Erros na classificação em função das condições topográficas e atmosféricas.

5.4 Acurácia do Dado digital

Em qualquer mapa temático obtido através da classificação digital pode ser efetuada a análise da acurácia referente as seguintes componentes (MONTGOMERY & SCHUCH, 1993):

- Posicional, que se refere a precisão espacial do mapa e que é variável de acordo com a escala do mapa. Por exemplo, mapas na escala 1:100.000, possuem um precisão espacial média de 50 metros considerando as mais diversas fontes de erros;
- Temático, refere-se a precisão da classificação das diferentes categorias que compõem o mapa. Estas devem de acordo com a escala e o grau de generalização da informação refletir no mapa a mesma classe no terreno.
- A Área Total obtida por categoria também é uma componente importante, no entanto não muito significativa.

As componentes citadas são interdependentes, no entanto a avaliação é geralmente realizada de forma independente, considerando se a precisão posicional verificada é compatível com a escala do levantamento e com a retificação da imagem.

Neste artigo serão tratados somente os procedimentos de avaliação temática proveniente da classificação digital. Isto não impede que os mesmos procedimentos possam ser utilizados na avaliação de outros mapas temáticos.

5.5 Pontos de Referência

A fim de avaliar a acurácia da classificação para a área de Guaraqueçaba foram selecionados pontos de referência de forma aleatória e estatigraficamente na imagem classificada. Estes pontos foram selecionados no centro de um raio de 10 pixels de mesmo valor, evitando assim amostras de um pixel.

Geralmente, a verificação é realizada em campo, no entanto devido a inviabilidade de acesso à área de estudo para identificar 377 pontos selecionados aleatoriamente a verificação foi efetuada por fotointerpretação da imagem original apoiada por controle de campo. Além disso, a defazagem temporal que existia entre a tomada das imagens, julho de 1993, e a data de verificação de campo prejudicaria a análise de algumas classes. Contudo, antes de verificar todos os pontos amostrados foi determinada a acurácia da fotointerpretação para algumas classes de vegetação, através de 47 pontos de referências amostrados, em campo, meio de GPS e mapas topográficos.

HORD et al. (1976) e RICHARDS (1994) apresentam a expressão para o cálculo do intervalo de confiança para uma amostra n

$$p\{-Z_{\alpha/2} < x - n\theta / \sqrt{n\theta(1-\theta)} < Z_{\alpha/2}\} = 1 - \alpha$$

onde:

X é o número de *pixels* corretamente identificados;

n é o tamanho da amostra;

θ é a acurácia total da identificação x/n ;

$1-\alpha$ é o limite de confiança.

De acordo com a expressão acima descrita, determinou-se a acurácia da fotointerpretação de uma amostra de 47 pontos coletadas em campo por GPS (PYE et al 1994). As amostras foram coletadas de forma aleatória para um conjunto de categorias (mangues, solo exposto, vegetação alterada, sub-montana, floresta aluvial e restingas).

Visto que a amostragem possuía mais de 30 pontos considerou-se a distribuição normal dentro de uma certeza 95%. Como resultado, a acurácia total da fotointerpretação ficou compreendida no intervalo de 84,7% a 91,75%, com uma probabilidade de 95% (ANTUNES 1996).

A verificação dos reflorestamentos de pinus é consideravelmente mais simples do que na floresta ombrófila. O acesso é mais fácil, assim como através da definição clara dos limites a localização das áreas de controle ocorre se problemas. Em função destas características optou-se pela verificação sistemática, onde foi determinado um raster de 700 x 700 m (FIG. 1a). Em cada ponto do cruzamento de linha no raster foi alocada uma área de controle ou referência representada por 13 elementos de imagem (FIG. 1b). Em função dos limites bem distintos as áreas de controle puderam ser facilmente localizadas em campo. Em campo a área de controle seria o talhão por ser uma unidade homogênea em termos de idade.

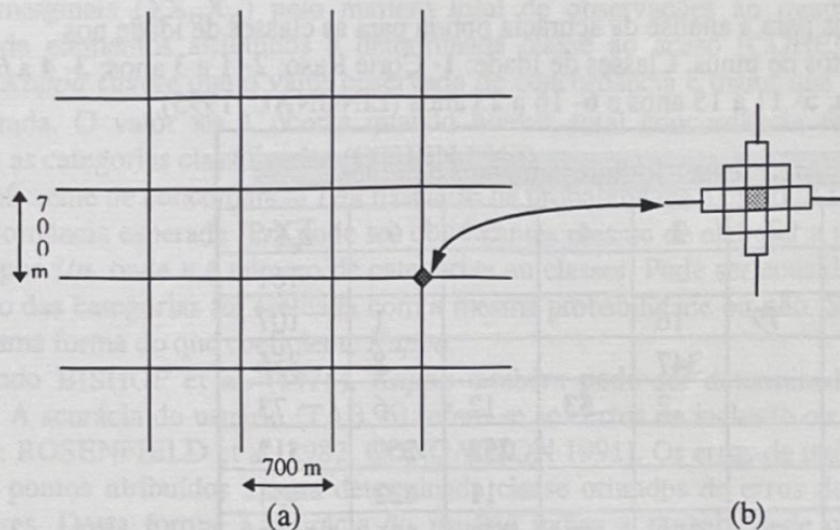


FIG. 1: Apresentação esquemática: (a) Raster e (b) Área de referência (LINGNAU, 1995).

Para avaliar a acurácia fez-se necessário a elaboração de imagem referência, a qual representasse todas as áreas de controle, para que possa ser comparada então com a imagem classificada. É possível que uma área de controle represente duas ou mais classes e neste caso seria uma área de controle impura e levaria a uma distorção da concordância. Para contornar este tipo de erro, a área de controle teria que ser então deslocada para o talhão mais próximo. A acurácia não foi por isso influenciada arbitrariamente, mas erros evidentes são eliminados.

5.6 Matriz de Erros

A matriz de erros ou confusão apresenta de forma resumida os resultados da classificação digital ou fotointerpretação em relação à imagem referência, onde teremos então uma distribuição binomial. Cada frequência observada na matriz corresponderá também a uma frequência esperada, a qual é calculada sob uma determinada hipótese de acordo com as regras da probabilidade (SPIEGEL, 1993). A frequência observada na diagonal (X_{ii}) apresenta a concordância entre o esperado e observado em cada categoria, ou seja, aqueles corretamente classificados. A coluna marginal é o número total de elementos de imagem para cada classe (X_{i+}). A linha marginal o número de pontos atribuídos a cada classe (X_{+i}).

Os resultados da combinação das imagens classificadas com as imagens referência está expressa nas matrizes de erros para ambas as áreas de estudo (TAB. 3 e TAB. 4)

TAB. 3: Matriz de Erros para Floresta Ombrófila Densa (FDO) e Formações Pioneiras: 1- FOD Submontana; 2- FPI Fluvial; 3- FPI Fluvio-Marinha; 4- FPI. Marinha; 5- FOD das Terras Baixas; 6- Vegetação Alterada, (Fonte: ANTUNES 1996).

Área de Controle	Área Teste - Floresta Ombrófila						ΣX_{+i}
	1	2	3	4	5	6	
1	80				1	2	83
2		59	5	1	6		71
3		1	51	3			55
4			1	47	2	2	52
5		1	2	1	55	1	60
6				1	2	53	56
ΣX_{+i}	80	61	59	53	66	58	377

TAB. 4: Matriz de erros para a análise da acurácia obtida para as classes de idade nos reflorestamentos de pinus. Classes de Idade: 1- Corte Raso; 2- 1 a 3 anos; 3- 4 a 6 anos; 4- 7 a 10 anos; 5- 11 a 15 anos e 6- 16 a 25 anos (LINGNAU, 1995).

Área de Controle	Área Teste - Reflorestamentos de Pinus						ΣX_{+i}
	1	2	3	4	5	6	
1	161						161
2	11	79	16			1	107
3	1		347			8	356
4			2	53	12	6	73
5					253	259	512
6			1		11	422	434
ΣX_{+i}	173	79	366	53	276	696	1643

5.7 Coeficientes de Concordância

A avaliação da acurácia pode ser obtida por meio de coeficientes de concordância, sendo que estes podem ser expressos para o mapa ou para classes individuais. Alguns dos coeficientes de concordância comumente usados para avaliar a acurácia total do mapa temático estão apresentados na TAB. 5.

COHEN (1960) define *Kappa* como um coeficiente de concordância para escala nominais que pede a proporção de concordância depois que a concordância atribuída a casualidade é retirada de consideração (TAB. 5). O coeficiente *Kappa* considera todos os elementos da matriz de erros ao invés de apenas aqueles que se situam na diagonal principal da mesma, ou seja, estima a soma da coluna e linha marginais. Segundo COHEN (1960) o coeficiente *Kappa* mede o grau de concordância em escalas nominais assumindo que :

- As unidades são independentes;
- As classes ou categorias da escala nominal são independentes e mutuamente exclusivas;
- O classificador e os pontos de referência operam de forma independente.

TAB. 5: Coeficientes de Concordância para a avaliação da acurácia total.

Acuracia	Fórmula	Interpretação
Concordância Total (HELLDEN, et al., 1980)	$P_o = \frac{\sum X_{ii}}{N}$	$0 < P_o < +1$
Coeficiente de concordância Kappa (COHEN, 1960)	$\kappa = \frac{P_o - P_e}{1 - P_e}$	$-1 < \kappa < +1$
Coeficiente de Concordância Tau (KLECKA, 1980)	$\tau = \frac{P_o - P_r}{1 - P_r}$	$-1 < \tau < +1$

P_o para toda a matriz de erros é dado pela razão do somatório da diagonal principal ($\sum X_{ii}$) pelo número total de observações (N). P_o pode ser traduzido pela proporção de unidades (pontos de referência) que concordam. P_e é expresso pelo somatório do produto dos elementos das linhas e colunas marginais ($\sum X_{i+} X_{+i}$) pelo número total de observações ao quadrado. P_e traduz a proporção de elementos atribuídos a determinada classe ao acaso (COHEN 1960). O valor positivo de Kappa sugere que o valor observado de concordância é maior que a concordância ao acaso esperada. O valor $\kappa = 1$ ocorre quando houver total concordância entre os pontos de referência e as categorias classificadas (COHEN 1960).

O coeficiente de concordância Tau baseia-se na probabilidade a priori (KLECKA, 1980), ou seja, a concordância esperada (P_r) pode ser obtida antes mesmo de elaborar a matriz de erros. P_r é expresso por $1/n$, onde n é número de categorias ou classes. Pode ser considerado, ainda, se a classificação das categorias foi efetuada com a mesma probabilidade ou não. Sua interpretação é feita da mesma forma do que coeficiente Kappa.

Segundo BISHOP et al. (1975), Kappa também pode ser determinado para categorias individuais. A acurácia do usuário (TAB. 6) refere-se aos erros de inclusão ou comissão (HORD et al., 1976; ROSENFELD et al. 1982; CONGALTON 1991). Os erros de inclusão são relativos a todos os pontos atribuídos a uma determinada classe oriundos de erros de classificação das outras classes. Desta forma, a acurácia do usuário indica a probabilidade que uma categoria classificada (mapa/imagem) representasse a mesma categoria no campo. A acurácia do produtor (TAB. 6) leva em consideração os erros de exclusão ou omissão, onde não são atribuídos a determinada classe os erros de classificação das outras classes. Da mesma forma, a acurácia do produtor referiu-se a probabilidade de determinada categoria ter sido corretamente classificada de acordo com os pontos de referência (ROSENFELD et al. 1986; STORY et al. 1986).

O índice de acurácia média (HELLDEN et al., 1980) (TAB. 6) apresenta a concordância entre um ponto no mapa e no terreno. Trata-se de um índice lógico, ou seja, não é desenvolvido em bases probabilísticas ou matemáticas (ROSENFELD et al. 1986). O índice de acurácia de mapa (SHORT, 1982) (TAB. 6) é similar à equação de JACCARD (1908), o qual é baseado em uma distribuição hipergeométrica, sendo que a distribuição binomial é uma aproximação para amostras grandes.

O coeficiente de concordância Kappa para as classes individuais é calculado através do mesmo princípio do que para acurácia total. o mesmo vale para a interpretação dos resultados.

TAB. 6: Coeficientes de Concordância para avaliação da acurácia para classes individuais.

Acurácia	Fórmula	Interpretação
Coeficiente de Concordância Kappa (LIGHT, 1971 e BISHOP et al., 1975)	$\kappa_i = \frac{N X_{ii} - X_{i+} X_{+i}}{N X_{i+} - X_{i+} X_{+i}}$	$-1 < \kappa > +1$
Índice Médio de Acurácia (HELLDEN et al., 1980)	$\mu a_i = \frac{2 X_{ii}}{X_{i+} + X_{+i}}$	$0 < \mu a > +1$
Índice de Acurácia do Mapeamento (SHORT, 1982)	$\mu p_i = \frac{X_{ii}}{X_{i+} + X_{+i} - X_{ii}}$	$0 < \mu p > +1$
Acurácia do Produtor	$A p_i = \frac{X_{ii}}{X_{i+}}$	$0 < A p > +1$
Acurácia do Usuário	$A u_i = \frac{X_{ii}}{X_{+i}}$	$0 < A u > +1$

A avaliação da acurácia através de diferentes tipos de coeficientes de concordância permitem obtenção de um grau de confiabilidade das categorias classificadas na imagem, sob diferentes perspectivas e metodologias.

6. RESULTADOS E DISCUSSÃO

Com base na matriz de erros (Tab. 3 e TAB. 4) pode então ser calculada o índice de acurácia total. Na TAB. 7 estão apresentados os índices de acurácia através de três métodos diferentes.

TAB. 7: Índices de acurácia total para as áreas de estudo de Floresta Ombrófila e Reflorestamentos de pinus.

Área de Estudo	Coeficiente de Concordância (%)		
	Concordância Total	Kappa (κ)	Tau (τ)
Floresta Ombrófila	91,51	89,83	89,80
Reflorestamentos de <i>Pinus</i> spp.	80,04	74,16	76,04

A acurácia total do mapa expressa pela razão entre total de pontos de referência e os pontos corretamente classificados ($\sum X_{ii}/N$) foi de 91,51%. O Kappa total para o mapa foi 89,83% menor que a acurácia Total, pois o seu cálculo levou em conta todos os elementos da matriz de erros. Como pode ser observado na TAB. 7 os valores da acurácia total obtida por quatro coeficientes diferentes não apresentam diferenças significativas entre os valores. No entanto, o coeficiente de

concordância *Kappa* e *Tau* são mais consistentes por envolver no valor final todas as células da matriz de erros. Como o coeficiente de concordância *Tau* e *Kappa* apresentam valores aproximados ambos podem ser recomendados na avaliação da acurácia de mapas temáticos, não obstante de ser *Tau* o mais simples de se determinar.

A avaliação da confiabilidade temática a nível de classes individuais está apresentada nas TAB. 8 e TAB. 9 para as áreas de Floresta Ombrófila e Reflorestamentos respectivamente. os resultados obtidos podem ser melhor interpretados se forem comparados com a matriz de erros.

TAB. 8: Índices de acurácia: 1- FOD Submontana; 2- FPI Fluvial; 3- FPI Fluvio-Marinha; 4- FPI. Marinha; 5- FOD das Terras Baixas e 6- Vegetação alterada, (Fonte: ANTUNES 1996).

CLASSES	A_p , %	A_u , %	μ_{ui} %	μ_{pi} %	κ_i %
1	100,00	96,38	98,16	96,38	95,41
2	96,72	83,10	89,34	80,82	79,83
3	86,44	92,72	89,37	80,95	91,38
4	88,68	90,38	87,85	79,66	88,81
5	83,33	91,67	87,30	77,46	89,90
6	91,38	94,64	92,98	86,66	93,67

Tomando como exemplo a classe 4, Formação Pioneira de Influência Marinha ou *restingas* (TAB. 8), observou-se um *Kappa* de 88,81%. Isto significa que 88,81% da amostragem estavam em concordância com os pontos de referência. Considerando os índices de acurácia do usuário e produtor, verificou-se que 88,68 % das áreas de *restingas* foram corretamente identificadas como *restingas*. No entanto, 90,38% das áreas denominadas *restingas* representaram *restingas* no mapa. Uma análise mais cuidadosa na matriz de erros permitiu avaliar uma certa confusão em discriminar *restingas* de terras baixas e vegetação alterada. Dentre os índices apresentados (A_u , usuário e A_p produtor) o coeficiente *Kappa* de concordância pareceu ser o mais apropriado na verificação da acurácia das categorias individuais (STORY et al. 1986; CONGALTON 1991). Levando em consideração que a fotointerpretação apresentou um erro médio de 11,75%, o índice *Kappa* total (κ) seria reduzido a 79,27% (ANTUNES 1996).

TAB. 9: Índice de acurácia para classes de idade nos reflorestamentos de pinus.

Classe de Idade	A_p , %	A_u , %	μ_{ui} %	μ_{pi} %	κ_i %
1. Corte Raso	93.06	100.00	96.41	93.06	92.31
2.1 a 3 anos	100.00	73.83	84.95	73.83	100.00
3.4 a 6 anos	94.81	97.47	96.12	92.53	93.37
4.7 a 10 anos	100.00	72.60	84.13	72.60	100.00
5.11 a 15 anos	91.67	49.41	64.21	47.29	87.89
6.16 a 25 anos	60.63	97.23	74.69	59.60	46.50

Os índices de acurácia para classes de idade (TAB. 9) mostram que para cinco índices as classes 1 a 4 apresentam valores altos, onde conclui-se que foi obtida uma concordância relativamente alta. Erros de omissão e comissão quase não foram observados (TAB. 4). Assim como pode ser observado na TAB. 4 uma confusão maior ocorre entre as classes 5 e 6, qual ocorreu devido a semelhança na assinatura espectral entre as mesmas. Os cinco diferentes coeficientes de concordância expressam a confusão ocorrida, no entanto não de que forma. O valor de 49,41% de A_{u5} mostra que houve um erro de comissão significativo, mas através dos índices de μ_{u5} e μ_{p5} não é possível identificar que tipo de erro (omissão ou comissão) pode ter

ocorrido. O valor de *Kappa* para K_6 de 46,50% indica que devido a alta proporção desta classe no mapa temático a concordância observada foi baixa em relação a esperada.

7. CONCLUSÃO

1. Inventário ou a classificação digital possuem um valor muito limitado se a acurácia ou a precisão não forem apresentadas ao usuário. Aqui vale lembrar que não basta simplesmente apresentar um valor de acurácia, mas sim também os métodos e parâmetros utilizados para avaliar a acurácia;
2. Erros sistemáticos não podem ser mensurados e nem quantificados através do erro padrão, desta forma devem ser evitados ao máximo em todas as fases do desenvolvimento de dados;
3. As características da área de estudo são relevantes na definição do método de verificação. O acesso a área ou até mesmo a complexidade entre as classes a serem definidas na classificação são fatores importantes que determinam os dados a serem coletados como referência;
4. A apresentação da matriz de erros para o usuário é fundamental, pois através desta é possível visualizar as confusões ocorridas entre as classes;
5. Conforme a finalidade do mapa, a acurácia total pode ser suficiente, no entanto, quando algumas classes específicas são importantes ou interessam mais do que as outras, é importante que o usuário tenha a acurácia por classe individual;
6. Quando a acurácia por classes individuais for apresentada, então deve constar também o método pela qual foi obtida;
7. O coeficiente de acurácia *Kappa* para classes individuais, parece ser o mais indicado na avaliação da classificação de imagens de sensoriamento remoto, por levar em conta os erros *comissão* e *omissão* da matriz de erros no seu cálculo.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- ANTUNES, A. F. B. **Uso do Processamento Digital de Imagens na Segmentação de Tipologias Vegetais na APA de Guaraqueçaba-PR.** Dissertação de Mestrado, Curitiba: Curso de Pós-graduação em Engenharia Florestal, Universidade Federal do Paraná. No prelo.
- BISHOP, Y. M.; FEINBERG, S. E. & HOLLAND, P. W. **Discrete Multivariate Analysis - Theory and Practice.** Cambridge, Mass., the mit press, 357 p., 1975.
- COHEN, J. **A Coefficient of Agreement for Nominal Scales.** Educational and Measurement. Vol XX, Nº 1, p. 37-46, 1960.
- CONGALTON, R. **A Review of Assessing the Accuracy of Classification Remotely Sensed Data.** Remote Sensing of Environment, USA, Nº 37, p. 35-36, 1991.
- GEMEAL, C. **Introdução ao Ajustamento de Observações- Aplicações Geodésicas.** Curitiba: Editora da UFPR. 1994.
- HELLDEN, U. & STERN, M. **Evaluation of Landsat Imagery and Digital Data for Monitoring desertification Indicators in Tunisia.** Proc. 14th. Int. Symp. on Rem. Sens. of Environ., p. 1601-1611, 1980.
- HORD, R. M. & BROONER, W. **Land Use Map-Accuracy Criteria.** Photogrammetric Engineering and Remote Sensing, Vol. 42 (5), p. 671-677, 1976.
- IBGE. **Manuel Técnico da Vegetação Brasileira.** Série Manuais Técnicos, Nº 1. 1992.
- JACCARD, P. **Nouvelle Recherches sur la Distribution Florale.** Bull. Soc. Vaud. Sci., Nat., Vol. 44, p. 223-270, 1908.

- KLECKA, W. R. **Discriminant Analysis**. SAGE Publications, Inc., Beverly Hills, California, 71 p., 1980.
- LINGNAU, C. **Zustandserfassung von Kiefernplantagen in Südbrasilien mit Landsat-5 TM Daten als Grundlage für die Überwachung und Abschätzung künftiger Entwicklung von Holzwespenschäden**. Dissertação de Doutorado. Albert Ludwigs-Universität, Freiburg - Germany, 158 p., 1995.
- MONTGOMERY, G.; SCHUCH, C. **Gis Data Conversion Handbook**. Fort Collins: Gis Word Books. 1993.
- PYE, J.; HENSHE W. & EVANS K. **Control Points for Georegistering Images of the APA Guaraqueçaba**. Relatório interno, U.S.D.A. Forest Service, North Carolina, USA. 1994.
- RODERJAN C. & KUNIYOSHI, Y. **Macrozoneamento Florístico da Área de Proteção Ambiental APA-Guaraqueçaba**. Curitiba: FUPEF,UFPR,p., 1988.
- ROSENFELD, G. H; FITZPATRICK-LINS, K. & LING, S. H. **Sampling for Thematic Map Accuracy Testing**. Photogrammetric Engineering and Remote Sensing, Vol. 48 (1), p. 131-137. 1982.
- ROSENFELD, G. H. & FITZPATRICK-LINS, K. **A Coefficient of Agreement as Measure of Thematic Classification Accuracy**. Photogrammetric Engineering and Remote Sensing, Vol. 52 (2), p. 223-227, 1986.
- SPIEGEL, M. R. **Estatística**. McGrawhill, 3ª edição, 1993.
- STORY, M & CONGALTON, R. **Accuracy assessment: a user's perspective**. Photogrammetric engineering and remote sensing, vol 52, nº3. p379-399. 1986
- THAPA, K. & BOSSLER, D. J. **Accuracy of Spatial Data Used in Geographic Information Systems**. Photogrammetric Engineering and Remote Sensing, Vol. 58 (6), p. 835-841, 1992.