

Comparação de Desempenho entre Três Índices de Exatidão Aplicados a Classificações de Imagens Orbitais

Ricardo Seixas Brites¹
Vicente Paulo Soares¹
Carlos Antonio Álvares Soares Ribeiro¹

¹Universidade Federal de Viçosa / Departamento de Engenharia Florestal
Av. P. H. Rolfs s/n, 36571.000 Viçosa, Minas Gerais, Brasil
{brites, vicente}@brufv.bitnet.br

Abstract. In this paper three indexes, Overall Accuracy, Kappa and Tau, are compared in order to verify their sensitivity to differences among classifications based on the same classifier, same training samples set and different combinations of spectral channels, special transformations and sampling schemes..

Keywords: Remote Sensing, Overall Accuracy Index, Kappa, Tau.

1. Introdução

O teste estatístico mais simples para se julgar a exatidão dos mapas derivados de dados de sensores orbitais consiste em determinar um número ótimo de amostras a serem tomados no terreno e assumir uma margem de erro permissível. Então, pela comparação dos pontos amostrados no terreno com aqueles correspondentes obtidos no mapa gerado pelo processo de classificação, o número de erros pode ser determinado. Se o número de erros é menor ou igual à margem assumida, o mapa é aceito como exato dentro dos intervalos de confiança especificados (CONGALTON, 1984).

De acordo com CAMPBELL (1987), os erros estão presentes em qualquer tipo de classificação. Tais erros, quando decorrentes de processamento automatizado, são resultantes de interações complexas entre as estruturas espaciais da paisagem, a resolução do sensor, os algoritmos de pre-processamento e os procedimentos da própria classificação. Provavelmente as mais simples das causas de erro são relacionadas com o errôneo assinalamento de categorias informacionais às categorias espectrais. Todavia, as fontes da maioria das causas de erros são mais complexas.

Ainda segundo o mesmo autor, a forma padronizada para reportar erros em locais específicos é a chamada Matriz de Erros, também conhecida como Matriz de Confusão, por identificar não somente o erro global da classificação para cada categoria, mas também como se deram as confusões entre categorias. Essencial para qualquer estudo sério sobre a exatidão dos processos de classificação, uma vez que por ela é possível a identificação, para cada categoria, dos erros de omissão e comissão, esta matriz é um exemplo

específico de uma classe mais genérica de matrizes conhecidas como Tabelas de Contingência

A verificação quantitativa da Matriz de Erros é geralmente necessária após a realização de uma inspeção inicial para se detectar a natureza genérica dos erros presentes. Alguns dos procedimentos que foram desenvolvidos para a análise de tabelas de contingência, podem ser aplicados ao exame da Matriz de Erros.

Como sugerem GONG e HOWARTH (1990), um dos procedimentos que vem sendo usado para mensurar a exatidão das classificações temáticas é o índice Kappa. Ele vem sendo recomendado como uma medida apropriada da exatidão por representar inteiramente a Matriz de Confusão. Ele toma todos os elementos da Matriz de Confusão em consideração, ao invés de apenas aqueles que se situam na diagonal da mesma, o que ocorre quando se calcula a exatidão global da classificação.

ROSENFELD e FITZPATRICK-LINS (1986) comentam ser o índice Kappa um coeficiente de concordância para escalas nominais que mede o relacionamento entre a concordância, além da casualidade, e a discordância esperada. A estimativa de Kappa (\hat{K}) é a proporção de concordância depois que a concordância devido à casualidade é retirada de consideração, ou seja:

$$\hat{K} = (p_0 - p_c) / (1 - p_c)$$

onde:

p_0 = proporção de unidades que concordam, e
 p_c = proporção de unidades que concordam por casualidade.

Simplificando o raciocínio acima, CAMPBELL (1987), propõe:

$$\hat{K} = (\text{observado} - \text{esperado}) / (1 - \text{esperado})$$

Neste caso, entenda-se por **observado** o valor global para a percentagem correta, ou seja, o somatório dos elementos presentes na diagonal da matriz dividido pelo total de elementos. Por **esperado** entenda-se os valores calculados usando-se os totais de cada linha e de cada coluna da matriz, designados pelo autor como marginais. Os produtos dos marginais de linhas pelos das colunas estimam o número de elementos (pixels) assinalados para cada célula da matriz, dado que esses pixels sejam assinalados por chance a cada categoria.

O trabalho de FOODY (1992) afirma que de acordo com a relação

$$\hat{K} = (\text{observado} - \text{esperado}) / (1 - \text{esperado}),$$

o grau de concordância por chance pode ser superestimado pelo fato dele incluir, além da concordância por chance, a concordância real. Em consequência disso, a magnitude de Kappa não estaria refletindo a concordância presente na classificação, descontada a concordância resultante somente da casualidade. Considerando-se que não sejam fixadas probabilidades *a priori* para as categorias, a proporção marginal para cada uma delas seria dada por $1/m$, onde m é o número de categorias na classificação. Consequentemente, a probabilidade de concordância por chance será $1/m$, gerando um outro índice para medir a exatidão da classificação que seria definido por:

$$K_m = (\text{observado} - 1/m) / (1 - 1/m).$$

Em recente pesquisa, MA e REDMOND (1995) concordam com o que propõe FOODY (1992), para o caso de não se estabelecer probabilidades *a priori* e estabelecendo outro procedimento quando estas probabilidades são diferentes. Os autores batizam este índice de T_e (Tau) quando do caso da equiprobabilidade entre as categorias. Desta forma, T_e torna-se um ajuste do percentual de concordância (aquele que considera somente os elementos da diagonal da matriz de erros, também chamado de concordância global) pelo número de categorias. À medida que este número aumenta, T_e se aproxima da percentagem de concordância global.

Ainda de acordo com MA e REDMOND (1995), para o caso de se desejar testar diferentes classificações, as seguintes condições devem ser satisfeitas:

. Os pixels deverão ser amostrados independentemente;

. As categorias de cobertura de solo deverão ser independentes e mutuamente exclusivas; e

. As classificações deverão ser processadas independentemente.

Em virtude de a concordância global (P_0) seguir uma distribuição binomial quando o número de pixels usado para construir a matriz de erros é grande ($N > 100$), ela irá se aproximar de uma distribuição normal (STEEL and TORRIE, 1960, citados por MA e REDMOND, 1995). Dado que a distribuição de P_0 se aproxima da normalidade, a distribuição de Tau também se aproximará (COHEN, 1960, citados por MA e REDMOND, 1995).

Considerada a normalidade, uma forma geral para um intervalo de confiança de 95%, para um coeficiente de exatidão (C), pode ser expressa por intervalo de confiança (95%) = $C \pm 1,96\sigma$.

Para testar a significância entre dois coeficientes de exatidão (C_1 e C_2), derivados de duas classificações independentes, o teste estatístico tomará a forma de

$$Z = \frac{C_1 - C_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Em virtude da concordância por casualidade (P_c) ser independente dos elementos da matriz de erros e poder ser calculada antes que se processe a classificação, ela pode ser tratada como uma constante. A variância de Tau pode ser derivada da seguinte maneira:

$$\sigma^2(P_c) = 0$$

e

$$\sigma^2(P_0) = \frac{1}{N} P_0 (1 - P_0),$$

tal que

$$\begin{aligned}\sigma^2(T) &= \sigma^2\left(\frac{P_0 - P_c}{1 - P_c}\right) \\ &= \frac{\sigma^2(P_0 - P_c)}{(1 - P_c)^2} \\ &= \frac{\sigma^2(P_0) - \sigma^2(P_c)}{(1 - P_c)^2} \\ &= \frac{P_0(1 - P_0)}{N(1 - P_c)^2}\end{aligned}$$

Obtida a variância para Tau, testes podem ser feitos para determinar se o coeficiente é significativamente maior que zero ou se dois coeficientes Tau, derivados de classificações independentes, são significativamente diferentes.

Ainda de acordo com MA e REDMOND (1995), existem pelo menos três razões para justificar o uso do coeficiente Tau na verificação da exatidão das classificações baseadas em dados de sensores remotos:

. Conceitualmente, o coeficiente Tau é mais fácil de entender e interpretar do que Kappa. A concordância por chance (P_c), usada para calcular Tau, é uma medida direta do assinalamento casual; em razão disso Tau provê uma clara indicação de como a classificação se comporta quando comparada com o assinalamento casual dos pixels às categorias ou classes. Com Kappa, por outro lado, a concordância por chance (ou casual) pode resultar do julgamento do analista adicionado à casualidade, de tal forma que este coeficiente ou índice superestima a concordância por chance (P_c) (BRENNAN e PREDIGER (1981), citados por MA e REDMOND (1995));

. Tau e sua estimativa de variância são relativamente simples de serem calculados. Devido à concordância por chance, no cálculo de Tau, ser independente da matriz de erros, ela pode ser conhecida antes da construção da matriz;

. Diferentemente de Kappa, Tau toma em consideração o número de categorias ou classes presentes na classificação.

2. Metodologia

2.1 Localização da Área de Estudo

A área de estudo está localizada no Vale do Rio Doce, estado de Minas Gerais, compreendida entre as coordenadas UTM 745.595 N, 7.797.446 E e 758.315 N, 7.808.096 E. Esta compreende uma

superfície de aproximadamente 13.500 ha, incluindo florestas de Eucalyptus, florestas naturais, pastagens, lagos, rios e áreas de solo exposto. De acordo com GOLFARI (1975), a região possui terreno relativamente rugoso, com uma elevação média de 350 m, variando entre 150 e 650 m. As terras são de propriedade da Companhia Agrícola e Florestal Santa Bárbara que utiliza suas florestas cultivadas para a produção de carvão, com rotação de 7 anos.

2.2 A Coleta das Amostras de Treinamento

As amostras que foram utilizadas no treinamento do classificador foram coletadas de forma blocada, a partir da interpretação das fotografias aéreas, com o auxílio de um estereoscópio de espelhos e de oculares de aumento acopladas a ele. Do próprio conhecimento da região e pela interpretação das fotos aéreas, em estudo previamente conduzido, verificou-se que as seguintes categorias informacionais seriam capazes de expressar a variabilidade presente na área de estudo:

- Rebrota sobre Solo Vegetado (RSV) categoria caracterizada pela rebrota do Eucalypto sobre solo coberto por vegetação essencialmente herbácea-arbustiva;
- Rebrota sobre Solo Parcialmente Exposto (RSPE) caracterizada pela rebrota do Eucalypto sobre solo que apresenta partes totalmente expostas, sem qualquer tipo de vegetação;
- Rebrota sobre Solo Exposto (RSE) caracterizada pela rebrota do Eucalypto sobre solo totalmente exposto;
- Eucalypto com Fechamento de Copa (EUFECH) caracterizada por plantios de Eucalyptos, já em idade adulta, com fechamento de copa, sem que haja exposição do solo;
- Eucalypto sem Fechamento de Copa (EUABER) caracterizada por plantios de Eucalyptos, já em idade adulta, com falhas impedindo que ocorra o fechamento de copa e permitindo a exposição do solo, vegetado ou não.
- Sombras (SOMBRAS) categoria constituída por áreas da imagem que se encontravam sob sombra impossibilitando a correta identificação dos pixels;
- Herbácea-Arbustiva (HERB_ARB) categoria constituída por vegetação de natureza herbácea-arbustiva.
- Solo Exposto (SOLO_EXP) categoria constituída por solo nú, desprovido de cobertura vegetal;
- Água Não Turva (AGUALIMP) água com pouca ou nenhuma turbidez;
- Água Turva (AGUATURV)

água com muito material em suspensão, acarretando turbidez;

- Nativas (NATIVAS)
vegetação arbórea, nativa da região.

2.3 O Algoritmo de Classificação

Todas as classificações levadas a efeito no estudo foram produzidas aplicando-se um algoritmo de máxima verossimilhança, provido pelo sistema IDRISI (programa MAXVER), assumindo-se a equiprobabilidade de ocorrência das categorias listadas anteriormente. Este algoritmo exige como entrada a definição de probabilidades *a priori* e os nomes dos arquivos de assinaturas a serem usados no processamento. A não especificação de probabilidades para nenhuma das classes implica na equiprobabilidade.

2.4 Combinações de Bandas e Transformações Especiais

O algoritmo de máxima verossimilhança foi aplicado às seguintes combinações de bandas, gerando as classificações à sua direita:

COMBINAÇÃO	CLASSIFICAÇÃO
TM3, TM4, TM5	MAX345
TM3, TM4, TM7	MAX347
TM2, TM3, TM4	MAX234

Da mesma forma ele foi aplicado às componentes “verdor”, “brilho” e “umidade”, da transformação *Tasseled Cap* geradas a partir da aplicação dos coeficientes aplicados às bandas reflexivas da imagem, de acordo com (JENSEN, 1986), gerando a classificação denominada MAXTC.

Uma última transformação utilizada foi a resultante da análise de componentes principais. As três primeiras componentes principais foram escolhidas por contemplarem 99,15% da variância presente em toda a imagem, consideradas todas as bandas reflexivas. A classificação resultante da utilização destas componentes recebeu o nome de MAXPC123.

2.5 Amostragem Sobre os Dados de Referência

A amostragem sobre os dados de referência foram conduzidas de acordo com as estratégias blocada, sistemática estratificada, sistemática e casualizada.

No caso da amostragem blocada foram selecionados, sobre a imagem, polígonos representativos das 11 categorias informacionais em concordância com a interpretação das fotos aéreas. Estes polígonos foram delineados sem que pixels que já houvessem sido amostrados na etapa de treinamento fossem contemplados. Para que isso fosse conseguido foi sobreposto o arquivo vetorial, contendo os polígonos de treinamento, à imagem sobre a qual seriam coletados os novos polígonos que serviriam para a verificação da exatidão das classificações. Os polígonos coletados, no formato vetorial, foram então transformados para o

formato matricial ou “raster”, gerando um novo arquivo.

Para as amostragens não blocadas o procedimento foi bastante semelhante. Utilizando-se o algoritmo denominado SAMPLE, do sistema IDRISI, foram gerados arquivos vetoriais, um para cada amostragem não blocada, contendo, inicialmente, 1200 pontos, cada. Este número de pontos foi determinado empiricamente objetivando contemplar todas as categorias presentes na imagem, especialmente aquelas que ocorrem com pouca frequência ocupando pequenas áreas. De acordo com JENSEN (1986), o número mínimo de pontos pode ser determinado pela relação:

$$N = [4(p)(q)] / E^2$$

em que,

p = percentual de exatidão esperado;

q = 1 - p;

E = erro permissível.

No presente caso foi considerado um percentual de exatidão de 85% e um erro permissível de 5%, o que implica, após a aplicação da relação acima, em um número mínimo de 204 pixels a serem amostrados.

À semelhança do que ocorreu com a amostragem blocada, para as outras amostragens também foram desconsiderados aqueles pixels que foram tomados para o treinamento do classificador.

De posse dos arquivos vetoriais contendo os pontos a serem amostrados, eles foram sobrepostos a diferentes composições de bandas que permitissem a melhor observação dos diversos tipos de cobertura vegetal. Desta forma, ora foi utilizada a composição em falsa cor dos canais TM4, TM3 e TM2, ora a imagem resultante da primeira componente principal e outras. Cada ponto presente sobre a imagem orbital foi localizado nas fotografias e, com o auxílio de um estereoscópio de espelhos, verificada a que categoria informacional o pixel sob aquele ponto pertenceria. Para que a correspondência imagem/fotografia fosse a mais exata possível, lançou-se mão de operações de ampliação sobre a imagem e do uso de oculares de aumento com o estereoscópio e, também, de um reticulado auxiliar que subdividia a imagem em 16 sub-imagens para que o analista não se perdesse na tarefa de identificar cada um dos pontos. Desta forma, a cada ponto que era identificado, digitalizava-se um outro coincidente com o que havia sido sobreposto anteriormente, com o atributo referente à categoria informacional que ele pertencesse. Em resumo, o arquivo contendo os pontos amostrais foi sobreposto à imagem apenas para servir de orientação locacional na

coleta dos pontos que comporiam o arquivo de referência. Esta seqüência foi repetida para todos os procedimentos não bloqueados de amostragem, gerando novos arquivos vetoriais que, em seguida, foram submetidos a um processo de rasterização gerando, como resultado, arquivos matriciais. Estes arquivos foram então submetidos a uma operação na qual cada um deles foi multiplicado por um arquivo booleano onde os pixels no interior dos polígonos tomados como amostras de treinamento receberam o valor zero, e os demais valor 1. Desta forma os arquivos de referência passaram a conter unicamente pixels que não faziam parte das amostras de treinamento, estando prontos para serem usados nas verificações de exatidão.

2.6 O Cálculo dos Índices e de Suas Variâncias

Uma vez estabelecidas as 5 classificações (MAXTC, MAXPC123, MAX345, MAX347 e MAX234) e os quatro arquivos de referência, um para cada estratégia de amostragem (bloqueada, casualizada, sistemática estratificada e sistematizada), foi possível gerar as matrizes de erro que possibilitariam o cálculo dos índices de exatidão (Global, Kappa e Tau).

Foi gerada uma matriz de erro para cada classificação comparada com cada um dos quatro arquivos de referência resultando, ao final do processo, 4 matrizes de erro para cada classificação. Para cada uma destas matrizes foram calculados os três índices de exatidão e suas respectivas variâncias, de acordo com as relações específicas para cada índice (MA e ROLAND, 1995).

O índice Exatidão Global, bem como sua variância podem ser determinados pelas seguintes relações:

$$G = \frac{\sum_{i=1}^M n_{ii}}{N}$$

$$\sigma_G^2 = \frac{P_0(1 - P_0)}{N}$$

De forma semelhante, para o índice Kappa, tem-se:

$$K = \frac{P_0 - P_c}{1 - P_c}$$

$$\sigma_K^2 \approx \frac{P_0(1 - P_0)}{N(1 - P_c)^2}$$

Para o índice Tau, vem:

$$T = \frac{P_0 - 1/M}{1 - 1/M}$$

$$\sigma_T^2 = \frac{P_0(1 - P_0)}{N(1 - 1/M)^2}$$

As relações para P_0 e P_c podem ser explicitadas por:

$$P_0 = \frac{\sum_{i=1}^M n_{ii}}{N}, e$$

$$P_c = \frac{\sum_{i=1}^M n_{i+} n_{+i}}{N^2}, onde$$

N = número total de pixels contemplados pela matriz de erros;

n = elemento da matriz de erros, e

M = número de categorias informacionais presentes na matriz de erros.

De posse dos índices e de suas respectivas variâncias foi possível se efetuar comparações entre as classificações e índices.

2.7 As Comparações

As comparações foram feitas com base nos valores absolutos dos índices, com base nos valores de Z ao nível de 99% de probabilidade e, graficamente, com a finalidade de melhor visualizar o comportamento dos três índices em função da variação das estratégias de amostragem e classificações.

Para testar a significância da diferença entre quaisquer dois índices de exatidão (C_1 e C_2), o teste estatístico utilizado foi o de Z , especificado abaixo, como preconizam MA e REDMOND (1995):

$$Z = \frac{C_2 - C_1}{\sqrt{\sigma_{K_2}^2 + \sigma_{K_1}^2}}$$

3. Conclusão

O objetivo do presente trabalho é o de comparar três diferentes índices de exatidão para classificações. Dois deles, Global e Kappa, já são comumente usados na área de sensoriamento remoto. O terceiro, Tau, foi proposto recentemente à comunidade da área (MA e Redmond, 1995), tentando corrigir deficiências incorporadas no cálculo do índice Kappa.

Para que as comparações fossem levadas a efeito foi necessário o cálculo das variâncias para cada índice, em cada classificação, tomando-se todos os quatro arquivos de referência, gerados a partir das diferentes estratégias de amostragem.

Para que se pudesse proporcionar uma visão geral do comportamento dos três índices, simultaneamente, os seus valores foram plotados (figura

1). Da análise desta figura, percebe-se que os três índices apresentam curvas extremamente semelhantes, apenas deslocando-se umas em relação às outras. A posição das curvas pode ser explicada pela forma com que cada índice contempla a concordância casual em sua formulação. Assim, o índice Global (G) apresenta os maiores valores pelo fato de considerar apenas a diagonal principal da matriz de erros, ignorando a ocorrência de tal concordância. Por outro lado o índice Kappa (K), ao calcular a concordância casual, inclui nos cálculos os elementos da diagonal principal fazendo com que seja superestimada, reduzindo o valor do índice. Na formulação do índice Tau, o valor da concordância casual é estabelecido *a priori* evitando as falhas em que incorrem G e T. A colocação da curva representando T, entre G e K, indica um bom balanceamento na consideração das concordâncias real e casual, tomando-se em consideração os dois extremos que representam as outras duas curvas, ou seja: total descondição da casualidade e sua superestimação. Tal fato é um indicador de que talvez seja ele o índice que com mais fidedignidade represente a exatidão dos processos de classificação.

No que se refere às três curvas na porção relativa ao procedimento bloqueado de amostragem, depreende-se da análise visual que as diferenças entre os valores dos três índices são muito pequenas, muito pouco importando o índice que se eleja para se verificar a exatidão da classificação. Quando se considera os procedimentos não bloqueados as diferenças são bem mais representativas, reportando valores de exatidão bastante distintos, a depender do índice que seja escolhido.

A comparação dos valores dos índices obtidos a partir do uso de estratégias não bloqueadas e bloqueada mostra ser o tipo de amostragem o principal fator que concorre para a superestimação dos índices. Considerando-se cada um dos índices, no caso das amostragens não bloqueadas, para cada classificação, é notável a diferença entre seus valores absolutos, porém, à exceção da amostragem sistemática, as diferenças

entre os índices de mesma natureza para quaisquer duas classificações são bastante semelhantes para qualquer índice que se tome. Este aspecto evidencia que qualquer dos três índices utilizados deverão ter comportamento semelhante no que se refere ao estabelecimento de diferenças entre as classificações.

Com o objetivo de se mensurar a significância das diferenças entre classificações por estratégia de amostragem dos dados de referência, e por estratégia de amostragem considerando-se, individualmente, cada classificação, foi empregado o teste estatístico Z ao nível de 99 % de probabilidade, de acordo com o que já foi descrito na metodologia.

Os quadros 1 e 2 mostram estas comparações utilizando todas as possíveis combinações de classificações e amostragens. No primeiro deles, foram comparadas entre si todas as classificações geradas dentro de cada procedimento de amostragem. Nota-se, a partir de sua observação, que a sensibilidade do teste às diferenças nas classificações aumenta para as amostragens sistemática e bloqueada, nesta ordem, o que parece indicar o efeito da interferência do analista no processo amostral, se considerarmos as amostragens casualizadas como as que apresentam a maior liberdade para a casualização, visto que o analista se limita a informar o número de pontos a ser amostrados; a amostragem sistemática com o analista interferindo ao selecionar o ponto de partida para o estabelecimento da rede de pontos e, em virtude disso, violando a independência das observações, devido ao igual espaçamento entre elas; e a amostragem bloqueada como sendo desprovida de qualquer casualização.

Os índices, por sua vez, à exceção das amostragens sistemática e bloqueada, apresentaram a mesma capacidade de identificar diferenças, ou seja as classificações que foram diferentes entre si, o foram pelos três índices.

COMPORTAMENTO DE G, K e T

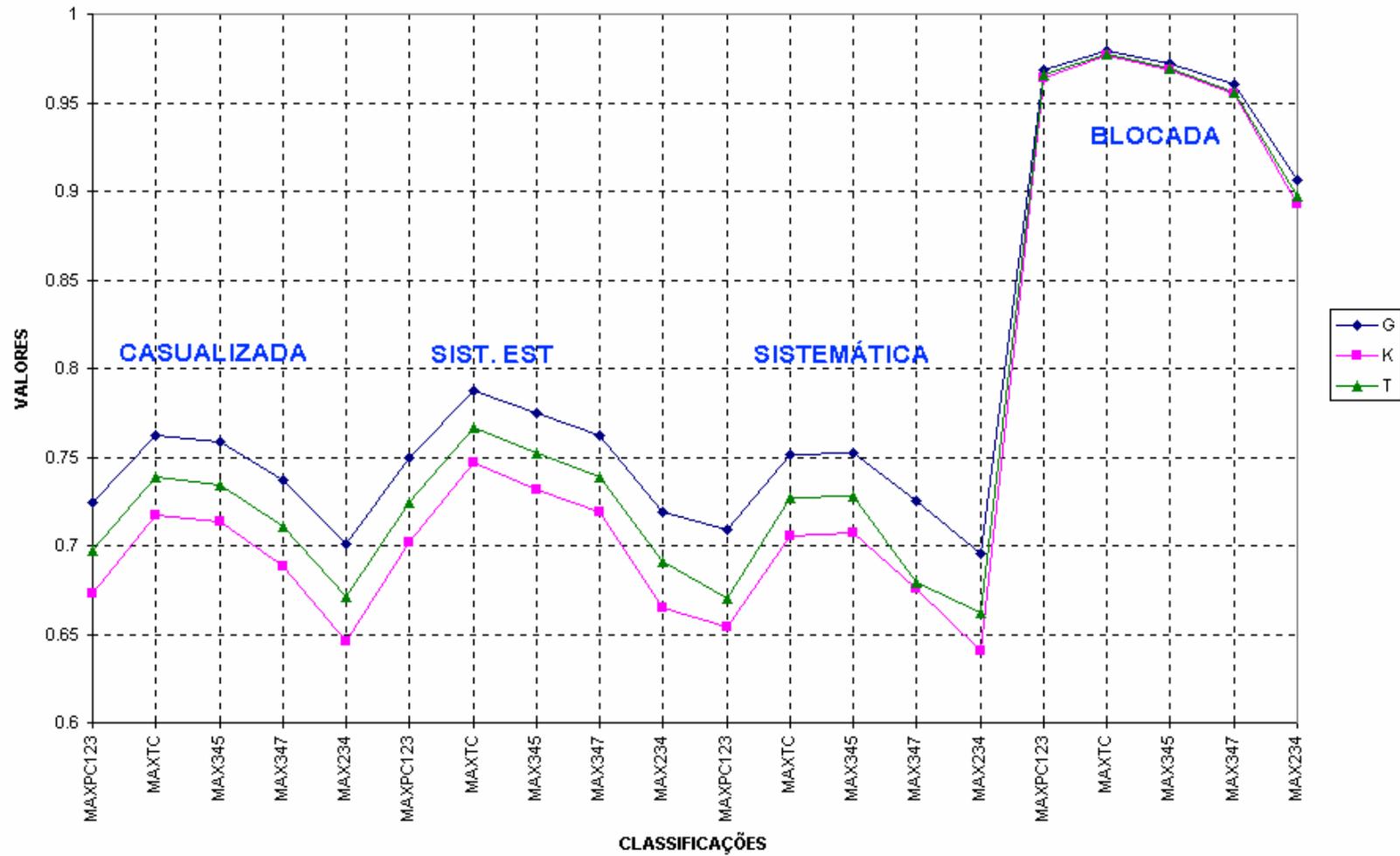


Figura 1 - Comportamento dos índices de exatidão Global (G), Kappa (K) e Tau (T) considerando-se todas as classificações e todas as estratégias de amostragem.

Quadro 1 - Diferenças, por índices, tomando por base as amostragens. As células assinaladas por um “X” significam que a diferença é significativa por Z, ao nível de 99% de probabilidade.

		AMOSTRAGEM					AMOSTRAGEM		
		G	K	T			G	K	T
Casualizada	MAXPC123 X MAXTC				Sist. Estratificada	MAXPC123 X MAXTC			
	MAXPC123 X MAX345					MAXPC123 X MAX345			
	MAXPC123 X MAX347					MAXPC123 X MAX347			
	MAXPC123 X MAX234					MAXPC123 X MAX234			
	MAXTC X MAX345					MAXTC X MAX345			
	MAXTC X MAX347					MAXTC X MAX347			
	MAXTC X MAX234	X	X	X		MAXTC X MAX234	X	X	X
	MAX345 X MAX347					MAX345 X MAX347			
	MAX345 X MAX234	X	X	X		MAX345 X MAX234	X	X	X
	MAX347 X MAX234					MAX347 X MAX234			

		AMOSTRAGEM					AMOSTRAGEM		
		G	K	T			G	K	T
Sistemática	MAXPC123 X MAXTC			X	Blocada	MAXPC123 X MAXTC	X	X	X
	MAXPC123 X MAX345	X	X	X		MAXPC123 X MAX345			
	MAXPC123 X MAX347					MAXPC123 X MAX347			
	MAXPC123 X MAX234					MAXPC123 X MAX234	X	X	X
	MAXTC X MAX345					MAXTC X MAX345			
	MAXTC X MAX347			X		MAXTC X MAX347	X	X	X
	MAXTC X MAX234	X	X	X		MAXTC X MAX234	X	X	X
	MAX345 X MAX347			X		MAX345 X MAX347	X	X	X
	MAX345 X MAX234	X	X	X		MAX345 X MAX234	X	X	X
	MAX347 X MAX234					MAX347 X MAX234		X	

No quadro 2, a análise é feita fixando-se cada uma das classificações e comparando-se os índices de exatidão obtidos a partir de cada estratégia de amostragem aplicadas sobre os dados de referência.

Da observação do quadro depreende-se que só ocorreu diferença significativa quando foram comparados procedimentos não bloqueados com aquele que utilizou a estratégia bloqueada, o que leva a concluir que não há diferença entre as amostragens não bloqueadas. Nesta situação, no caso de se adotar a amostragem pixel

a pixel, poderia ser recomendado, por questões de custo e de implementação, a amostragem sistemática, desde que as feições de interesse não apresentem o caráter de periodicidade na paisagem.

No que se refere, neste quadro, à comparação entre os três índices, todos se comportaram de forma idêntica, ou seja, apresentaram exatamente a mesma sensibilidade às comparações entre os diferentes tipos de amostragem.

Quadro 2 - Diferenças, por índices, tomando por base as classificações. As células assinaladas por um "X" significam que a diferença é significativa por Z, ao nível de 99% de probabilidade.

CLASSIFICAÇÕES											
MAXPC123			MAXTC			MAX345					
	G	K	T		G	K	T		G	K	T
CASUAL. X SIST.EST.				CASUAL. X SIST.EST.				CASUAL. X SIST.EST.			
CASUAL. X SIST.				CASUAL. X SIST.				CASUAL. X SIST.			
CASUAL. X BLOC.	X	X	X	CASUAL. X BLOC.	X	X	X	CASUAL. X BLOC.	X	X	X
SIST. EST. X SIST.				SIST. EST. X SIST.				SIST. EST. X SIST.			
SIST. EST. X BLOC.	X	X	X	SIST. EST. X BLOC.	X	X	X	SIST. EST. X BLOC.	X	X	X
SIST. X BLOC.	X	X	X	SIST. X BLOC.	X	X	X	SIST. X BLOC.	X	X	X

CLASSIFICAÇÕES									
MAX347			MAX234						
	G	K	T		G	K	T		
CASUAL. X SIST.EST.				CASUAL. X SIST.EST.					
CASUAL. X SIST.				CASUAL. X SIST.					
CASUAL. X BLOC.	X	X	X	CASUAL. X BLOC.	X	X	X		
SIST. EST. X SIST.				SIST. EST. X SIST.					
SIST. EST. X BLOC.	X	X	X	SIST. EST. X BLOC.	X	X	X		
SIST. X BLOC.	X	X	X	SIST. X BLOC.	X	X	X		

Referências

CAMPBELL, J. B. *Introduction to remote sensing*. New York, The Guilford Press, 1987. 551p.

CONGALTON, R.G. *A comparison of five sampling schemes used in assessing the accuracy of land cover/land use maps derived from remotely sensed data*. Blacksburg, Virginia Politechnic Institute and State University. 1984. 146 p. (Tese PhD.)

FOODY, G.M. On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing*, Bethesda, 58(10), 1459 - 1460. 1990.

GOLFARI, L., *Zoneamento ecológico do Estado de Minas Gerais para o reflorestamento*. Belo Horizonte, Centro de Pesquisas do IBDF na Região dos Cerrados, 1975. 65p.

GONG, P., HOWART, P.J. An assessment of some factors influencing multispectral land-cover classification. *Photogrammetric Engineering and Remote Sensing*, Bethesda, 56(5), p.597-603, 1990.

MA, Z., REDMOND, R.L. Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing*, Bethesda, 61(4), 453 - 439. 1995.

ROSENFELD, G.H., FITZPATRICK-LINS, K.A. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, Bethesda, 52(2), 223-227. 1986.