

GEÍZA COUTINHO FIGUEIREDO

**EXATIDÃO POSICIONAL E TEMÁTICA DE IMAGENS ORBITAIS**

Dissertação apresentada à  
Universidade Federal de Viçosa, como  
parte das exigências do Programa de  
Pós-Graduação em Engenharia Civil,  
para obtenção do título de *Magister  
Scientiae*.

VIÇOSA  
MINAS GERAIS – BRASIL  
2007

**Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV**

T

F475e  
2007

Figueiredo, Geíza Coutinho, 1981-  
Exatidão posicional e temática de imagens orbitais  
Geíza Coutinho Figueiredo. – Viçosa, MG , 2007.  
xv, 87f. : il. (algumas col.) ; 29cm.

Inclui apêndice.

Orientador: Antonio Simões Silva.

Dissertação (mestrado) - Universidade Federal de  
Viçosa.

Referências bibliográficas: f. 69-72.

1. Sensoriamento remoto. 2. Satélites artificiais em  
sensoriamento remoto. I. Universidade Federal de Viçosa.  
II.Título.

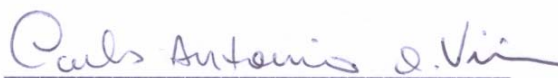
CDD 22.ed. 621.3678

GEÍZA COUTINHO FIGUEIREDO

**EXATIDÃO POSICIONAL E TEMÁTICA DE IMAGENS ORBITAIS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Magister Scientiae*.

APROVADA: 7 de março de 2007.



Prof. Carlos Antônio Oliveira  
Vieira  
(Co-orientador)



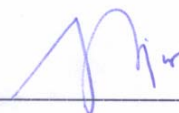
Prof. Dalto Domingos Rodrigues  
(Co-orientador)



Prof. Ricardo Seixas Brites



Prof. Elpídio Inácio Fernandes  
Filho



Prof. Antônio Simões Silva  
(Orientador)

A Deus por sua fidelidade.

À família pelo apoio.

Aos amigos pela companhia.

À igreja pelas orações.

Dedico

## **AGRADECIMENTOS**

À Deus, pela vida e por tudo o quanto me tem feito.

À Universidade Federal de Viçosa (UFV), especialmente ao Departamento de Engenharia Civil, pela oportunidade de realização deste curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de mestrado.

Ao professor Antônio Simões Silva pela orientação no desenvolvimento desse trabalho.

Ao professor Carlos Antônio Oliveira Vieira, pela colaboração, pela paciência, pelo apoio e orientação durante a realização deste trabalho e por ceder a imagem Landsat.

Ao professor Dalto Domingos Rodrigues, pelo apoio, pela contribuição e incentivo transmitido durante o curso.

Ao professor Vicente Paulo Soares, do Departamento de Engenharia Florestal, por ceder imagem Ikonos da região da Serra do Salitre, pelo uso do software ERDAS Imagine 8.5 e pela colaboração.

Ao Instituto Nacional de Pesquisas Espaciais – INPE, pela imagem CBERS.

Ao Núcleo de Geoprocessamento Professor Celestino Aspiazú - NUGEO, pela utilização do software Idrisi Kilimanjaro.

Ao Laboratório de Geoprocessamento da Universidade de Uberaba - LabGeo, pela utilização das cartas do serviço militar na escala 1:25.000, formato digital, da região de Uberaba.

A todos os colegas do curso de pós-graduação do Departamento de Engenharia Civil, pelo apoio e amizade, especialmente à Alessandra Braga, Leonardo Assis, Giuliano Marotta e Rafael Andrade pela colaboração.

Aos meus pais Iza (*in memoriam*) e Getúlio, e aos meus irmãos Geilson e Getúlio, pelo apoio, pelo amor e pelo incentivo, que foram de grande valia.

A minha segunda família, Waldir e Jeanny, por todo incentivo, carinho e atenção em todos os momentos.

Ao Julio Cesar G. Vendramini, pelo carinho, apoio, amor, paciência e compreensão.

Aos amigos Marcos do Nascimento, José Carlos, Webert, Leililene, Geraldo Júnior, Gisele, Rosinha e Laís, pela amizade e por me apoiarem nos momentos mais importantes.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

## **BIOGRAFIA**

GEÍZA COUTINHO FIGUEIREDO, filha de Getúlio Marques Figueiredo e Iza Coutinho Figueiredo, nasceu em 20 de agosto de 1981, na cidade de Vitória, Espírito Santo.

Em 2000, iniciou o Curso de Engenharia de Agrimensura na Universidade Federal de Viçosa, Viçosa-MG, graduando-se em 2005.

Em março de 2005, iniciou o Programa de Pós-Graduação em Engenharia Civil com ênfase em Informações Espaciais, em nível de mestrado, na Universidade Federal de Viçosa, submetendo-se à defesa de dissertação em 7 de março de 2007.

## SUMÁRIO

RESUMO .....	xii
ABSTRACT .....	xiv
1. INTRODUÇÃO .....	1
1.1 Generalidades .....	1
1.2 Problema e Motivação .....	3
1.3 Objetivos .....	4
1.4 Estrutura da dissertação .....	4
2. REVISÃO BIBLIOGRÁFICA .....	6
2.1 Exatidão Posicional .....	6
2.2 Exatidão Temática .....	10
2.2.1 Esquemas de Amostragem .....	10
2.2.2 Tamanho da Amostra .....	12
2.2.3 Matriz de Confusão (ou Matriz de Erro) .....	13
2.2.4 Medidas e Índices Derivados da Matriz de Confusão .....	14
2.2.5 Distribuição Espacial do Erro .....	18
3. EXATIDÃO POSICIONAL .....	20
3.1 Caracterização das áreas de estudo .....	20
3.2 Materiais .....	23
3.2.1 Dados Orbitais .....	23
3.2.2 Ferramentas Computacionais .....	25

3.3 Métodos empregados na avaliação da exatidão posicional .....	26
4. EXATIDÃO TEMÁTICA.....	30
4.1 Métodos aplicados para análise não-espacial da exatidão temática.....	30
4.2 Análise espacial .....	36
4.2.1 Caracterização da área de estudo .....	36
4.2.2 Materiais .....	36
4.2.2.1 Dado Orbital .....	36
4.2.2.2 Ferramentas Computacionais .....	37
4.2.3 Métodos aplicados para análise espacial da exatidão temática .....	38
5. RESULTADOS E DISCUSSÃO .....	41
5.1 Exatidão Posicional .....	41
5.1.1 Exatidão Posicional em imagens de altíssima resolução espacial ...	41
5.1.2 Exatidão Posicional em imagens de alta resolução espacial.....	44
5.1.3 Exatidão Posicional em imagens de média resolução espacial.....	53
5.2 Exatidão temática .....	55
5.2.1 Análise não-espacial da exatidão temática.....	55
5.2.2 Análise espacial da exatidão temática.....	62
6. CONCLUSÃO .....	67
REFERÊNCIA BIBLIOGRÁFICA .....	69
APÊNDICE A .....	73
APÊNDICE B .....	76
APÊNDICE C .....	79
APÊNDICE D .....	82
APÊNDICE E .....	85

## LISTA DE FIGURAS

Figura 1: Ilustração do princípio do Método do Ponto Gerado.....	8
Figura 2: Área entre duas feições. ....	9
Figura 3: Ponto $a_i$ e os quatro pontos mais próximos (a) antes da rotação e (b) após a rotação. ....	10
Figura 4: Localização de Viçosa - MG. ....	21
Figura 5: Localização de Uberaba - MG. ....	22
Figura 6: Localização da área de abrangência da EDR de Limeira - SP. ....	23
Figura 7: (a) Imagem Ikonos II; (b) Imagem Quickbird.....	24
Figura 8: (a) Imagem Landsat; (b) Imagem CBERS. ....	24
Figura 9: Imagem MODIS. ....	25
Figura 10: Localização da Serra do Salitre - MG. ....	36
Figura 11: Imagem Ikonos da região da Serra do Salitre.....	37
Figura 12: Variando o número de classes informacionais num nível de desempenho excelente.....	56
Figura 13: Variando o número de classes informacionais num nível de desempenho muito bom. ....	56
Figura 14: Variando o número de classes informacionais num nível de desempenho bom. ....	56
Figura 15: Variando o número de classes informacionais num nível de desempenho razoável.....	57

Figura 16: Variando o número de classes informacionais num nível de desempenho ruim. ....	57
Figura 17: Variando o total de amostras num nível de desempenho excelente. ....	58
Figura 18: Variando o total de amostras num nível de desempenho muito bom. ....	58
Figura 19: Variando o total de amostras num nível de desempenho bom. ....	58
Figura 20: Variando o total de amostras num nível de desempenho razoável. ....	59
Figura 21: Variando o total de amostras num nível de desempenho ruim. ....	59
Figura 22: Variando o total de pixels com classe predominante num nível de desempenho excelente. ....	60
Figura 23: Variando o total de pixels com classe predominante num nível de desempenho muito bom. ....	60
Figura 24: Variando o total de pixels com classe predominante num nível de desempenho bom. ....	60
Figura 25: Variando o total de pixels com classe predominante num nível de desempenho razoável. ....	61
Figura 26: Variando o total de pixels com classe predominante num nível de desempenho ruim. ....	61
Figura 27: Imagem de distância. ....	64
Figura 28: Imagem binária do erro. ....	65
Figura 29: Imagem de diferença entre as classificações MaxVer e RNA. ....	66

## LISTA DE TABELAS

Tabela 1 – Representação matemática de uma matriz de confusão. ....	14
Tabela 2 – Matriz de confusão.....	15
Tabela 3 – Índice Kappa e o correspondente desempenho da classificação. ..	16
Tabela 4 – Erro médio quadrático (RMS) pós correção geométrica. ....	26
Tabela 5 – Matriz de confusão utilizada para os exemplos a seguir.....	31
Tabela 6 – Classes na imagem de referência.....	38
Tabela 7 – Pontos homólogos referentes à imagem Ikonos de Viçosa-MG. ....	42
Tabela 8 – Pontos homólogos referentes à imagem Quickbird de Viçosa-MG.	42
Tabela 9 – Resultado do PEC, EP e $\sigma_P$ para a escala 1:10.000.....	43
Tabela 10 – Valores de $\chi^2_P$ usados na análise de exatidão para as imagens Ikonos e Quickbird. ....	43
Tabela 11 – Pontos homólogos referentes à imagem CBERS de Uberaba-MG. ....	45
Tabela 12 – Pontos homólogos referentes à imagem Landsat de Uberaba-MG. ....	46
Tabela 13 – Resultado do PEC, EP e $\sigma_P$ para a escala 1:25.000. ....	47
Tabela 14 – Valores de $\chi^2_P$ usados na análise de exatidão para as imagens CBERS e Landsat.....	47
Tabela 15 – Distâncias entre pontos equidistantes referentes à Uberaba – MG, coletados através de feições homólogas. ....	49

Tabela 16 – Valores de $\chi^2_D$ utilizados na análise de exatidão pelo método do Ponto Gerado.....	50
Tabela 17 – Áreas e perímetros originados da sobreposição de feições homólogas .....	51
Tabela 18 – Pontos homólogos referentes a abrangência da EDR de Limeira-SP.....	53
Tabela 19 – Resultado do PEC, EP e $\sigma_P$ para a escala 1:400.000.....	54
Tabela 20 – Valores de $\chi^2_P$ usados na análise de exatidão para a imagem MODIS.....	54
Tabela 21 – Matriz de confusão resultante da classificação MaxVer e os índices de exatidão. ....	62
Tabela 22 – Matriz de confusão resultante da classificação RNA e os índices de exatidão. ....	63

## RESUMO

FIGUEIREDO, Geíza Coutinho, M.Sc., Universidade Federal de Viçosa, março de 2007. **Exatidão posicional e temática de imagens orbitais**. Orientador: Antônio Simões Silva. Co-Orientadores: Carlos Antônio Oliveira Vieira e Dalto Domingos Rodrigues.

A presente dissertação teve como finalidade avaliar a exatidão posicional e representar espacialmente a exatidão temática em produtos de sistemas imageadores, empregando como metodologias o *Padrão de Exatidão Cartografia* (PEC), para avaliar a exatidão posicional, e um mapa de confiabilidade temático, para representar a distribuição espacial do erro, de forma a apresentar ao usuário uma indicação visual da confiabilidade nas atribuições dos rótulos de cada *pixel* na imagem. Também foi realizado um estudo para verificar o comportamento dos índices de concordância, amplamente usados para avaliar a exatidão temática, quando da variação do número de classes informacionais e/ou do número de padrões de validação ou mesmo se houver classe predominante entre os valores amostrados, para validação, entre as classes informacionais. Pelos resultados obtidos, pode-se concluir que os dados provenientes do sensoriamento remoto possuem exatidão posicional para serem utilizados para fins de atualização, desde que estejam atendendo às normas estabelecidas para o controle de qualidade na escala de trabalho desejada. Com relação à exatidão temática, pôde-se

verificar que o comportamento dos índices de exatidão temática são mais afetados quando ocorre discrepância entre os valores amostrados e que o mapa de confiabilidade temático mostra ser uma forma eficaz para representar a exatidão temática, pois fornece uma boa representação visual das áreas que possuem baixa precisão dentro da imagem classificada.

## ABSTRACT

FIGUEIREDO, Geíza Coutinho, M.Sc., Universidade Federal de Viçosa, March of 2007. **Positional and thematic accuracy of orbital images.** Adviser: Antônio Simões Silva. Co-Advisers: Carlos Antônio Oliveira Vieira and Dalto Domingos Rodrigues.

This dissertation had the purpose of evaluating the positional accuracy and representing, spatially, the thematic accuracy in products of imaging systems by using as methodologies the *Map Accuracy Standards* (MAS), to evaluate the positional accuracy, and a thematic reliability map, to represent the space distribution of errors, in way to present to the user a visual indication of the reliability in the attributions of the labels of each pixel in the image. Also, a study was conducted to verify the behavior of the agreement indexes, which are largely used to evaluate the thematic accuracy, when the variation of the number of informational classes and/or the number of validation patterns or even if there is predominant class among the sample used for validation, among the informational classes. Based on the obtained results, it can conclude that the coming data of the remote sensing have positional accuracy to be used for purpose updating, since they are assisting the established norms for the quality control in the work scale wanted. Regarding thematic accuracy, it could be verified that the behavior of the indexes of thematic accuracy are more affected

when it occurs discrepancy among the sample values and that the thematic reliability map seems to be an efficient form to represent the thematic accuracy, because it provides good visual representation of the areas that have low precision inside of the classified image.

# 1. INTRODUÇÃO

## 1.1 Generalidades

Nas últimas décadas assistiu-se a uma notável evolução das tecnologias destinadas à observação e medição de feições e fenômenos que ocorrem na superfície da Terra. O sensoriamento remoto é uma das tecnologias que apresentou um alto nível de desenvolvimento, por receber grandes investimentos. Hoje em dia, encontram-se disponíveis imagens oriundas dos mais diversos tipos de sensores, com diferentes resoluções espaciais, espectrais, radiométricas e temporais; abrindo-se um enorme leque de opções para profissionais que trabalham com imageamento (Alves e Vergara, 2005).

Segundo Tisot et al. (2005), a produção de mapas temáticos, como os de uso e cobertura do solo, através da classificação de imagens, é uma das aplicações mais utilizadas do sensoriamento remoto. O aumento da exatidão de classificação é um dos principais enfoques de pesquisa na área de classificação de imagens.

As imagens orbitais servem como uma fonte de dados para os mais diversos estudos e levantamentos, representando uma das únicas formas viáveis para o monitoramento ambiental, nas mais diversas escalas, devido à rapidez, eficiência e periodicidade, que as caracterizam (Crósta, 1993). O

sensoriamento remoto é uma das principais ferramentas aplicadas no monitoramento dos fenômenos que ocorrem na superfície terrestre e, atualmente, passou a ser de grande importância quando se pensa em atualização e mapeamento sistemático.

Porém, com toda essa facilidade de aquisição de dados provenientes das imagens do sensoriamento remoto, surge o problema da verificação da qualidade dos produtos cartográficos originados dessas imagens. Estas imagens estão sujeitas às distorções advindas da instabilidade das plataformas orbitais e necessitam ser avaliadas e corrigidas antes de serem utilizadas (Lunetta et al., 1991).

No contexto do sensoriamento remoto, o erro pode ser definido como alguma discrepância existente entre as representações na imagem gerada (mapa) e a realidade no campo (Vieira e Mather, 2005). Quando se avalia a qualidade de uma mensuração, é comum a utilização dos termos precisão e exatidão. O termo precisão implica no quanto os valores de uma série de medidas estão próximos uns dos outros, sendo normalmente expressa em termos do desvio padrão ou da variância das medições (Amorim, 2004). Já o termo exatidão expressa o grau de concordância entre uma medição e o valor tomado como referência (Neto, 2006).

Basicamente existem dois componentes importantes, quando da avaliação da qualidade dos produtos cartográficos oriundos do sensoriamento remoto: exatidão posicional, relacionada com a posição relativa e absoluta de feições na carta imagem e seus homólogos no terreno, e a exatidão temática, que se refere à exatidão dos atributos gerados a partir do processo de classificação da imagem (Vieira, 2000).

Atualmente a demanda por mapeamento é crescente nas mais diversas áreas que envolvem projeto ou planejamento. Muitas vezes esses mapeamentos não se encontram atualizados, estão em uma escala inadequada ou simplesmente não existem. Com o desenvolvimento de novas tecnologias, atender a demanda de mapeamento em meio digital tornou-se possível, facilitando a atualização e o gerenciamento dos produtos gerados. Entretanto, pouco controle de qualidade é realizado nesses documentos ocasionando grandes conseqüências, como, por exemplo, um

erro existente e desconhecido pode causar implicações na execução ou no processo de manutenção de um empreendimento (Nero, 2005).

## 1.2 Problema e Motivação

Embora o Brasil possua uma legislação que garante o controle e a qualidade dos produtos cartográficos, o Padrão de Exatidão Cartográfica (PEC) - Decreto 89.817/84, muitos produtos cartográficos não apresentam uma indicação da sua qualidade ao usuário. É importante que os analistas, que diretamente trabalham na área do sensoriamento remoto, estejam cientes da necessidade de realizar o controle da qualidade, pois as imagens orbitais sofrem uma série de distorções que afetam a qualidade do mapeamento.

Uma das formas mais eficientes para representar a exatidão de uma classificação temática está baseada na matriz de confusão (ou matriz de erro), de onde é possível derivar índices de concordância, sendo que estes podem ser expressos como concordância total ou para classes individuais. Apesar desses índices de concordância serem amplamente usados pela comunidade científica, não existem estudos que indiquem como esses índices se comportam, quando da variação do número de classes informacionais e/ou do número de padrões de validação - considerando que esses índices levam em consideração ou o número total de classes ou o número total de padrões de validação; ou mesmo se houver discrepância entre os valores amostrados, para validação, entre as classes informacionais. Porém, essas medidas para representar a exatidão temática, oriundas da estatística não-espacial, não indicam o padrão de erro espacialmente distribuídos na imagem.

A validação é essencial na maior parte dos mapeamentos. O problema é justamente representar a distribuição espacial dos *pixels* erroneamente classificados, que auxiliará na identificação de áreas que possuem baixa precisão dentro da imagem (Vieira, 2000). Conhecer o padrão de distribuição espacial dos erros, numa imagem temática, é indispensável para avaliar a qualidade dos dados em um banco de dados espacial.

### **1.3 Objetivos**

Esta pesquisa propõe-se a avaliar a exatidão posicional e representar espacialmente a exatidão temática em produtos de sistemas imageadores, empregando como metodologias o *Padrão de Exatidão Cartográfica* (PEC) e um mapa de confiabilidade temático para representar a distribuição espacial do erro, realizando uma avaliação quantitativa da exatidão de cada *pixel* na imagem.

Neste trabalho objetivou-se especificamente apresentar ao usuário uma indicação visual da confiabilidade nas atribuições dos rótulos de cada *pixel* na imagem.

Na validação dessa pesquisa explorou-se algumas perguntas específicas:

- Quais os métodos para acessar a exatidão posicional das imagens orbitais?
- É aceitável, em termos cartográficos, a exatidão posicional do produto da classificação de imagens do sensoriamento remoto?
- O número de padrões de validação e/ou de classes informacionais influenciam no cálculo dos índices de exatidão temática?
- A presença de uma classe informacional predominante, na validação, pode afetar o resultado da exatidão temática?
- Quais seriam as possíveis formas de expressar espacialmente a exatidão temática?

### **1.4 Estrutura da dissertação**

Esta dissertação está organizada como segue: no capítulo 2, é apresentada uma revisão bibliográfica sobre exatidão posicional e temática. No capítulo 3, é feita uma caracterização das áreas de estudo, material e metodologia empregados para realizar a avaliação da exatidão posicional; e no capítulo 4 apresenta-se a caracterização da área de estudo, material e metodologia utilizados para a avaliação da exatidão temática tanto na análise não-espacial, quanto na análise espacial. Os resultados e discussão

são apresentados no capítulo 5 e o capítulo 6 corresponde às conclusões a que se chegou e sugestões para trabalhos futuros.

## **2. REVISÃO BIBLIOGRÁFICA**

Segundo Lunetta et al. (1991), os erros associados à aquisição de dados, processo, análise, conversão e apresentação do produto final podem ser calculados a cada passo de sua entrada no sistema, mas na prática o erro só é avaliado na conclusão da análise dos dados. Devido aos gastos com o custo e o tempo para avaliar a exatidão, normalmente na maioria dos projetos avalia-se a exatidão posicional (geométrica) e a exatidão temática, deixando de avaliar outros aspectos importantes como, por exemplo, o método de amostragem, o levantamento de dados de referência e a escala de trabalho (Banko, 1998).

Os tópicos seguintes apresentam uma revisão bibliográfica sobre as formas de expressar a exatidão posicional e a exatidão temática, nos produtos do sensoriamento remoto, que já se encontram difundidas na comunidade científica.

### **2.1 Exatidão Posicional**

O Padrão de Exatidão Cartográfica (PEC), conforme especifica o Decreto 89.817/84, normalmente é aplicado para avaliação da exatidão posicional que garante o controle e a qualidade dos produtos cartográficos no Brasil. Na aplicação do PEC, referentes a dados planimétricos, devem-se

comparar as coordenadas dos pontos facilmente identificados na imagem orbital geometricamente corrigida com as coordenadas homólogas de referência, coletadas em carta ou em levantamento de campo, desde que o erro não seja superior a 1/3 do erro padrão esperado para a classe da carta (Vieira et al., 2002). De posse das coordenadas pode-se avaliar as discrepâncias e calcular as estatísticas das amostras, e a partir destas estatísticas amostrais são realizados testes para avaliação de tendência e exatidão (Galo et al., 2001).

Segundo Merchant (1982), a análise estatística da exatidão planimétrica é composta por duas fases: a análise de tendências, com o objetivo de se averiguar a presença de erros sistemáticos, e a análise de exatidão.

A partir das análises das discrepâncias médias amostrais aplica-se o teste de tendência, onde se essas discrepâncias forem estatisticamente iguais a zero, dentro de certo nível de significância, admite-se a não existência de tendência, e o teste de exatidão baseia-se na comparação da variância das discrepâncias amostrais com os valores pré-estabelecidos no PEC, para determinar em qual classe se enquadra o produto (Galo et al., 2001).

Estes testes são realizados a partir de testes de hipóteses sobre a média e o desvio padrão amostral das discrepâncias obtidas em cada uma das coordenadas planimétricas. Segundo o Decreto 89.817/84, noventa por cento dos pontos testados não podem ter discrepâncias superiores ao padrão de exatidão cartográfica estabelecido.

Os testes de tendência e exatidão representam uma das abordagens que podem ser utilizadas. Vieira et al. (2002) utilizou outros métodos como o Método do Ponto Gerado, o Método das Áreas e o Método dos Retângulos Equivalentes, que são baseados em características, para avaliar a correspondência de uma imagem do sensoriamento remoto com o mapa de referência. Segundo Vieira e Mather (2005), como estes métodos utilizam distâncias relativas entre pontos homólogos, não é necessário realizar o teste de tendência para verificar a presença de erros sistemáticos.

No trabalho de Galo et al. (2001), além dos três métodos utilizados por Vieira et al. (2002), ele também usou o Método Baseado na Aproximação por Splines, proposto por Lugnani (1986).

No Método do Ponto Gerado, é assumido que os pontos iniciais em cada uma das feições são homólogos, e a partir destes gera-se pontos equidistantes em cada feição (Figura 1). Pode-se então comparar as distâncias entre os pontos gerados e aplicar o PEC. Posteriormente avalia-se a exatidão comparando os resultados com o valor do Erro Padrão pré-estabelecido no Decreto 89.817/84 (Marotta e Vieira, 2005). Na aplicação deste método, Galo et al. (2001) fez uma colocação referente à digitalização, onde esta deve ser feita com uma grande taxa de amostragem e que seja sujeita apenas a efeitos aleatórios.



Figura 1: Ilustração do princípio do Método do Ponto Gerado.  
Fonte: Galo et al. (2001).

A Figura 2 apresenta uma ilustração do princípio Método das Áreas que baseia-se no cálculo da área ocasionada pela sobreposição das feições digitalizadas em um documento de referência ou coletadas em campo e as respectivas feições digitalizadas na imagem (Marotta e Vieira, 2005). A partir da área ( $S$ ) ocasionada pela sobreposição das feições e do perímetro ( $p$ ), chega-se a um estimador do erro em área.

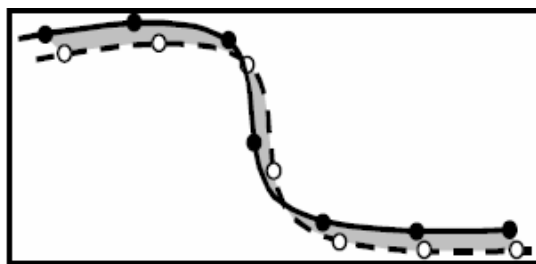


Figura 2: Área entre duas feições.  
Fonte: Galo et al. (2001).

Utilizando o conceito do Método das Áreas, Ferreira e Cintra (1999) propuseram o Método dos Retângulos Equivalentes. Através dos valores da área ( $S$ ) e do perímetro ( $2p$ ) obtidos pelo método das áreas, originados da sobreposição das feições homólogas, admite-se que existe um retângulo de lados  $x_1$  e  $x_2$  que tenha a mesma área ( $S$ ) e perímetro ( $2p$ ). Assim, usam-se as dimensões dos retângulos equivalentes como estimadores das discrepâncias entre as representações de uma dada feição. Segundo Ferreira e Cintra (1999), este método oferece bons resultados e com rapidez.

O Método Baseado na Aproximação por Splines foi apresentado por Lugnani (1986) e tem como princípio obter para todos os pontos de uma das feições homólogas, exceto os extremos, o ponto correspondente na outra feição, ou seja, dado um ponto genérico  $a_i$  na feição  $A$  pesquisa-se os quatro pontos mais próximos de  $a_i$  na feição  $B$ . Desses quatro pontos busca-se os dois pontos ( $b'_{i1}$  e  $b'_{i2}$ ) mais próximos ao ponto  $a_i$ , (Figura 3a). Obtém-se o ângulo entre o eixo  $ox$  e o segmento de reta que liga os pontos  $b'_{i1}$  e  $b'_{i2}$ , para aplicar uma rotação nos quatro pontos e no ponto  $a_i$ , de modo que o eixo coincida com a direção dos pontos  $b'_{i1}$  e  $b'_{i2}$ , gerando-se os pontos  $a''_i$ ,  $b''_{i1}$  e  $b''_{i2}$  (Figura 3b). Daí, obtém-se a equação da *spline* que passa pelos quatro pontos e calcula-se o valor da discrepância ( $W$ ) para o ponto  $i$ . Neste método admiti-se algumas hipóteses básicas:

- as digitalizações apresentam o mesmo grau de generalização;
- todas as discrepâncias entre os pontos das curvas são consideradas como erros;
- as feições contínuas são suaves;
- as feições estão representadas no mesmo referencial; e

- ambas as seqüências de pontos digitalizados são livres de erros grosseiros e de grandes erros sistemáticos.

Segundo Lugnani (1986), os resultados obtidos são compatíveis com os métodos convencionais.

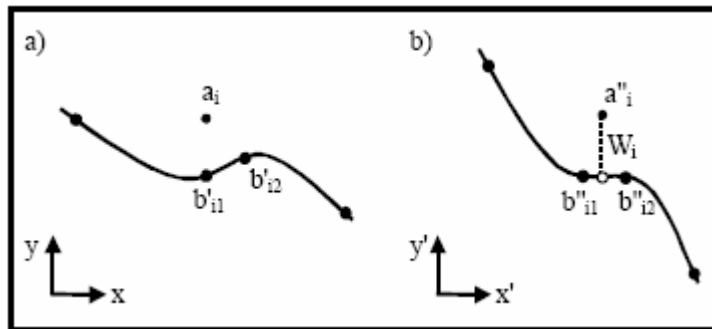


Figura 3: Ponto  $a_i$  e os quatro pontos mais próximos (a) antes da rotação e (b) após a rotação.

Fonte: Galo et al. (2001).

## 2.2 Exatidão Temática

Avaliar a exatidão obtida no processo de análise dos dados do sensoriamento remoto é uma etapa fundamental, pois os usuários desses produtos necessitam saber sobre a confiabilidade dos dados derivados de mapas obtidos por meio desta ferramenta.

Congalton (1991) sugere que os seguintes fatores sejam considerados: fontes de erro, esquema de amostragem (esquema de amostra, número de amostras e unidade de amostra, que é a coleção de dados de campo) e tamanho da amostra. Cada um destes fatores fornece informação essencial para a qualidade da avaliação da exatidão dos produtos do sensoriamento remoto.

### 2.2.1 Esquemas de Amostragem

Como não é possível conferir todos os *pixels* das áreas mapeadas, a amostragem é o meio pelo qual a exatidão do mapeamento pode ser avaliada. Os esquemas de amostragem que geralmente têm sido usados são os que incluem amostragem aleatória simples, amostragem aleatória estratificada, amostragem sistemática e amostragem sistemática não alinhada.

Numa amostra aleatória simples a premissa é de que cada unidade de amostra na área de estudo tem a mesma chance de ser escolhida para compor a amostra, fornecendo dados representativos escolhidos sem influências e satisfazendo exigências fundamentais de inferência estatística (Congalton e Green, 1999).

No caso da amostragem aleatória estratificada a área de estudo é dividida em extratos e em seguida é selecionada uma amostra aleatória de cada extrato, requerendo um pouco de conhecimento *a priori* sobre o local (Vieira, 2000).

Na amostragem sistemática as amostras são tomadas uniformemente, tendo como desvantagem que cada unidade na área não tem uma chance igual de ser incluída na amostra (Liang, 2004). Na amostragem sistemática normalmente a primeira amostra é selecionada aleatoriamente e cada amostra sucessiva é tomada em um intervalo específico (Congalton e Green, 1999).

Por fim, amostragem sistemática estratificada não alinhada combina as vantagens da aleatoriedade e da estratificação, com aspectos da amostragem sistemática, evitando as possibilidades de influências devido à presença de periodicidades (Banko, 1998).

A escolha da técnica de amostragem dependerá de vários fatores, inclusive o tamanho da área de estudo, o tipo e distribuição das características a serem mapeadas, e os custos para aquisição dos dados de verificação (Banko, 1998).

Congalton (1991) sugeriu, para uso no sensoriamento remoto, uma combinação de amostragem estratificada e aleatória simples. Congalton e Green (1999) também verificaram experimentalmente que a amostragem aleatória estratificada pode ser usada em substituição a amostragem aleatória simples. A amostragem estratificada pode ser feita junto com a coleção de dados de treinamento em uma fase inicial do projeto. Depois do resultado da primeira classificação, a amostragem aleatória estratificada completa a coleção de dados necessários para avaliação da exatidão. Segundo Banko (1998), uma aproximação de amostragem aleatória deveria ser escolhida com a finalidade de usar o coeficiente Kappa na análise da matriz de confusão.

Brites et al. (1996) fez uma comparação entre as classificações feitas utilizando as estratégias de amostragem estratificada, sistemática estratificada, sistemática e aleatória. Ele aplicou o teste estatístico Z para medir a significância das diferenças entre classificações por estratégia de amostragem. Comparando entre si todas as classificações geradas dentro de cada procedimento de amostragem, ele observou que o teste foi mais sensível às diferenças nas classificações para a amostragem sistemática e amostragem estratificada, o que pode indicar que quando o analista tem pouca interferência no processo amostral os resultados talvez não sejam bons.

Uma outra análise feita por Brites et al. (1996) foi a comparação das amostragens entre si, onde observou que só ocorreu diferença significativa quando foram comparados os procedimentos de amostragens sistemática estratificada, sistemática e aleatória com o procedimento que utilizou a amostragem estratificada. O mesmo autor ainda relata que o tipo de amostragem mostra ser o principal fator que contribui para a superestimação dos índices de exatidão empregados para verificar a exatidão da classificação.

### **2.2.2 Tamanho da Amostra**

O tamanho da amostra é estabelecido de forma a minimizar os custos com amostras de campo, mas também satisfazendo a exigência de um tamanho mínimo de amostra que seja representativo e estatisticamente confiável, as quais serão usadas para a validação da classificação. Em geral, quanto maior o tamanho da amostra e representatividade delas em relação as classes informacionais, maior a confiança que se pode ter nas avaliações baseadas naquela amostra (Dicks e Lo, 1990).

Dependendo do objetivo da avaliação da exatidão o número de amostras pode ser calculado por métodos diferentes. Se necessitar somente de uma avaliação booleana (certo/errado), então a *distribuição binomial* pode ser usada para calcular o tamanho da amostra (Hay, 1979; Rosenfield et al., 1982). Mas, se além da avaliação booleana o objetivo também for avaliar o erro entre todas as classes em questão, então uma *distribuição*

*multinomial* deveria ser usada para calcular o tamanho da amostra (Rosenfield, 1982).

Para o cálculo do tamanho da amostra, a maioria dos pesquisadores utiliza equações baseadas na aproximação normal à *distribuição binomial*, e estas equações geralmente são baseadas na proporção de amostras corretamente classificadas e um erro permissível.

O método da *distribuição binomial* é aceitável quando se deseja selecionar o número total de observações a ser amostrada, mas ele não foi projetado para selecionar um tamanho de amostra que produza uma matriz de confusão confiável (Congalton, 1991). Em dados do sensoriamento remoto lida-se com grande quantidade de *pixels*, o que inviabiliza o uso do método tradicional para determinar o tamanho da amostra. Segundo recomendação de Congalton e Green (1999) usar a *distribuição multinomial* seria uma alternativa.

Em determinados casos seria interessante obter uma maior quantidade de amostras para as classes informacionais que apresentem maior variabilidade e ocupam maior área, e reduzir o número de amostras em categorias que apresentam pouca variação (Vieira, 2000).

### **2.2.3 Matriz de Confusão (ou Matriz de Erro)**

No processo de análise dos dados do sensoriamento remoto, um passo fundamental é a avaliação da exatidão temática. Os usuários necessitam saber quão confiáveis são os dados provenientes dos mapas temáticos derivados da classificação de um produto do sensoriamento remoto, e através da matriz de confusão é possível derivar medidas e consequentemente verificar erros oriundos do processo de atribuição dos pixels a determinadas classes (Vieira, 2000).

O método padrão para avaliação da exatidão temática atualmente tem sido índices derivados da matriz de confusão. A matriz de confusão fornece a base para descrever a exatidão da classificação e caracterizar os erros, ajudando a refinar a classificação. De uma matriz de confusão podem ser derivadas várias medidas de exatidão da classificação, sendo a exatidão global uma das mais conhecidas (Foody, 2002).

A matriz de confusão, também conhecida como matriz de erro, é formada por um arranjo quadrado de números dispostos em linhas e colunas que expressam o número de unidades de amostras de uma categoria particular relativa – inferida por um classificador (ou regra de decisão), comparado com a categoria atual verificada no campo (Congalton, 1991). Normalmente, ao longo das colunas representa-se o conjunto de dados de referência, que é comparado com os dados do produto da classificação, sendo estes representados ao longo das linhas. A Tabela 1 apresenta a representação de uma matriz de confusão. Os elementos da diagonal principal (em negrito) indicam o nível de acerto (ou concordância) entre os dois conjuntos de dados e  $n$  é o número total de amostras coletadas.

Tabela 1 – Representação matemática de uma matriz de confusão.

Classificação	Dados de referência			Total nas linhas $x_{i+}$
	1	2	c	
1	<b><math>x_{11}</math></b>	$x_{12}$	$x_{1c}$	$x_{1+}$
2	$x_{21}$	<b><math>x_{22}</math></b>	$x_{2c}$	$x_{2+}$
c	$x_{31}$	$x_{32}$	<b><math>x_{3c}</math></b>	$x_{c+}$
Total nas colunas $x_{+j}$	$x_{+1}$	$x_{+2}$	$x_{+c}$	$n$

Fonte: Adaptada de Bernardes (2006).

#### 2.2.4 Medidas e Índices Derivados da Matriz de Confusão

As medidas derivadas da matriz de confusão são: a exatidão global, exatidão de classe individual, exatidão de produtor, exatidão de usuário e índice Kappa, entre outros. A exatidão global ( $G$ ) é calculada dividindo a soma da diagonal principal da matriz de erros  $x_{ij}$ , pelo número total de amostras coletadas  $n$ . A distribuição da exatidão ao longo das categorias individuais não é apresentada na exatidão global, entretanto a exatidão de uma categoria individual é obtida através da divisão do número total de amostras classificadas corretamente naquela categoria pelo número total de amostras daquela categoria. Congalton e Green (1999) descrevem os cálculos associados com estas medidas.

A exatidão de produtor e de usuário são maneiras de representar a exatidão de uma categoria individual. A exatidão de produtor indica a probabilidade da amostra ser classificada corretamente, sendo realmente uma medida de erro de omissão, e a exatidão de usuário indica a

probabilidade que um *pixel* classificado na imagem de fato representa aquela categoria no campo (Vieira, 2000). Assim, a exatidão de produtor na

classe  $(j) = \frac{x_{jj}}{\sum_{i=1}^c x_{ij}}$  e a exatidão de usuário na classe  $(j) = \frac{x_{jj}}{\sum_{i=1}^c x_{ji}}$ . Para

exemplificar, tomamos uma matriz de confusão apresentada na Tabela 2 a seguir.

Tabela 2 – Matriz de confusão.

Classes	Dados de referência			total	Usuário (%)
	1	2	3		
1	303	3	12	318	95,3
2	22	305	9	336	90,8
3	37	19	440	496	88,7
total	362	327	461	1150	
Produtor (%)	83,7	93,3	95,4		

A análise do índice Kappa é uma técnica multivariada discreta usada na avaliação da exatidão temática. Segundo Vieira (2000), o coeficiente Kappa ( $K$ ) é uma medida da concordância real (indicado pelos elementos diagonais da matriz de confusão) menos a concordância por chance (indicado pelo produto total da linha e coluna, que não inclui entradas não reconhecidas), ou seja, é uma medida do quanto à classificação está de acordo com os dados de referência. Além do coeficiente Kappa, pode ser proveitoso usar o coeficiente Kappa condicional ( $K_i$ ) para calcular o acerto dentro de uma categoria individual.

Congalton (1991) relata que o uso do coeficiente Kappa é satisfatório na avaliação da exatidão de uma classificação temática, pelo fato de levar em consideração toda a matriz de confusão no seu cálculo, inclusive os elementos de fora da diagonal principal, os quais representam as discordâncias na classificação, diferentemente da exatidão global, por exemplo, que utiliza somente os elementos diagonais (concordância real).

Embora o coeficiente Kappa seja muito utilizado na avaliação da exatidão de mapeamento, não existe uma fundamentação teórica para recomendar quais os níveis mínimos aceitáveis deste coeficiente numa classificação. Entretanto, a Tabela 3 apresenta níveis de desempenho da

classificação para o valor de Kappa obtido, normalmente aceitos pela comunidade científica.

Tabela 3 – Índice Kappa e o correspondente desempenho da classificação.

<b>Índice Kappa</b>	<b>Desempenho</b>
$< 0$	Péssimo
$0 < k \leq 0,2$	Ruim
$0,2 < k \leq 0,4$	Razoável
$0,4 < k \leq 0,6$	Bom
$0,6 < k \leq 0,8$	Muito Bom
$0,8 < k \leq 1,0$	Excelente

Fonte: Fonseca (2000).

Ao redor do valor de Kappa podem ser calculados intervalos de confiança usando a variância da amostra (*var*) e o fato de que a distribuição estatística do Kappa é normalmente assintótica. Congalton e Green (1999) sugerem meios de testar a significância estatística do Kappa para uma única matriz de confusão, através da variância de Kappa, a fim de determinar se o nível de acerto da classificação e os dados de referência são significativamente maior que zero. Semelhantemente, através de dois índices de exatidão, tomando duas classificações independentes, podemos testar se dois classificadores independentes são significativamente diferentes.

Com relação ao coeficiente Kappa, Foody (1992) observou que o grau de concordância por chance poderia estar sendo superestimado, pelo fato de incluir também a concordância real, e por causa disso a magnitude de Kappa não refletiria a concordância presente na classificação, apenas descontada a casualidade. Na tentativa de corrigir essa deficiência no cálculo do índice Kappa, Ma e Redmond (1995) propuseram um outro índice para a medição da exatidão temática da classificação, o índice Tau (*T*).

De acordo com Ma e Redmond (1995), o índice Tau que fornece uma medida quantitativa relativamente precisa e intuitiva sobre a exatidão da classificação. Este índice é uma medida da concordância real (indicado pelos elementos diagonais da matriz de confusão) menos a concordância casual (indicado pela distribuição marginal dos dados de referência e as probabilidades *a priori* para cada classe).

Ma e Redmond (1995) sugeriram denotar  $T_e$  quando Tau for utilizado numa classificação baseada em equi-probabilidade entre as categorias e  $T_p$  quando for utilizado em uma classificação baseada em probabilidades desiguais. As probabilidades desiguais podem ser obtidas por levantamento prévio em campo, por estimativas através de fotografias aéreas ou por uma decisão arbitrária pela importância da classe.

O mesmo teste sugerido por Congalton e Green (1999) para testar a significação estatística do Kappa, pode ser aplicado para o índice Tau (Ma e Redmond, 1995). Segundo Lingnau e Antunes (2003) os coeficientes Kappa e Tau podem apresentar valores entre -1 e +1, sendo interpretados da seguinte forma: quando o valor do coeficiente for 1 significa que a classificação foi completamente correta (100% de concordância); quando o valor for zero, indica que não houve nenhum acordo entre a classificação e os dados de referência; quando o valor for menor que zero significa que não houve discriminação entre as categorias; e quando o valor for maior que zero significa que a concordância real foi maior do que a concordância por chance (ou aleatória, no caso do índice Tau).

Segundo Vieira (2000), quando utilizamos algum procedimento de cálculo estatístico, esperamos obter um conjunto de intervalos de confiança. E se for o caso de se utilizar a *distribuição binomial*, pode-se facilmente avaliar intervalos de confiança.

De acordo com Mather (1999), os limites de confiança podem ser interpretados da seguinte forma: se um número muito grande de amostras de tamanho  $n$  é escolhido e se a verdadeira proporção  $P$  de resultados bem sucedidos é  $P_T$ , então 95% de todos os valores de amostra estarão entre o limite inferior ( $P_L$ ) e superior ( $P_U$ ) ao redor de  $P_T$ . Os valores de limites de confiança superior e inferior dependem (i) do nível de probabilidade empregado e (ii) o tamanho da amostra  $n$ . Os limites de confiança tornam-se amplos com o aumento do nível de probabilidade para 100% e/ou quando o tamanho da amostra  $n$  se torna menor. Mather (1999) também fornece a fórmula para o cálculo do limite de confiança inferior baseado no valor da exatidão da classificação.

Estes limites de confiança também podem ser derivados da matriz de confusão quando o mapa marginal de proporções  $\pi_j$  também é conhecido. O

procedimento para gerar limites de confiança da matriz de confusão pode ser encontrado em Congalton e Green (1999).

Segundo Vieira (2000), a classificação pode sofrer influência de vários fatores (como por exemplo, falha humana, software, sensor, etc.) e deve-se lembrar que a derivação de um intervalo de confiança para uma classificação representa tais combinações de influências.

### **2.2.5 Distribuição Espacial do Erro**

Apesar dos índices derivados da matriz de confusão não considerarem a distribuição espacial do erro implícita ou explicitamente, existem técnicas de representação e visualização da confiabilidade espacial - *pixel a pixel* - que fornecem uma avaliação quantitativa da exatidão na imagem temática. Vieira (2000) introduziu algumas dessas técnicas de representação de um mapa de confiabilidade temático que serão descritas a seguir.

Segundo o autor, pode-se usar a Distância de Mahalanobis ( $MD^2$ ) para gerar uma imagem de distância, em escala de cinza (0-255), que permitirá visualizar as discrepâncias de *pixels* individuais dos meios das classes para as quais foram atribuídos, onde os *pixels* mais escuros indicam que são mais prováveis de serem classificados corretamente por estarem próximos ao centróide da classe, e já os *pixels* em tons de cinza altos são mais prováveis de serem classificados erroneamente.

Uma outra forma de espacializar os erros de uma classificação é através da comparação das imagens temáticas com os respectivos mapas de referência (ou verdade de campo), que resultaria em uma imagem binária do erro, sendo que o *pixel* que foi rotulado corretamente recebe o valor "0" e o rotulado erroneamente recebe o valor "1". Ainda a partir desta comparação, pode-se gerar uma imagem de diferença, onde duas imagens temáticas diferentes (geradas por métodos diferentes) podem ser comparadas com o mapa de referência.

Vieira (2000) propôs uma forma inovadora de representar um mapa de confiabilidade temático. Nesta proposta, a Distância de Mahalanobis é transformada em estimativas de probabilidades *a posteriori*  $P_i$  (probabilidade

de um *pixel* pertencer à classe  $i$ ). De posse dessas probabilidades, gera-se um novo valor ( $NV$ ) através da multiplicação do rótulo da classe nomeada por 10 e somando a probabilidade calculada pela Distância de Mahalanobis multiplicada por 10. Na representação temática da confiança, uma cor separada é atribuída a cada classe na imagem e dentro de cada classe são separados níveis de precisão onde estes níveis recebem tons separados da cor alocada àquela classe. Essa representação permite ao usuário ter uma visualização quantitativa das precisões distribuídas dentro da imagem classificada.

### **3. EXATIDÃO POSICIONAL**

#### **3.1 Caracterização das áreas de estudo**

Para comparar os procedimentos de verificação da exatidão posicional, bem como a influência da resolução espacial dos sensores orbitais em questão (Ikonos, Quickbird, Landsat TM, CBERS e MODIS), utilizou-se três áreas de estudo: Viçosa-MG, Uberaba-MG e a área de abrangência do Escritório de Desenvolvimento Regional (EDR) de Limeira-SP.

As imagens Ikonos e Quickbird foram utilizadas na primeira área de estudo correspondente à Viçosa no Estado de Minas Gerais (Figura 4). Segundo a Prefeitura de Viçosa, o município está situado na região da Zona da Mata, entre as Serras da Mantiqueira, do Caparaó e da Piedade, portanto uma cidade de relevo acidentado com 85% de montanhas, 12% ondulado e 3% plano. Cidade da cultura e do conhecimento devido a presença da Universidade Federal de Viçosa e de mais três outras instituições de ensino superior; a vida da cidade gira em torno da educação.

A área de estudo no município de Viçosa está delimitada pelas seguintes coordenadas Geográficas:

- Canto superior esquerdo: 20°45'25,58" S, 42°52'20,45" W;
- Canto inferior direito: 20°45'56,88" S, 42°51'48,45" W.

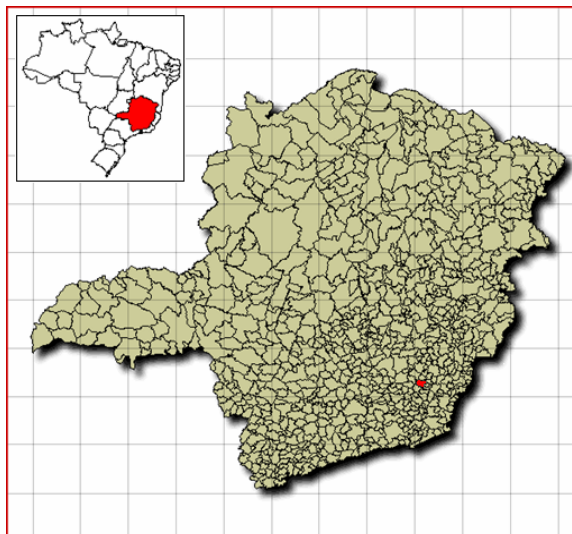


Figura 4: Localização de Viçosa - MG.

Fonte: IBGE, 2007.

As imagens Landsat TM e CBERS são da segunda área de estudo, correspondendo a região de Uberaba no Estado de Minas Gerais (Figura 5). De acordo com a Prefeitura de Uberaba, o município está localizado na micro-região do Triângulo Mineiro. Uberaba é um centro universitário, possuindo várias instituições de ensino superior e um Centro de Ensino Federal e Tecnológico. O município é também uma das maiores bacias arqueológicas do Brasil, principalmente no distrito denominado Peirópolis, onde são encontrados fósseis de mais de 85 milhões de anos. A cidade também é conhecida como a capital mundial do gado Zebu, espécie que foi introduzida por criadores da cidade no final do século XIX.

A área de estudo no município de Uberaba está delimitada pelas seguintes coordenadas Geográficas:

- Canto superior esquerdo: 19°39'22,77" S, 48°00'08,51" W;
- Canto inferior direito: 19°48'07,89" S, 47°44'21,34" W.

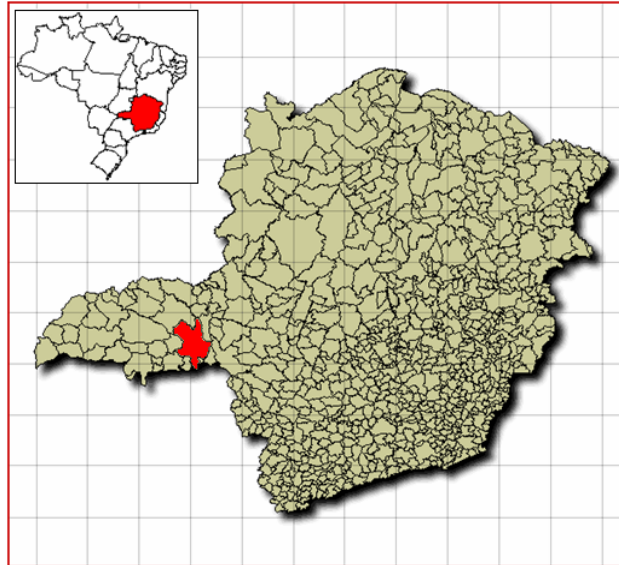


Figura 5: Localização de Uberaba - MG.

Fonte: IBGE, 2007.

A imagem MODIS foi utilizada para a terceira área de estudo que corresponde à abrangência do Escritório de Desenvolvimento Regional (EDR) de Limeira, uma das divisões administrativas da Coordenadoria de Assistência Técnica Integral (CATI) do Estado de São Paulo (Figura 6), compreendendo os seguintes municípios: Porto Ferreira, Pirassununga, Analândia, Santa Cruz da Conceição, Araras, Cordeirópolis, Corumbataí, Ipeúna, Iracemápolis, Itirapina, Leme, Limeira, Rio Claro e Santa Gertrudes. Esta área situa-se a leste no Estado de São Paulo, cuja área estimada chega a 497.598,3 ha. Para Arce et al. (2005), a agricultura é a principal atividade da região, onde cultiva-se principalmente cana de açúcar, frutas cítricas, milho e soja. Outra atividade importante é o gado bovino, para o qual são destinadas áreas importantes para os pastos. Finalmente, os solos nesta área têm um uso de floresta com plantações importantes de *pinus* e eucalipto, para industrialização da celulose.

A área da EDR de Limeira está delimitada pelas seguintes coordenadas Geográficas:

- Canto superior direito: 22°45'33,72" S, 47°59'07,21" W;
- Canto inferior esquerdo: 21°44'47,09" S, 47°09'20,76" W.

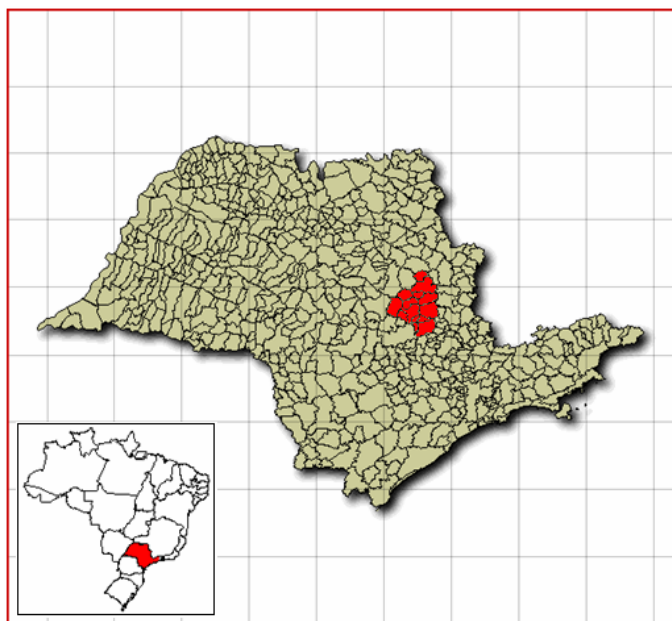


Figura 6: Localização da área de abrangência da EDR de Limeira - SP.

Fonte: IBGE, 2007.

## 3.2 Materiais

### 3.2.1 Dados Orbitais

Para a região de Viçosa-MG foram utilizadas as imagens dos satélites Ikonos II e Quickbird (Figura 7). A imagem Ikonos II corresponde a um trecho da imagem comercializada para o projeto “Modelagem integrada de processos Hidrossedimentológicos e Hidrológicos e a avaliação de cenários de desenvolvimento em bacia da Zona da Mata de Minas Gerais”. Essa imagem Ikonos possuía uma composição colorida RGB (vermelho: 0,63-0,69  $\mu\text{m}$ ; verde: 0,52-0,60  $\mu\text{m}$ ; azul: 0,45-0,52  $\mu\text{m}$ ), com resolução espacial de 4m, mas foi comercializada já reamostrada com a banda pancromática (0,45-0,90 $\mu\text{m}$ ), passando de 4m para 1m de resolução espacial. A imagem Quickbird, possuía apenas a banda pancromática (0,45-0,90  $\mu\text{m}$ ) com resolução espacial de 61cm.

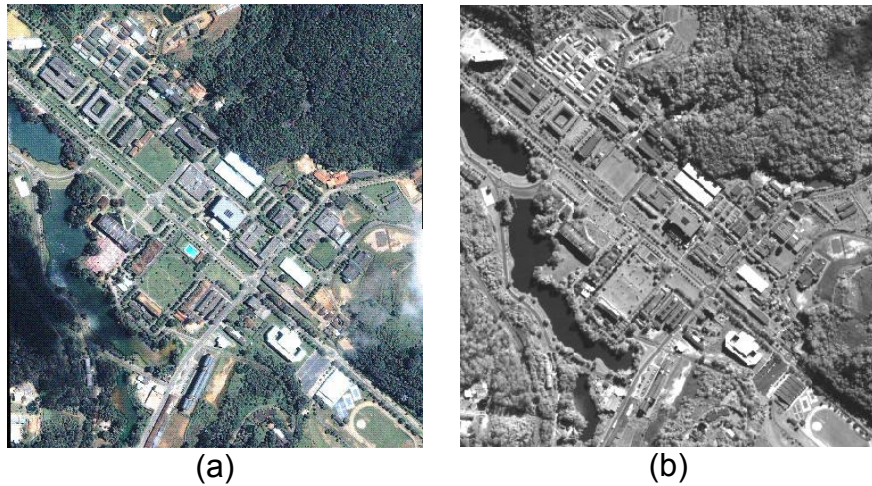


Figura 7: (a) Imagem Ikonos II; (b) Imagem Quickbird.

Da região de Uberaba-MG, as imagens Landsat TM e CBERS utilizadas possuíam uma única banda espectral ( $0,45-0,52 \mu\text{m}$ , ambas) com resolução espacial de 30m e 20m respectivamente (Figura 8). Para coleta dos pontos de controle e dos pontos homólogos utilizou-se como referência as cartas do serviço militar na escala 1:25.000 cujo Datum horizontal é o Córrego Alegre, formato digital, cedidas pelo LabGeo da Universidade de Uberaba.

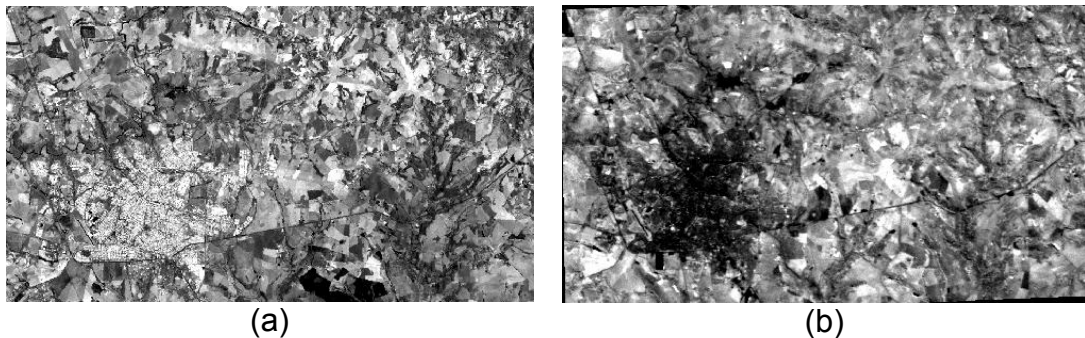


Figura 8: (a) Imagem Landsat; (b) Imagem CBERS.

A imagem MODIS, da região de Limeira-SP, foi cedida pelo Instituto Nacional de Pesquisas Espaciais - INPE (Figura 9). Uma característica das imagens do sensor MODIS é que elas já vêm georreferenciadas e corrigidas para os efeitos atmosféricos. Estes dados são disponibilizados na forma de produtos, e para esta pesquisa utilizou-se o produto MOD13, com as respectivas bandas: azul ( $0,459-0,479 \mu\text{m}$ ), vermelho ( $0,620-0,670 \mu\text{m}$ ),

infravermelho próximo (0,841-0,876  $\mu\text{m}$ ), além de uma banda localizada na faixa do infravermelho médio (1,230-1,250  $\mu\text{m}$ ). As bandas do vermelho e infravermelho próximo possuem resolução de 250m e as outras duas bandas foram fornecidas reamostradas de 500m para 250m. Devido a baixa resolução espacial das imagens MODIS não foi possível utilizar cartas topográficas para coleta dos pontos homólogos, pois não foi possível identificar feições homólogas na carta e na imagem, portanto utilizou-se uma carta imagem Landsat também cedida pelo INPE.

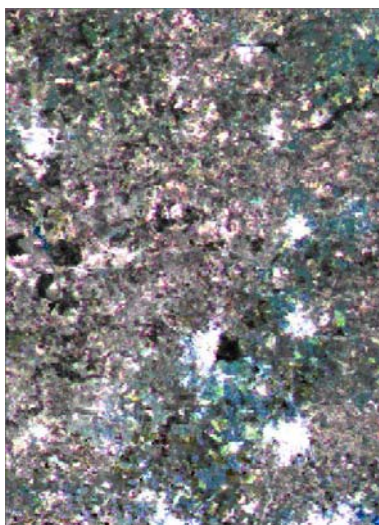


Figura 9: Imagem MODIS.

### **3.2.2 Ferramentas Computacionais**

Foi empregado levantamento topográfico para coleta de pontos de controle (para correção geométrica) e pontos homólogos em campo para realizar a verificação da exatidão posicional nas imagens Ikonos e Quickbird. Utilizou-se uma estação total Topcon GTS 212 para coleta dos pontos e o software Topograph para o processamento, ambos do Departamento de Engenharia Civil da UFV. As coordenadas foram calculadas no sistema WGS84 com base nos pontos homologados pelo IBGE no campus da UFV.

O software ERDAS Imagine 8.5 foi utilizado para manipulação das imagens e para realizar o processo de correção geométrica das mesmas utilizando o método de reamostragem do vizinho mais próximo com polinômio de primeiro grau. Todas as imagens corrigidas estão referenciadas ao sistema de projeção UTM e zona 23 sul. A Tabela 4 apresenta o erro

médio quadrático (RMS) em *pixels* associado a cada imagem após a realização da correção geométrica; e o número de pontos de controle utilizados na correção. A imagem MODIS já vem georreferenciada, logo não foi necessário realizar correção geométrica.

Tabela 4 – Erro médio quadrático (RMS) pós correção geométrica.

<b>Imagem</b>	Ikonos	Quickbird	Landsat	CBERS
<b>RMS</b>	0,4834	0,4632	0,4855	0,4734
<b>Nº. pontos controle</b>	5	5	7	7

Utilizou-se o software Idrisi Kilimanjaro para manipulação do banco de dados geográfico de Uberaba contendo as informações: rodovias, ferrovias, hidrografia, curvas de nível e lagoas. O software AutoCad 2000 foi utilizado na geração de pontos equidistantes, cálculo das áreas e cálculo dos perímetros, dados estes empregados nos métodos baseados em características, para avaliar a correspondência de uma imagem do sensoriamento remoto com o mapa de referência.

### 3.3 Métodos empregados na avaliação da exatidão posicional

O Decreto-lei nº. 89.817 de 1984, nos artigos 8º a 11º, diz respeito ao PEC que é definido pelas normas de regulamentação e classificação de produtos cartográficos, tratando da classificação de uma carta quanto à sua exatidão. Na aplicação do PEC, utilizou-se pontos bem distribuídos na imagem ( $X_i, Y_i$ ) e seus homólogos no documento de referência ( $X_r, Y_r$ ). De posse das coordenadas pôde-se avaliar as discrepâncias ( $\Delta_E = X_r - X_i, \Delta_N = Y_r - Y_i, \Delta_P = \sqrt{(\Delta_E^2 + \Delta_N^2)}$ ) e calcular as estatísticas das amostras como a média ( $\overline{\Delta_E}, \overline{\Delta_N}$ ), desvio padrão da direção ( $S_E, S_N$ ) e o desvio padrão da posição ( $S_P$ ).

É importante ressaltar a questão do número de pontos necessários para realizar uma análise confiável, porque não é apropriado ter um número de pontos que possa fornecer uma análise não confiável ou então ter um número de pontos excessivos fornecendo uma análise confiável mas de custo inviável. Merchant (1982) sugere que vinte pontos bem distribuídos por

todos os quadrantes da carta (ou da imagem que se esteja avaliando) seria um número mínimo aceitável.

De posse das estatísticas amostrais, foram realizados testes para avaliação de tendência e exatidão. O teste de tendência é aplicado usando a hipótese nula, onde se as discrepâncias médias amostrais forem estatisticamente iguais a zero (não significativos), dentro de um certo nível de significância ( $\alpha$ ), admite-se a não existência de tendência (ou erro sistemático). Para a análise de tendência normalmente utiliza-se o teste *t* de *Student*, a certo nível de confiança ( $1-\alpha$ ) e com grau de liberdade ( $n-1$ ), onde  $n$  corresponde ao número de amostras coletadas. A imagem orbital está livre de erros sistemáticos nas coordenadas  $E$  e  $N$  se os valores de  $t_E$  e  $t_N$  calculados forem inferiores ao valor limite  $t_{n-1,\alpha/2}$  tabelado. Os valores de  $t_E$  e  $t_N$  são calculados a partir das seguintes equações:

$$|t_E| = \frac{\overline{\Delta_E} \times \sqrt{n}}{S_E} \quad (1)$$

$$|t_N| = \frac{\overline{\Delta_N} \times \sqrt{n}}{S_N} \quad (2)$$

Na análise de exatidão compara-se a variância das discrepâncias amostrais com os valores pré-estabelecidos do Erro Padrão (EP) no Decreto-lei nº. 89.817/84, com objetivo de verificar em qual classe o produto analisado se enquadra segundo a Lei. O teste de exatidão é aplicado usando a hipótese nula, onde se a variância da discrepância amostral do erro de posição do ponto ( $S_p^2$ ) for estatisticamente igual a variância obtida a partir do Erro Padrão ( $\sigma_p^2$ ), dentro de um certo nível de significância ( $\alpha$ ), admite-se que há exatidão posicional.

O teste estatístico utilizado para a análise de exatidão foi o teste *Qui-quadrado* ( $\chi^2$ ), onde a precisão nas imagens orbitais são atendidas quando o valor calculado de  $\chi^2_{P\ n-1}$  for inferior ao valor limite  $\chi^2_{\alpha,\ n-1}$  tabelado. O valor de  $\chi^2_{P\ n-1}$  é calculado a partir da seguinte equação:

$$\chi^2_{P\ n-1} = (n-1) \times (S_p^2 / \sigma_p^2) \quad (3)$$

Onde  $\sigma_p$  varia em função da escala da carta sendo calculado através da equação:  $\sigma_p = EP/\sqrt{2}$ , segundo a NBR13133 de 1994.

Os testes de tendência e exatidão representam uma das abordagens que podem ser utilizadas para avaliar a exatidão posicional. Existem outros métodos como o Método do Ponto Gerado, o Método das Áreas e o Método dos Retângulos Equivalentes, que são baseados em características (feições), para avaliar a correspondência de uma imagem do sensoriamento remoto com o documento de referência. Esses métodos foram aplicados nas imagens Landsat e CBERS para apresentar o uso dessas metodologias. Para isto, digitalizou-se feições homólogas no documento de referência e na imagem.

No Método do Ponto Gerado, admite-se que os pontos iniciais em cada uma das feições são homólogos, e a partir destes gera-se pontos equidistantes em cada uma das feições. Em seguida, calcula-se as distâncias euclidianas ( $D$ ) entre os pontos homólogos gerados e aplica-se o PEC, utilizando para análise de exatidão o teste *Qui-quadrado*, aplicando a hipótese nula sobre a variância do conjunto de distâncias euclidianas obtidas. Analisa-se a exatidão comparando os resultados com os valores pré-estabelecidos do Erro Padrão (EP).

O Método das Áreas baseia-se no cálculo da área ocasionada pela sobreposição das feições digitalizadas em um documento de referência (ou coletadas em campo) e as respectivas feições digitalizadas na imagem. A partir do somatório de todas as áreas ( $S$ ), ocasionada pela sobreposição das feições, e do somatório dos perímetros ( $p$ ), chega-se a um estimador do erro em área ( $e_a$ ), dado pela Equação (5). Este valor representa uma estimativa das discrepâncias entre as representações de uma dada feição.

$$e_a = \frac{S}{p} \quad (5)$$

Utilizando o conceito do Método das Áreas, Ferreira e Cintra (1999) propuseram o Método dos Retângulos Equivalentes. Através dos valores da área ( $S$ ) e do perímetro ( $2p$ ), obtidos pelo método das áreas, originados da sobreposição das feições homólogas, admite-se que existe um retângulo de

lados  $x_1$  e  $x_2$  que tenha a mesma área ( $S$ ) e perímetro ( $2p$ ), resultando nas seguintes relações:

$$\begin{cases} x_1 \times x_2 = S \\ 2x_1 + 2x_2 = 2p \end{cases} \quad (6)$$

Assim, usa-se as dimensões dos retângulos equivalentes como estimadores das discrepâncias entre as representações de uma dada feição.

## 4. EXATIDÃO TEMÁTICA

### 4.1 Métodos aplicados para análise não-espacial da exatidão temática

Objetivando analisar o comportamento dos índices de exatidão global, Kappa e Tau, foram simuladas várias matrizes de confusão (não foram realizadas classificações propriamente ditas) variando:

- 1º - O número de classes informacionais para cada classificação,
- 2º - O total de amostras utilizados para gerar a matriz de confusão, e,
- 3º - O total de pixels e ainda ocorrendo a presença de classe informacional predominante entre os valores amostrados, para validação.

Estas variações foram realizadas para verificar qual seria o comportamento dos três índices dentro dos cinco níveis de desempenho da classificação (níveis apresentados na Tabela 3, seção 2.2.4). Os valores dos índices Kappa, Global e Tau, obtidos das matrizes de confusão simuladas para o experimento são apresentados nos Apêndices A, B, C, D e E.

Para realizar o experimento, foram calculados os valores dos índices para cada matriz de confusão, através das suas respectivas equações, para posteriormente realizar a análise do comportamento dos mesmos. A Tabela 5 apresenta uma matriz de confusão para exemplificar os cálculos dos índices a seguir.

Tabela 5 – Matriz de confusão utilizada para os exemplos a seguir.

Classes	Dados de referência							total
	1	2	3	4	5	6	7	
1	125	5	3	5	0	1	0	139
2	0	83	2	11	9	0	7	112
3	13	12	196	15	7	9	0	252
4	29	6	72	71	4	12	17	211
5	0	5	0	0	174	18	22	219
6	8	0	5	2	41	91	3	150
7	0	2	8	1	0	2	204	217
total	175	113	286	105	235	133	253	1300

A exatidão global é calculada dividindo a soma da diagonal principal da matriz de erros  $x_{ii}$ , pelo número total de amostras coletadas  $n$ . Para a matriz de confusão da Tabela 5 temos que a exatidão global é igual a:

$$G = \frac{\sum_{i=1}^c x_{ii}}{n} = \frac{(125 + 83 + 196 + 71 + 174 + 91 + 204)}{1300} * 100 = 72,6\% \quad (7)$$

O coeficiente Kappa pode ser calculado através da Equação (8). Lembrando que para os casos onde existem padrões que não foram reconhecidos na classificação, estes não entram no cálculo do Kappa quando se calcula o total nas colunas ( $x_{+i}$ ).

$$\sum_{i=1}^c x_{ii} = (125 + 83 + 196 + 71 + 174 + 91 + 204) = 944 ,$$

$$x_{i+} = \sum_{j=1}^c x_{ij} = (125 + 5 + 3 + 5 + 0 + 1 + 0) = 139 \text{ de igual modo para as demais}$$

linhas, e

$$\begin{aligned} \sum_{i=1}^c x_{i+} \times x_{+i} &= (139 \times 175) + (112 \times 113) + (252 \times 286) + (211 \times 105) + (219 \times 235) \\ &+ (150 \times 133) + (217 \times 253) = 257524 \end{aligned}$$

$$K = \frac{n \sum_{i=1}^c x_{ii} - \sum_{i=1}^c x_{i+} x_{+i}}{n^2 - \sum_{i=1}^c x_{i+} x_{+i}} = \frac{(1300 \times 944) - 257524}{(1300)^2 - 257524} = 0,677 \quad (8)$$

Onde  $K$  é uma estimativa do coeficiente Kappa,  $x_{ii}$  é o valor na linha  $i$  e coluna  $i$ ,  $x_{i+}$  é a soma da linha  $i$  e  $x_{+i}$  é a soma da coluna  $i$  da matriz de confusão,  $n$  é o número total de amostras e  $c$  o número total de classes. Também pode-se adotar a seguinte forma equivalente para o coeficiente Kappa:

$$K = \frac{P_0 - P_c}{1 - P_c} \quad (9)$$

Onde  $P_0$  seria a concordância real e  $P_c$  a concordância por chance

expressos da seguinte forma:  $P_0 = \frac{\sum_{i=1}^c x_{ii}}{n}$  e  $P_c = \frac{\sum_{i=1}^c x_{i+} x_{+i}}{n^2}$ .

Congalton e Green (1999) sugerem meios de testar a significação estatística do Kappa para uma única matriz de confusão, através da variância de Kappa, a fim de determinar se o nível de acerto da classificação e os dados de referência são significativamente maior que zero a um determinado nível de significância. O teste estatístico para testar a significação de uma única matriz de confusão é determinado pela equação a seguir:

$$Z = \frac{K}{\sqrt{\text{var}(K)}} = \frac{0,677}{\sqrt{0,000207}} = 47,076 \quad (10)$$

Onde  $Z$  é unificado e normalmente distribuído e  $\text{var}$  é a grande variância da amostra do coeficiente Kappa, que pode ser calculado usando o método de *Delta* como segue:

$$\text{var}(K) = \frac{1}{n} \left[ \frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2^2)}{(1-\theta_2)^4} \right] \quad (11)$$

Onde  $\theta_1 = \frac{1}{n} \sum_{i=1}^c x_{ii}$ ,  $\theta_2 = \frac{1}{n^2} \sum_{i=1}^c x_{i+} x_{+i}$ ,  $\theta_3 = \frac{1}{n^2} \sum_{i=1}^c x_{ii} (x_{i+} + x_{+i})$  e

$$\theta_4 = \frac{1}{n^3} \sum_{i=1}^c \sum_{j=1}^c x_{ij} (x_{j+} + x_{+j})^2$$

Se  $Z \geq Z_{\alpha/2}$  a classificação é significativamente melhor que uma distribuição aleatória, onde  $\alpha/2$  é o nível de confiança nos dois lados da curva no teste  $Z$  e o número de graus de liberdade é assumido ser infinito.

Semelhantemente, através de dois índices de exatidão, tomando duas classificações independentes, podemos testar se dois classificadores independentes são significativamente diferentes. Para testar se os dois índices de exatidão ( $K_1$  e  $K_2$ ) são significativamente diferentes, o teste estatístico é expresso pela Equação (12) onde  $Z$  é unificado e normalmente distribuído. Se  $Z \geq Z_{\alpha/2}$ , então os dois classificadores independentes são significativamente diferentes. Assim, é possível comparar estatisticamente dois métodos, dois algoritmos, ou então comparar uma imagem em duas datas diferentes para verificar qual apresenta maior exatidão (Vieira, 2000).

$$Z = \frac{|K_1 - K_2|}{\sqrt{\text{var}(K_1) + \text{var}(K_2)}} \quad (12)$$

O coeficiente Tau ( $T$ ) é superficialmente similar ao Kappa e pode ser calculado como segue:

$$T = \frac{P_0 - P_r}{1 - P_r} \quad (13)$$

Onde  $P_r = \sum_{i=1}^c P_{i+} P_i = \frac{1}{n^2} \sum_{i=1}^c x_{i+} x_i$

Sendo  $P_r$  a concordância casual, onde  $P_{i+} = (x_{i+} / n)$  a distribuição marginal dos dados de referência,  $P_i = (x_i / n)$  as probabilidades *a priori* para cada classe e  $P_0$  já foi definido anteriormente. Quando as probabilidades *a priori* para as classes forem iguais, ou seja,  $x_i = n/c$  temos:

$$P_r = \frac{1}{n^2} \sum_{i=1}^c x_{i+} x_i \quad (14)$$

$$P_r = \frac{1}{n^2} \sum_{i=1}^c x_{i+} \frac{n}{c} \quad (15)$$

$$P_r = \frac{1}{c} \quad (16)$$

Assim,

$$T = \frac{P_0 - \frac{1}{c}}{1 - \frac{1}{c}} \quad (17)$$

Considerando a equiprobabilidade para as classes, já que não se conhece as probabilidades *a priori*, o valor do coeficiente Tau para os dados da Tabela 5 seria então:

$$T = \frac{P_0 - \frac{1}{c}}{1 - \frac{1}{c}} = \frac{0,726 - \frac{1}{7}}{1 - \frac{1}{7}} = 0,681$$

Ma e Redmond (1995) sugeriram denotar  $T_e$  quando Tau for utilizado numa classificação baseada em equiprobabilidade entre as categorias e  $T_p$  quando for utilizado em uma classificação baseada em probabilidades desiguais.

Segundo Ma e Redmond (1995), devido a concordância casual ( $P_r$ ) ser independente dos elementos da matriz de erros e poder ser calculada antes da classificação, ela é tratada como uma constante. Entretanto, a variância de Tau é derivada da seguinte maneira:

$$\text{Sendo } \sigma^2(P_r) = 0 \text{ e } \sigma^2(P_0) = \frac{1}{n} P_0(1 - P_0)$$

temos que

$$\sigma^2(T) = \sigma^2\left(\frac{P_0 - P_r}{1 - P_r}\right) \quad (18)$$

$$\sigma^2(T) = \frac{\sigma^2(P_0 - P_r)}{(1 - P_r)^2} \quad (19)$$

$$\sigma^2(T) = \frac{\sigma^2(P_0) - \sigma^2(P_r)}{(1 - P_r)^2} \quad (20)$$

$$\sigma^2(T) = \frac{P_0(1 - P_0)}{n(1 - P_r)^2} \quad (21)$$

Com base nos dados da Tabela 5, considerando novamente a equiprobabilidade para as classes, a variância de Tau é igual a:

$$\sigma^2(T) = \frac{P_0(1-P_0)}{n(1-P_r)^2} = \frac{0,726(1-0,726)}{1300(1-0,143)^2} = 0,000208$$

Para testar a significação estatística do índice Tau, procede-se da mesma forma descrita para o índice Kappa.

Para Vieira (2000) quando utilizamos algum procedimento de cálculo estatístico esperamos obter um conjunto de intervalos de confiança. Mather (1999) fornece uma fórmula para o cálculo do limite de confiança inferior baseado no valor da exatidão da classificação. Dados o valor de  $P$  e  $n$ , a fórmula para determinar o limite inferior é:

$$s = P - \left[ z \sqrt{\frac{P(100-P)}{n}} + \frac{50}{n} \right] \quad (22)$$

Onde  $P$  é a proporção verdadeira dos resultados,  $n$  é o número de amostras,  $z$  é o  $(100 - r)/100$  ponto da distribuição normal padrão. Assim, se  $r$  for igual a 95% então o valor de  $z$  requerido será 0,05 sob a curva normal padrão. A tabela de valores  $z$  para este ponto é 1,645. Para exemplificar, considere que, de 1300 pixels de teste, 944 foram corretamente classificados, chegando a uma exatidão da classificação aparente de 72,615%. Se requerido 95% de limite de confiança inferior,  $z$  é igual a 1,645 e

$$s = 72,615 - \left[ 1,645 \sqrt{\frac{72,615(100-72,615)}{1300}} + \frac{50}{1300} \right] = 70,542\%$$

Enfim, este resultado indica que 95% das amostras de treinamento com precisões observadas de 72,615% tenham verdadeiras precisões de 70,542% ou maior. Mather (1999) cita que este procedimento também pode ser aplicado a classes individuais do mesmo modo como foi descrito, com a exceção que  $P$  é o número de pixels corretamente classificado na classe  $i$  de uma amostra de teste de  $n_i$  pixels.

## 4.2 Análise espacial

### 4.2.1 Caracterização da área de estudo

A área de estudo está localizada no município da Serra do Salitre (Figura 10), Estado de Minas Gerais, e fica na microrregião de Patrocínio pertencente à mesorregião do Triângulo Mineiro e Alto Paranaíba. Essa região possui uma paisagem predominantemente rural, mas vem sofrendo modificações em função da crescente industrialização e exploração de sua riqueza mineral.

A área de estudo no município da Serra do Salitre está delimitada pelas seguintes coordenadas Geográficas:

- Canto superior esquerdo: 19°15'48,8" S, 46°41'54,97" W;
- Canto inferior direito: 19°16'57,41" S, 46°40'39,5" W.

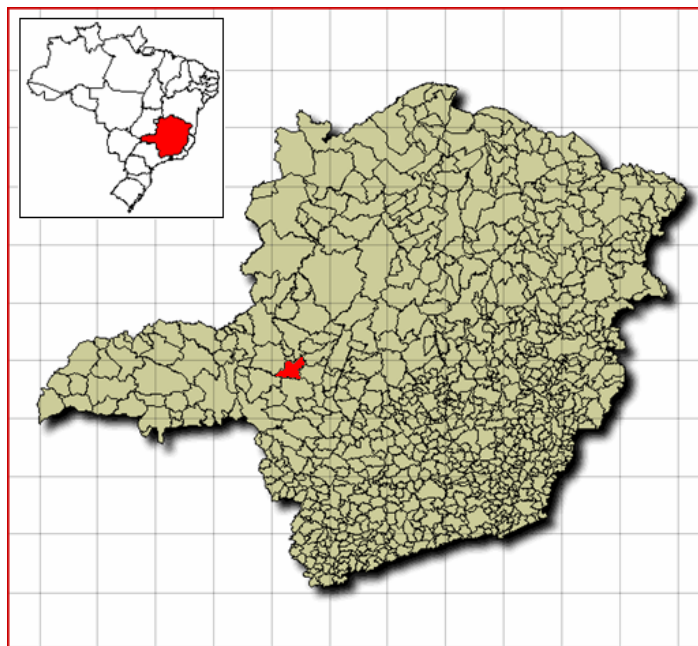


Figura 10: Localização da Serra do Salitre - MG.

Fonte: IBGE, 2007.

### 4.2.2 Materiais

#### 4.2.2.1 Dado Orbital

Para realizar a avaliação da distribuição espacial do erro na classificação de imagens do sensoriamento remoto, utilizou-se uma imagem

obtida pelo satélite Ikonos II nas bandas azul (0,45-0,52 $\mu$ m), verde (0,52-0,60 $\mu$ m), vermelho (0,63-0,69 $\mu$ m) e infravermelho próximo (0,76-0,90  $\mu$ m), com resolução espacial de 4 metros, de tamanho 280x280 *pixels*, obtida em 2 de março de 2002 (Figura 11).

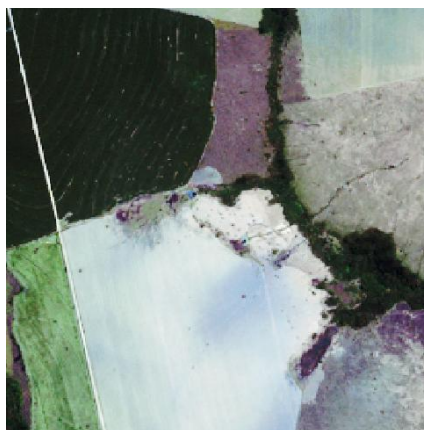


Figura 11: Imagem Ikonos da região da Serra do Salitre.

#### **4.2.2.2 Ferramentas Computacionais**

O software Idrisi Kilimanjaro foi utilizado para realizar a coleta das amostras e para a edição dos mapas finais. O software ERDAS Imagine 8.5 foi utilizado para manipulação das imagens e importação/exportação para formatos compatíveis com os programas utilizados para classificar a imagem. n

Para realizar a classificação da imagem e gerar as imagens de espacialização do erro, foram utilizados os seguintes programas desenvolvidos por Vieira (2000) em linguagem C:

- Programa “Transform.exe”, para gerar os arquivos de normalização e os arquivos \*.pat (para classificação por redes neurais) ;
- Programa “Funcpow.exe”, para gerar a matriz de confusão e realizar a classificação pelo algoritmo da Máxima Verossimilhança (MaxVer); e
- Programa “Reliab\_sup.exe”, para gerar as imagens para visualização da confiança temática.

No processo de classificação por redes neurais artificiais (RNA), utilizou-se o simulador de redes neurais “*Java Neural Network Simulator*”, desenvolvido pelo Instituto Wilhem-Schickard para Ciência da Computação

(WSI) em Tübingen, Alemanha. Este simulador é baseado no *Stuttgart Neural Network Simulator 4.2* (Zell et al., 1996), com uma nova interface gráfica.

#### 4.2.3 Métodos aplicados para análise espacial da exatidão temática

Para realizar as classificações criou-se uma imagem de referência com amostras mais homogêneas possível. Partindo do conhecimento de campo sobre as classes presentes na imagem, foi realizada uma digitalização de polígonos sobre a imagem, gerando uma imagem de referência com 7 classes informacionais com as respectivas proporções (Tabela 6).

Tabela 6 – Classes na imagem de referência.

<b>Classes</b>	<b>Pixels</b>	<b>Percentual (%)</b>
Campo Limpo	7202	13,92
Campo Sujo	2718	5,26
Café	14157	27,38
Braquiária	3955	7,65
Mata Ciliar	2591	5,01
Pastagem	3619	7,00
Solo Exposto	17460	33,77
<b>Total</b>	<b>51702</b>	<b>100</b>

No cálculo do tamanho da amostra para a validação da classificação, Congalton e Green (1999) recomendam usar a *distribuição multinomial* que seria uma alternativa para gerar uma matriz de confusão confiável.

Para gerar uma matriz de erro válida utilizando uma amostra aleatória simples, o tamanho da amostra  $n$  exigido pode ser obtido da seguinte equação:

$$n = B \prod_i (1 - \prod_i) / b_i^2 \quad (23)$$

Onde  $b_i$  é a precisão absoluta da amostra  $i$ ,  $k$  o número de classes,  $\prod_i$  a proporção de área da imagem na classe  $i$  ( $i = 1, \dots, k$ ) e  $B$  é o superior  $(\alpha/k)$  x 100 da distribuição de  $\chi^2$  com 1 grau de liberdade, sendo que  $\alpha$  é o erro admissível. No caso desta dissertação há sete classes ( $k = 7$ ), o nível de

confiança desejado foi 95%, o erro admissível de 5% ( $\alpha = 0,05$ ), e a classe solo cobre 33,77% da área ( $\prod_i = 0,337$ ). O valor de  $B$  deve ser determinado de uma tabela de distribuição qui-quadrado com 1 grau de liberdade e  $\alpha/k$  (0,007142857). O valor de  $B$  é então  $\chi^2_{(1,0.992857)} = 7,348571$ . Logo, o tamanho da amostra seria:

$$n = 7,348571(0,337)(1 - 0,337)/(0,05)^2 = 657$$

Sendo o total de 657 amostras, aproximadamente 94 amostras por classe devem ser tomadas. Quando não se conhecem os valores do  $\prod_i$ 's, pode-se assumir que  $\prod_i = 1/2$  (50%) e  $b_i = b$  para  $i = 1, \dots, k$ . Neste caso menos apropriado, o tamanho da amostra  $n$  é calculado a partir da equação:

$$n = B / 4b^2 \quad (24)$$

Da Equação (24) teríamos um total de 735 amostras ou 105 amostras por classe. O desenvolvimento e derivação das Equações (23) e (24) o leitor pode encontrar em Congalton e Green (1999).

Neste trabalho, no processo das classificações (MaxVer e RNA) utilizou-se um total de 840 amostras (120 amostras por classe). A quantidade de amostras de treinamento e validação dos classificadores foram iguais, porém foram tomadas em locais diferentes.

Realizada a classificação, o próximo passo foi caracterizar a distribuição espacial do erro nesta classificação. Para isto utilizou-se a Distância de Mahalanobis (Equação 25) para gerar uma imagem de distância, em escala de cinza (0-255), permitindo visualizar as discrepâncias de *pixels* individuais dos meios das classes para as quais foram atribuídos.

$$MD_i^2 = (X - M_i)^T S_i^{-1} (X - M_i) \quad (25)$$

Onde  $X$  é o vetor desconhecido,  $M_i$  o centróide da classe  $i$  e  $S_i$  é a variância-covariância da classe específica em questão.

Uma segunda forma utilizada para espacializar o erro foi através da comparação da imagem temática com o respectivo mapa de referência (ou verdade de campo), resultando em uma imagem binária do erro, sendo que

o *pixel* rotulado corretamente recebeu o valor "0" e o rotulado erroneamente recebeu o valor "1". Ainda a partir desta comparação, utilizando a Distância de Mahalanobis, foi possível gerar uma imagem de diferença, onde duas imagens temáticas (uma gerada pela classificação MaxVer e a outra por RNA) foram comparadas com o mapa de referência, sendo também uma forma de espacializar o erro.

Um método alternativo de representar um mapa de confiabilidade temático foi proposto por Vieira (2000). Neste método, a Distância de Mahalanobis é transformada em estimativas de probabilidades *a posteriori*  $P_i$  (probabilidade de um *pixel* pertencer à classe  $i$ ) através da seguinte equação:

$$P_i = \frac{1}{\sum_{j=1}^c \frac{1}{MD_j^2}} \quad (26)$$

Onde  $MD^2$  é a Distância de Mahalanobis entre o *pixel*  $i$  e os valores médios de *pixel* (ou protótipos) de cada classe  $C_i$ .

De posse dessas probabilidades, gera-se um novo valor ( $NV$ ) através da multiplicação do rótulo da classe nomeada por 10 e somando a probabilidade calculada pela Distância de Mahalanobis multiplicada por 10, sendo  $NV = (\text{rótulo da classe} * 10) + (P * 10)$ . Como exemplo, tendo-se um *pixel* atribuído a classe 5 com probabilidade 0,58 tem-se um  $NV$  de 56. Assim, os *pixels* da classe 1 abrangem um  $NV$  de 10 a 19, a classe 2 abrange de 20 a 29, e de igual modo para as outras classes. Para a representação temática da confiança, uma cor separada é atribuída a cada classe na imagem e dentro de cada classe separa-se níveis de precisão onde estes níveis receberam tons separados da cor alocada àquela classe.

Estes procedimentos descritos acima, para caracterizar a distribuição espacial do erro na classificação, foram realizados no programa "Reliab\_sup.exe" implementados por Vieira (2000).

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Exatidão Posicional

Os resultados das análises de exatidão posicional das imagens orbitais a seguir estão separadas em três tópicos: imagens de altíssima resolução espacial, de alta resolução espacial e de média resolução espacial. Entende-se que resolução espacial seja a habilidade que um sensor remoto possui de distinguir objetos na superfície terrestre, então, quando a área representada por *pixels* no terreno for menor, maior será a resolução espacial. Entretanto, espera-se que os sensores de altíssima resolução apresentem boas precisões para escalas de trabalho maiores.

#### **5.1.1 Exatidão Posicional em imagens de altíssima resolução espacial**

Foram coletados 14 pontos nas imagens Ikonos e Quickbird corrigidas geometricamente e seus respectivos pontos homólogos no levantamento em campo para realizar a avaliação da exatidão posicional (Tabelas 7 e 8). Os pontos foram testados para a escala 1:10.000.

Tabela 7 – Pontos homólogos referentes à imagem Ikonos de Viçosa-MG.

Viçosa - MG								
	Campo		IKONOS		Resíduos			
	E (m)	N (m)	E (m)	N (m)	$\Delta E$ (m)	$\Delta N$ (m)	$\Delta P$ (m)	
1	722350,439	7702688,219	722357,211	7702684,452	-6,772	3,767	7,749	
2	721795,388	7702787,337	721795,422	7702787,587	-0,034	-0,250	0,252	
3	721839,730	7702842,394	721839,977	7702843,724	-0,247	-1,330	1,353	
4	721908,649	7702932,228	721908,397	7702932,604	0,252	-0,376	0,453	
5	721711,735	7702948,192	721711,517	7702948,385	0,218	-0,193	0,291	
6	721597,512	7702889,730	721598,491	7702889,842	-0,980	-0,112	0,986	
7	721833,306	7703137,085	721830,526	7703137,044	2,780	0,041	2,780	
8	721772,359	7703065,206	721767,301	7703065,052	5,058	0,154	5,060	
9	721717,200	7703009,960	721720,963	7703008,921	-3,763	1,039	3,904	
10	722051,196	7702660,354	722050,832	7702660,809	0,364	-0,456	0,583	
11	722130,643	7702752,184	722129,914	7702752,668	0,729	-0,484	0,875	
12	722215,518	7702860,352	722214,882	7702858,516	0,636	1,836	1,943	
13	722009,350	7702619,598	722007,283	7702620,075	2,067	-0,477	2,121	
14	721946,777	7702532,170	721944,907	7702533,845	1,870	-1,675	2,510	
					<b>Máximo</b>	5,058	3,767	7,749
					<b>Mínimo</b>	-6,772	-1,675	0,252
					<b>Média</b>	0,1556	0,1061	2,2043
					<b>Desvio Padrão</b>	2,8143	1,3589	2,1382

Tabela 8 – Pontos homólogos referentes à imagem Quickbird de Viçosa-MG.

Viçosa - MG								
	Campo		Quickbird		Resíduos			
	E (m)	N (m)	E (m)	N (m)	$\Delta E$ (m)	$\Delta N$ (m)	$\Delta P$ (m)	
1	722350,439	7702688,219	722350,498	7702688,520	-0,059	-0,301	0,307	
2	721795,388	7702787,337	721795,064	7702787,358	0,324	-0,021	0,325	
3	721839,730	7702842,394	721839,985	7702842,534	-0,255	-0,140	0,291	
4	721908,649	7702932,228	721908,596	7702932,193	0,053	0,036	0,064	
5	721711,735	7702948,192	721711,720	7702948,148	0,014	0,044	0,046	
6	721597,512	7702889,730	721597,340	7702889,788	0,171	-0,058	0,181	
7	721833,306	7703137,085	721833,086	7703137,204	0,220	-0,119	0,250	
8	721772,359	7703065,206	721772,380	7703065,161	-0,021	0,045	0,050	
9	721717,200	7703009,960	721717,415	7703010,457	-0,215	-0,497	0,541	
10	722051,196	7702660,354	722051,134	7702660,371	0,063	-0,018	0,065	
11	722130,643	7702752,184	722130,430	7702752,602	0,213	-0,418	0,469	
12	722215,518	7702860,352	722215,912	7702860,104	-0,394	0,247	0,465	
13	722009,350	7702619,598	722009,179	7702619,727	0,171	-0,129	0,215	
14	721946,777	7702532,170	721946,281	7702532,119	0,496	0,051	0,498	
					<b>Máximo</b>	0,496	0,247	0,541
					<b>Mínimo</b>	-0,394	-0,497	0,046
					<b>Média</b>	0,0558	-0,0913	0,2691
					<b>Desvio Padrão</b>	0,2381	0,1999	0,1760

A partir dos resultados da estatística descritiva, realizou-se o teste de tendência *t* de *Student* para as imagens Ikonos e Quickbird ao nível de

confiança de 90% ( $\alpha = 10\%$ ), com  $n = 14$  amostras e 13 graus de liberdade. O valor crítico tabelado foi  $t_{10\%(13)} = 1,77$ .

Para o cálculo do  $t$  nas direções E e N aplicou-se as Equações (1) e (2), mencionadas no item 3.3. Assim, temos:

(a) Para imagem Ikonos:

$$t_E = 0,2069$$

$$t_N = 0,2922$$

$t_{\text{calculado}} < t_{\text{tabelado}}$  logo, não há tendência nas direções E e N na imagem Ikonos.

(b) Para imagem Quickbird:

$$t_E = 0,8765$$

$$t_N = 1,7084$$

$t_{\text{calculado}} < t_{\text{tabelado}}$  logo, não há tendência nas direções E e N na imagem Quickbird.

Para a análise da exatidão nas imagens Ikonos e Quickbird aplicou-se o teste *Qui-quadrado* ( $\chi^2$ ), ao nível de confiança de 90% ( $\alpha = 10\%$ ) e 13 graus de liberdade. O valor crítico tabelado foi  $\chi^2_{10\%, 13} = 19,812$ .

Para o cálculo de  $\chi^2$  da posição utilizou-se a Equação (3) do item 3.3. O valor de  $\sigma_P$  foi obtido por meio do Erro Padrão (EP) segundo a Tabela 9.

Tabela 9 – Resultado do PEC, EP e  $\sigma_P$  para a escala 1:10.000.

CLASSE	PLANIMETRIA		ESCALA 1:10.000
	PEC	EP	$\sigma_P$ (m)
<b>A</b>	0,5 mm	0,3 mm	2,1213
<b>B</b>	0,8 mm	0,5 mm	3,5355
<b>C</b>	1,0 mm	0,6 mm	4,2426

A Tabela 10 apresenta os resultados do cálculo de  $\chi^2$  referentes as imagens Ikonos e Quickbird.

Tabela 10 – Valores de  $\chi^2_P$  usados na análise de exatidão para as imagens Ikonos e Quickbird.

CLASSE	Ikonos	Quickbird
<b>A</b>	13,2082	0,0895
<b>B</b>	4,7549	0,0322
<b>C</b>	3,3020	0,0224

Comparando os valores da Tabela 10 com o valor do  $\chi^2$  tabelado, conclui-se que as imagens Ikonos e Quickbird testadas têm exatidão equivalente à classe **A**, para a escala de 1:10.000 a um nível de confiança de 90%. Essa imagem Quickbird possui uma excelente qualidade, tanto posicional quanto visual, e é possível conseguir resultados bons a escalas de trabalho ainda maiores que 1:10.000.

### ***5.1.2 Exatidão Posicional em imagens de alta resolução espacial***

As Tabelas 11 e 12 a seguir apresentam os 26 pontos coletados nas imagens CBERS e Landsat e seus respectivos homólogos na carta de 1:25.000.

Tabela 11 – Pontos homólogos referentes à imagem CBERS de Uberaba-MG.

Uberaba - MG							
Pto	Carta - 1:25.000		CBERS		Resíduos		
	E (m)	N (m)	E (m)	N (m)	ΔE (m)	ΔN (m)	ΔP (m)
1	188408,091	7811115,968	188386,823	7811125,598	21,268	-9,630	23,347
2	198654,217	7811897,378	198674,646	7811884,034	-20,429	13,344	24,401
3	198057,781	7814729,097	198067,351	7814739,016	-9,570	-9,919	13,783
4	186113,987	7821073,061	186156,707	7821040,389	-42,720	32,672	53,782
5	189547,074	7819970,090	189553,356	7819959,541	-6,282	10,549	12,278
6	198847,042	7820237,990	198805,868	7820194,105	41,174	43,885	60,176
7	187530,861	7813335,764	187556,992	7813311,371	-26,131	24,393	35,747
8	187265,908	7809302,979	187284,710	7809283,644	-18,802	19,335	26,970
9	191301,943	7822852,361	191278,001	7822826,548	23,942	25,813	35,207
10	203212,158	7810388,959	203239,004	7810417,210	-26,846	-28,251	38,972
11	204701,048	7812386,811	204686,084	7812402,061	14,964	-15,250	21,365
12	198813,460	7820939,624	198786,250	7820977,928	27,210	-38,304	46,985
13	202247,781	7819699,993	202262,335	7819714,258	-14,554	-14,265	20,379
14	207828,891	7812964,902	207790,451	7812982,857	38,440	-17,955	42,427
15	210833,601	7814573,122	210877,561	7814559,759	-43,960	13,363	45,946
16	208516,074	7822524,710	208530,119	7822510,111	-14,045	14,599	20,258
17	206824,474	7823357,371	206784,910	7823371,113	39,564	-13,742	41,883
18	199357,126	7822740,799	199315,815	7822769,317	41,311	-28,518	50,198
19	187664,441	7819620,722	187635,301	7819590,668	29,140	30,054	41,861
20	190325,392	7821184,110	190336,051	7821141,852	-10,659	42,258	43,582
21	191201,791	7821817,766	191221,770	7821796,289	-19,979	21,477	29,333
22	201061,136	7815815,637	201050,743	7815798,711	10,393	16,926	19,862
23	194473,917	7817439,043	194503,881	7817410,330	-29,964	28,713	41,500
24	196673,796	7814947,545	196650,065	7814915,632	23,731	31,913	39,769
25	197341,536	7813539,933	197312,374	7813554,218	29,162	-14,285	32,473
26	205775,819	7813946,481	205751,848	7813983,236	23,971	-36,755	43,881
				<b>Máximo</b>	41,311	43,885	60,176
				<b>Mínimo</b>	-43,960	-38,304	12,278
				<b>Média</b>	3,0896	5,4777	34,8602
				<b>Desvio Padrão</b>	27,5786	24,9648	12,6929

Tabela 12 – Pontos homólogos referentes à imagem Landsat de Uberaba-MG.

Uberaba - MG								
Pto	Carta - 1:25.000		Landsat		Resíduos			
	E (m)	N (m)	E (m)	N (m)	$\Delta E$ (m)	$\Delta N$ (m)	$\Delta P$ (m)	
1	188408,091	7811115,968	188425,085	7811099,139	-16,994	16,829	23,917	
2	198654,217	7811897,378	198621,880	7811875,534	32,337	21,844	39,024	
3	198057,781	7814729,097	198037,116	7814686,389	20,665	42,708	47,445	
4	186113,987	7821073,061	186133,627	7821038,694	-19,640	34,367	39,583	
5	189547,074	7819970,090	189581,943	7819937,996	-34,869	32,094	47,391	
6	198847,042	7820237,990	198862,324	7820202,790	-15,282	35,200	38,374	
7	187530,861	7813335,764	187503,599	7813367,986	27,262	-32,222	42,208	
8	187265,908	7809302,979	187254,109	7809323,907	11,799	-20,928	24,025	
9	191301,943	7822852,361	191336,667	7822812,114	-34,724	40,247	53,156	
10	203212,158	7810388,959	203202,589	7810400,300	9,569	-11,341	14,839	
11	204701,048	7812386,811	204717,694	7812341,076	-16,646	45,735	48,670	
12	198813,460	7820939,624	198828,148	7820920,601	-14,688	19,023	24,034	
13	202247,781	7819699,993	202262,767	7819677,377	-14,986	22,616	27,130	
14	207828,891	7812964,902	207790,560	7812992,057	38,331	-27,155	46,975	
15	210833,601	7814573,122	210867,771	7814602,601	-34,170	-29,479	45,129	
16	208516,074	7822524,710	208563,937	7822565,449	-47,863	-40,739	62,853	
17	206824,474	7823357,371	206783,780	7823322,744	40,694	34,627	53,432	
18	199357,126	7822740,799	199379,258	7822780,501	-22,132	-39,702	45,454	
19	187664,441	7819620,722	187685,604	7819594,851	-21,163	25,871	33,424	
20	190325,392	7821184,110	190365,583	7821145,220	-40,191	38,890	55,926	
21	191201,791	7821817,766	191158,141	7821846,246	43,650	-28,480	52,119	
22	201061,136	7815815,637	201083,719	7815834,175	-22,583	-18,538	29,217	
23	194473,917	7817439,043	194487,883	7817454,719	-13,966	-15,676	20,995	
24	196673,796	7814947,545	196643,712	7814922,833	30,084	24,712	38,932	
25	197341,536	7813539,933	197330,131	7813564,125	11,405	-24,192	26,746	
26	205775,819	7813946,481	205796,009	7813935,252	-20,190	11,229	23,103	
					<b>Máximo</b>	43,650	45,735	62,853
					<b>Mínimo</b>	-47,863	-40,739	14,839
					<b>Média</b>	-4,7804	6,0592	38,6193
					<b>Desvio Padrão</b>	27,6329	29,8096	12,7685

Realizou-se o teste de tendência  $t$  de *Student* para as imagens CBERS e Landsat ao nível de confiança de 90% ( $\alpha = 10\%$ ), com  $n = 26$  amostras e 25 graus de liberdade. O valor crítico tabelado foi  $t_{10\%(25)} = 1,71$ .

Para o cálculo do  $t$  nas direções E e N aplicou-se as Equações (1) e (2), mencionadas no item 3.3. Assim, temos:

(a) Para imagem CBERS:

$$t_E = 0,5712$$

$$t_N = 1,1188$$

$t_{\text{calculado}} < t_{\text{tabelado}}$  logo, não há tendência nas direções E e N na imagem CBERS.

(b) Para imagem Landsat:

$$t_E = 0,8821$$

$$t_N = 1,0364$$

$t_{\text{calculado}} < t_{\text{tabelado}}$  logo, não há tendência nas direções E e N na imagem Landsat.

Para a análise da exatidão nas imagens CBERS e Landsat aplicou-se o teste *Qui-quadrado* ( $\chi^2$ ), ao nível de confiança de 90% ( $\alpha = 10\%$ ) e 25 graus de liberdade. O valor crítico tabelado foi  $\chi^2_{10\%, 25} = 34,382$ .

Para o cálculo de  $\chi^2$  da posição utilizou-se a Equação (3) do item 3.3. O valor de  $\sigma_P$  foi obtido por meio do Erro Padrão (EP) segundo a Tabela 13.

Tabela 13 – Resultado do PEC, EP e  $\sigma_P$  para a escala 1:25.000.

CLASSE	PLANIMETRIA		ESCALA 1:25.000
	PEC	EP	$\sigma_P$ (m)
A	0,5 mm	0,3 mm	5,3033
B	0,8 mm	0,5 mm	8,8388
C	1,0 mm	0,6 mm	10,6066

A Tabela 14 apresenta os resultados do cálculo de  $\chi^2$  referentes as imagens CBERS e Landsat.

Tabela 14 – Valores de  $\chi^2_P$  usados na análise de exatidão para as imagens CBERS e Landsat.

CLASSE	CBERS	Landsat
A	143,2096	144,9201
B	51,5554	52,1712
C	35,8024	36,2300

Comparando os valores da Tabela 14 com o valor do  $\chi^2$  tabelado, conclui-se que as imagens CBERS e Landsat testadas não tem exatidão posicional, para a escala de 1:25.000 a um nível de confiança de 90%.

Para avaliar a exatidão das imagens CBERS e Landsat também foram aplicados o método do Ponto Gerado, o método das Áreas e o método dos Retângulos Equivalentes. Para o método do Ponto Gerado, estabeleceu-se

os pontos iniciais homólogos e foram coletadas 71 amostras utilizando equidistância de 500 metros em toda a extensão de cada uma das feições. A Tabela 15 apresenta as distâncias, em metros, entre os pontos equidistantes.

Como o método do Ponto Gerado trabalha com distâncias relativas entre pontos homólogos, não há necessidade de avaliar tendência para aferir se há presença de erros sistemáticos nas direções E e N. Assim, procedeu-se a avaliação da exatidão a partir da média e do desvio padrão do conjunto de distâncias euclidianas (D) obtidas. Aplicou-se o teste *Qui-quadrado* ( $\chi^2$ ), ao nível de confiança de 90% ( $\alpha = 10\%$ ) e 70 graus de liberdade. O valor crítico tabelado foi  $\chi^2_{10\%, 70} = 85,527$ .

No cálculo de  $\chi^2$  para as distâncias, os valores do desvio padrão ( $\sigma_D$ ) para cada classe são os mesmos especificados na Tabela 13. A Tabela 16 apresenta os resultados do cálculo de  $\chi^2$  para as distâncias, referentes as imagens CBERS e Landsat, na análise de exatidão pelo método do Ponto Gerado.

Tabela 15 – Distâncias entre pontos equidistantes referentes à Uberaba – MG, coletados através de feições homólogas.

Uberaba - MG							
Equidistância de 500m							
CBERS				Landsat			
Pto		Pto		Pto		Pto	
1	3,001	37	75,410	1	1,548	37	64,147
2	0,582	38	50,237	2	8,628	38	40,902
3	1,084	39	46,968	3	15,869	39	33,116
4	1,692	40	46,411	4	23,012	40	16,368
5	7,645	41	44,541	5	29,962	41	16,269
6	14,994	42	45,365	6	35,377	42	22,631
7	16,438	43	47,451	7	21,382	43	20,676
8	17,674	44	49,858	8	26,836	44	27,974
9	18,274	45	50,538	9	36,216	45	21,292
10	12,733	46	45,670	10	27,146	46	19,468
11	9,401	47	44,220	11	32,764	47	15,110
12	8,064	48	48,731	12	10,422	48	12,433
13	15,895	49	44,416	13	9,572	49	14,166
14	35,385	50	43,756	14	26,212	50	12,047
15	56,869	51	43,603	15	22,032	51	12,260
16	34,709	52	43,116	16	27,877	52	11,729
17	29,043	53	43,449	17	18,911	53	13,118
18	27,547	54	43,275	18	13,405	54	11,587
19	37,998	55	44,194	19	14,003	55	16,389
20	56,699	56	42,064	20	14,640	56	11,340
21	68,373	57	38,427	21	9,057	57	13,045
22	53,405	58	26,277	22	12,007	58	20,870
23	44,786	59	25,314	23	18,468	59	4,614
24	76,335	60	25,802	24	14,165	60	5,017
25	39,612	61	25,978	25	22,965	61	4,166
26	45,116	62	26,776	26	54,291	62	14,318
27	69,588	63	24,206	27	41,706	63	9,474
28	83,428	64	23,955	28	40,911	64	12,563
29	67,159	65	24,142	29	44,750	65	10,100
30	53,788	66	25,297	30	40,728	66	11,934
31	63,470	67	23,582	31	41,467	67	8,281
32	56,208	68	21,902	32	37,859	68	8,742
33	57,910	69	44,332	33	37,776	69	9,027
34	21,234	70	53,315	34	37,051	70	18,029
35	45,384	71	33,289	35	36,238	71	35,262
36	80,207			36	50,390		
<b>Média</b>			38,276	<b>Média</b>			21,889
<b>Desvio Padrão</b>			19,647	<b>Desvio Padrão</b>			13,289

Tabela 16 – Valores de  $\chi^2_D$  utilizados na análise de exatidão pelo método do Ponto Gerado.

<b>CLASSE</b>	<b>CBERS</b>	<b>Landsat</b>
<b>A</b>	960,7149	439,54606
<b>B</b>	345,8574	158,23658
<b>C</b>	240,1787	109,88652

Para o método do Ponto Gerado, comparando os valores da Tabela 16 com o valor do  $\chi^2$  tabelado, conclui-se que as imagens CBERS e Landsat não tem exatidão posicional para a escala de 1:25.000. A baixa qualidade visual dessas imagens, especialmente a CBERS, dificultou digitalizar com clareza as feições, podendo ter influenciado neste resultado. Apesar disto, para escalas menores, essas imagens geralmente apresentam exatidão posicional e podem ser utilizadas para representações em síntese de grandes extensões territoriais, revelando com fidelidade a situação do uso e ocupação do solo de um estado ou município, dando suporte ao trabalho de gestão ambiental e territorial, e ao monitoramento de áreas de preservação.

No método das Áreas foram utilizadas 24 amostras para a imagem CBERS e 21 amostras para a imagem Landsat de áreas e perímetros calculados a partir da sobreposição das feições homólogas da região de estudo (Tabela 17).

Tabela 17 – Áreas e perímetros originados da sobreposição de feições homólogas na região de Uberaba-MG.

Uberaba - MG					
CIBERS			Landsat		
Pto	Área (m <sup>2</sup> )	Perímetro (m)	Pto	Área (m <sup>2</sup> )	Perímetro (m)
1	3023,818	2248,330	1	1099,688	789,860
2	2033672,033	22766,406	2	131437,735	11632,803
3	216918,516	8968,245	3	113,677	294,355
4	174006,911	16218,356	4	130238,906	14440,926
5	2116,440	1238,939	5	149442,073	13684,227
6	24,797	202,265	6	684,232	457,137
7	336,314	717,715	7	15215,831	1967,399
8	2936,073	1149,236	8	39143,175	7943,466
9	25566,488	3647,759	9	622,460	748,023
10	1086,365	602,771	10	9884,914	3386,875
11	617,231	450,579	11	2244,365	1278,051
12	162,014	186,027	12	16186,077	2384,840
13	1482,314	660,022	13	273,726	354,040
14	27,389	87,632	14	141,532	221,538
15	433,685	360,997	15	704,284	371,963
16	2963,536	1041,316	16	2800,269	751,902
17	12997,061	2507,506	17	276,078	252,733
18	174,558	335,664	18	31,682	75,670
19	1160,826	918,018	19	18,653	59,685
20	5478,442	1781,985	20	23274,889	5261,923
21	72,189	267,977	21	74408,971	7271,718
22	2863,045	1397,841	<b>Total</b>	<b>598243,214</b>	<b>73629,135</b>
23	191,345	158,119			
24	57902,252	3575,466			
<b>Total</b>	<b>2546213,641</b>	<b>71489,170</b>			

A partir do total das áreas e do total dos perímetros, aplicou-se a Equação (5), citada no item 3.3, para estimar o erro em área. Assim temos:

(a) Para a imagem CBERS:

$$\text{Área (S)} = 2546213,641 \text{ m}^2$$

$$\text{Perímetro (P)} = 71489,170 \text{ m}$$

$$\text{Erro em área (e}_a\text{)} = 35,617 \text{ m}$$

$$\text{Na escala de 1:25.000, e}_a = 1,425 \text{ mm}$$

(b) Para a imagem Landsat:

$$S = 598243,214 \text{ m}^2$$

$$P = 73629,135 \text{ m}$$

$$e_a = 8,125 \text{ m.}$$

$$\text{Na escala de 1:25.000, e}_a = 0,325 \text{ mm}$$

Esses resultados obtidos pelo método das Áreas nos fornecem apenas um estimador do erro em área ( $e_a$ ).

Para o método dos Retângulos Equivalentes, foram utilizados os dados referentes ao método das Áreas para a resolução do sistema mencionado anteriormente no item 3.3 (Equação 6):

(a) Para a imagem CBERS:

$$\begin{cases} x_1 \times x_2 = 2546213,641 \\ 2x_1 + 2x_2 = 2 \times 71489,170 \end{cases}$$

$$X_1 = 71453,5353 \text{ m}$$

$$X_2 = 35,6345 \text{ m}$$

Na escala de 1:25.000 tem-se:

$$X_1 = 2,8581 \text{ m}$$

$$X_2 = 0,1425 \text{ cm}$$

(b) Para a imagem Landsat:

$$\begin{cases} x_1 \times x_2 = 598243,214 \\ 2x_1 + 2x_2 = 2 \times 73629,135 \end{cases}$$

$$X_1 = 8,1260 \text{ m}$$

$$X_2 = 73621,0094 \text{ m}$$

Na escala de 1:25.000 tem-se:

$$X_1 = 0,0325 \text{ cm}$$

$$X_2 = 2,9448 \text{ m}$$

Os resultados obtidos pelo método dos Retângulos Equivalentes representam estimativas das discrepâncias entre as representações de uma dada feição.

Esses métodos, que são baseados em características (feições) para avaliar a exatidão posicional, são uma alternativa em situações onde a presença de pontos bem identificáveis e distribuídos sejam insuficientes, o que prejudicaria a análise através da comparação por pontos. Porém, os métodos das áreas e dos retângulos equivalentes ainda não possuem valores críticos (tabelados) para se estabelecer um teste estatístico avaliando se o resultado é satisfatório ou não.

### 5.1.3 Exatidão Posicional em imagens de média resolução espacial

A Tabela 18 apresenta os pontos coletados aleatoriamente na imagem MODIS e seus respectivos homólogos na carta imagem Landsat na escala de 1:400.000.

Tabela 18 – Pontos homólogos referentes a abrangência da EDR de Limeira-SP.

EDR Limeira-SP							
Pto	LANDSAT / ETM+		MODIS		Resíduos		
	E (m)	N (m)	E (m)	N (m)	$\Delta E$ (m)	$\Delta N$ (m)	$\Delta P$ (m)
1	236813,643	7525858,170	236814,306	7525859,143	-0,663	-0,973	1,177
2	240780,637	7514453,543	240769,162	7514445,130	11,475	8,413	14,229
3	247801,270	7510344,364	247777,766	7510340,090	23,504	4,274	23,889
4	248321,011	7504584,840	248328,442	7504583,022	-7,431	1,818	7,650
5	243426,539	7501181,196	243422,419	7501228,905	4,120	-47,709	47,887
6	209659,358	7536879,817	209660,666	7536912,900	-1,308	-33,083	33,109
7	249692,355	7568922,325	249715,080	7568924,146	-22,725	-1,821	22,798
8	250733,501	7561879,278	250720,131	7561888,792	13,370	-9,514	16,410
9	259289,679	7566036,046	259288,542	7566042,402	1,137	-6,356	6,457
10	242675,306	7581598,181	242683,887	7581590,801	-8,581	7,380	11,318
11	242417,358	7578463,124	242418,149	7578477,863	-0,791	-14,739	14,760
12	242556,432	7522585,042	242571,374	7522605,148	-14,942	-20,106	25,050
13	215780,175	7545093,762	215710,871	7545063,512	69,304	30,250	75,618
14	203372,433	7534463,179	203341,400	7534450,768	31,033	12,411	33,423
15	225726,777	7524764,918	225820,425	7524704,130	-93,648	60,788	111,647
16	207055,520	7540637,885	207065,944	7540619,646	-10,424	18,239	21,008
17	210711,644	7543813,461	210722,842	7543797,883	-11,198	15,578	19,185
18	262493,805	7567600,802	262494,811	7567622,302	-1,006	-21,500	21,524
19	250595,089	7582668,911	250589,525	7582656,920	5,564	11,991	13,219
20	243642,088	7581881,294	243622,059	7581873,080	20,029	8,214	21,648
21	247169,803	7526140,459	247187,331	7526134,641	-17,528	5,818	18,468
22	260740,566	7490598,531	260730,810	7490591,716	9,756	6,815	11,901
23	201957,619	7538205,977	201970,345	7538199,041	-12,726	6,936	14,493
24	255174,524	7541807,484	255179,353	7541793,844	-4,829	13,640	14,470
25	257226,919	7496592,400	257242,910	7496610,902	-15,991	-18,502	24,455
				<b>Máximo</b>	69,304	60,788	111,647
				<b>Mínimo</b>	-93,648	-47,709	1,177
				<b>Média</b>	-1,3800	1,5305	25,0317
				<b>Desvio Padrão</b>	27,1964	21,2970	23,3439

A partir dos resultados da estatística descritiva, realizou-se o teste de tendência  $t$  de *Student* para a imagem MODIS ao nível de confiança de 90% ( $\alpha = 10\%$ ), com  $n = 25$  amostras e 24 graus de liberdade. O valor crítico tabelado foi  $t_{10\%(24)} = 1,71$ .

Para o cálculo do  $t$  nas direções E e N aplicou-se as Equações (1) e (2), mencionadas no item 3.3. Assim temos:

$$t_E = 0,2537$$

$$t_N = 0,3593$$

$t_{\text{calculado}} < t_{\text{tabelado}}$  logo, não há tendência nas direções E e N na imagem MODIS.

Para a análise da exatidão na imagem MODIS aplicou-se o teste *Qui-quadrado* ( $\chi^2$ ), ao nível de confiança de 90% ( $\alpha = 10\%$ ) e 24 graus de liberdade. O valor crítico tabelado foi  $\chi^2_{10\%, 24} = 33,196$ .

Para o cálculo de  $\chi^2$  da posição utilizou-se a Equação (3) do item 3.3. O valor de  $\sigma_P$  foi obtido por meio do Erro Padrão (EP) segundo a Tabela 19.

Tabela 19 – Resultado do PEC, EP e  $\sigma_P$  para a escala 1:400.000.

CLASSE	PLANIMETRIA		ESCALA 1:400.000
	PEC	EP	$\sigma_P$ (m)
A	0,5 mm	0,3 mm	84,8528
B	0,8 mm	0,5 mm	141,4214
C	1,0 mm	0,6 mm	169,7056

A Tabela 20 apresenta os resultados do cálculo de  $\chi^2$  referente a imagem MODIS.

Tabela 20 – Valores de  $\chi^2_P$  usados na análise de exatidão para a imagem MODIS.

CLASSE	MODIS
A	1,8165
B	0,6539
C	0,4541

Comparando os valores da Tabela 20 com o valor do  $\chi^2$  tabelado, conclui-se que a imagem MODIS testada tem exatidão equivalente à classe **A**, na escala de 1:400.000 a um nível de confiança de 90%.

O sensor MODIS fornece uma cobertura quase que diária da superfície terrestre. Foi desenvolvido com o objetivo de servir como base de

dados para a quantificação e detecção das mudanças da cobertura terrestre e dos processos naturais. Uma característica da imagem do sensor MODIS é que ela é fornecida já georreferenciada, e pelo resultado obtido da análise de exatidão pode-se confirmar que, para a escala 1:400.000, apresentou um ótimo resultado.

## **5.2 Exatidão temática**

### ***5.2.1 Análise não-espacial da exatidão temática***

Para analisarmos o comportamento dos índices, em cada caso foram geradas cinco matrizes de confusão, de onde foram derivados os índices de exatidão global ( $G$ ), Kappa ( $K$ ) e Tau ( $T$ ) (registrados nos Apêndices), obtidos utilizando as equações (7), (8) e (17), respectivamente. Os resultados são representados nos gráficos a seguir, pelas diferenças absolutas entre os índices:  $\Delta GK$  – diferença entre o índice de exatidão global e o Kappa;  $\Delta GT$  - diferença entre o índice de exatidão global e o Tau; e  $\Delta KT$  - diferença entre o índice de Kappa e o Tau; de onde extraiu-se a média e o desvio-padrão. Chamamos a atenção do leitor para a variação dos valores no eixo Y dos gráficos.

Variando o número de classes informacionais pode-se notar, analisando as Figuras 12 a 16, que  $\Delta GK$  e  $\Delta GT$  diminuem a medida que aumenta o número de classes, mas o valor da variação aumenta a medida que cai o nível de desempenho da classificação (níveis apresentados na Tabela 3, seção 2.2.4). A variação de  $\Delta KT$  foi muito pequena em todos os níveis, e manteve-se constante.

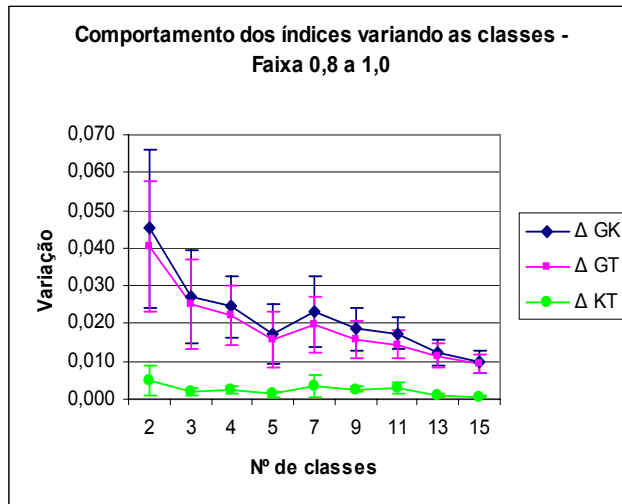


Figura 12: Variando o número de classes informacionais num nível de desempenho excelente.

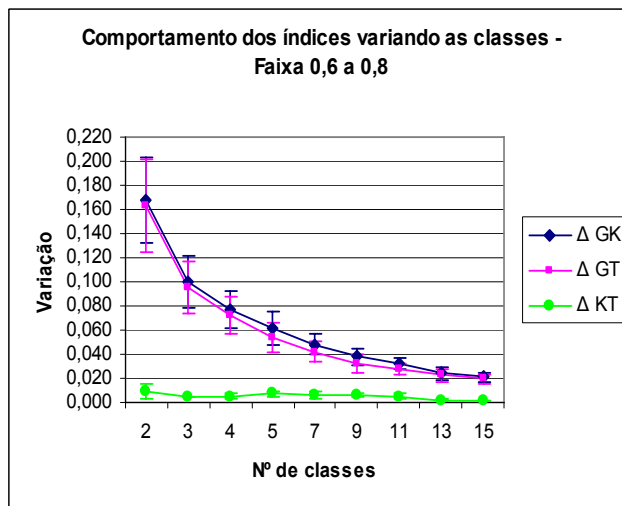


Figura 13: Variando o número de classes informacionais num nível de desempenho muito bom.

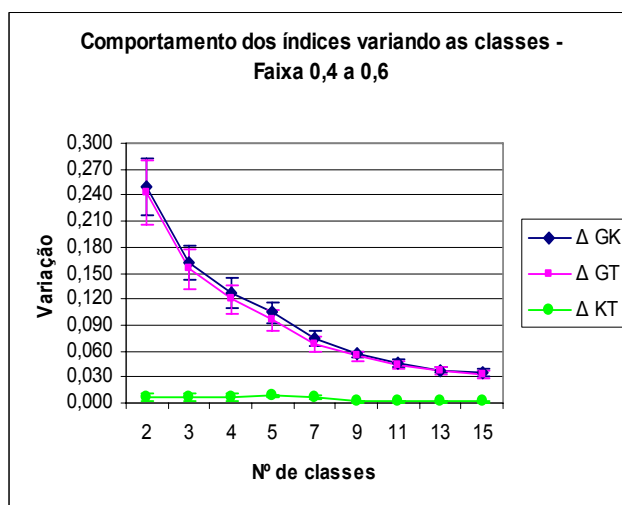


Figura 14: Variando o número de classes informacionais num nível de desempenho bom.

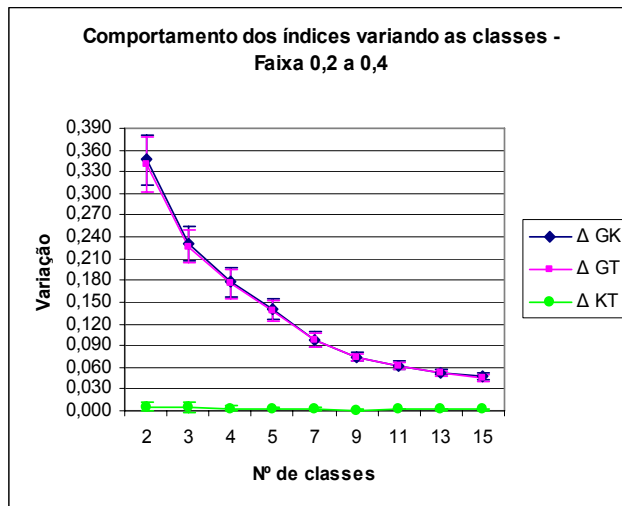


Figura 15: Variando o número de classes informacionais num nível de desempenho razoável.

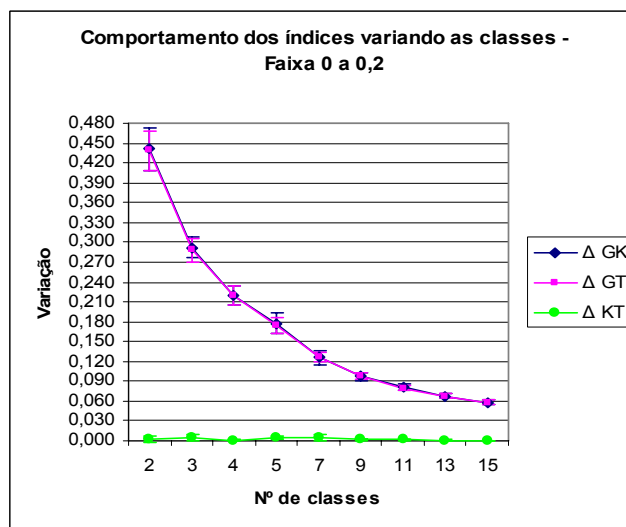


Figura 16: Variando o número de classes informacionais num nível de desempenho ruim.

Variando o número total de amostras pode-se notar, a partir das Figuras 17 a 21, que  $\Delta$ GK é maior que  $\Delta$ GT e o valor da variação também aumentou a medida que caiu o nível de desempenho da classificação, sendo que a variação entre  $\Delta$ KT também foi baixa em todos os níveis.

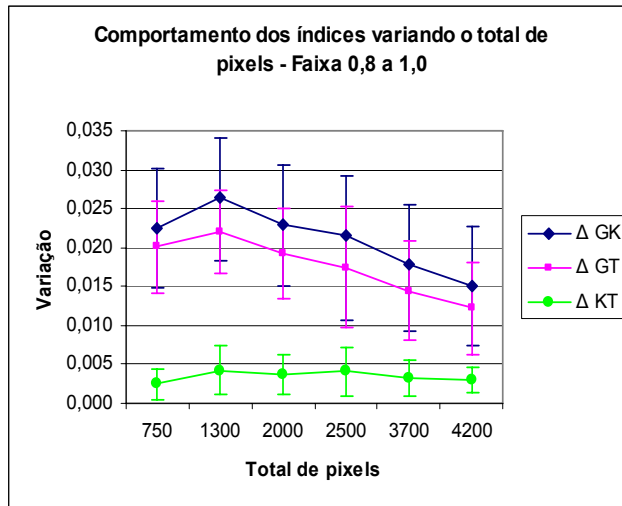


Figura 17: Variando o total de amostras num nível de desempenho excelente.

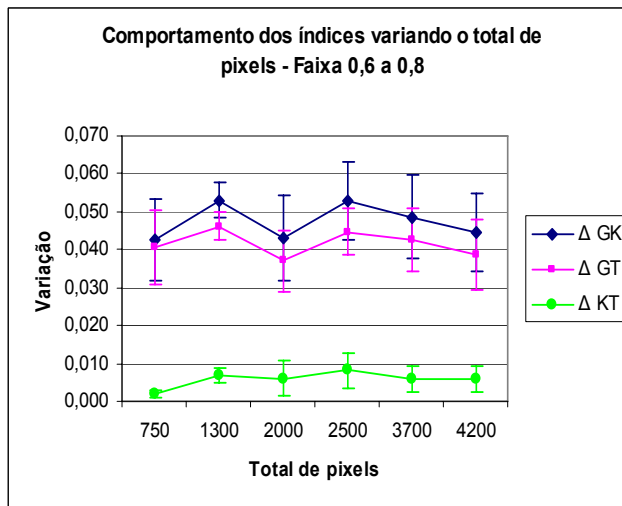


Figura 18: Variando o total de amostras num nível de desempenho muito bom.

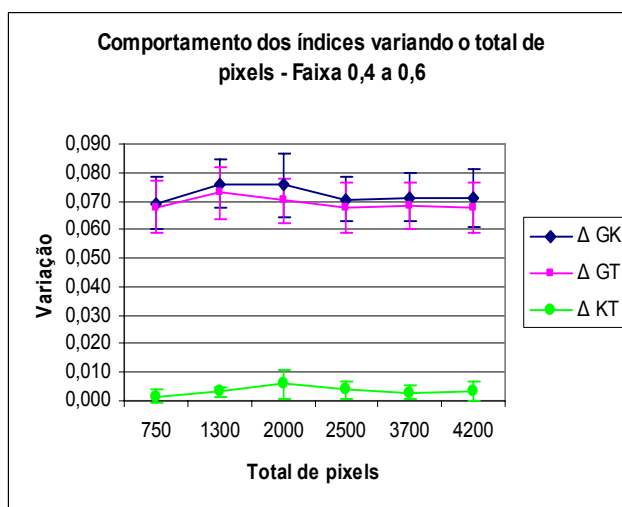


Figura 19: Variando o total de amostras num nível de desempenho bom.

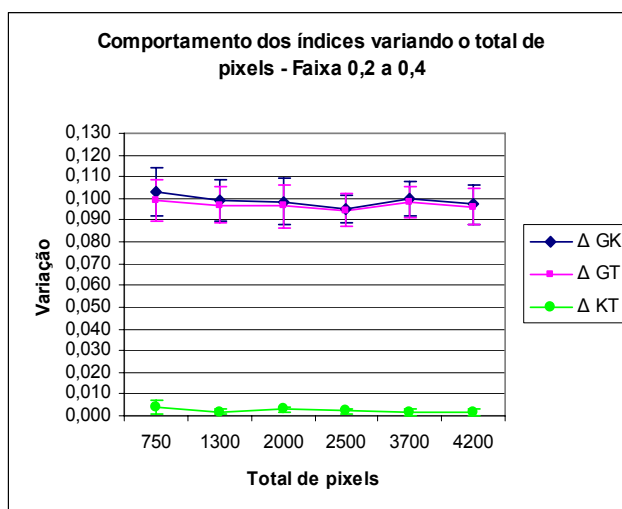


Figura 20: Variando o total de amostras num nível de desempenho razoável.

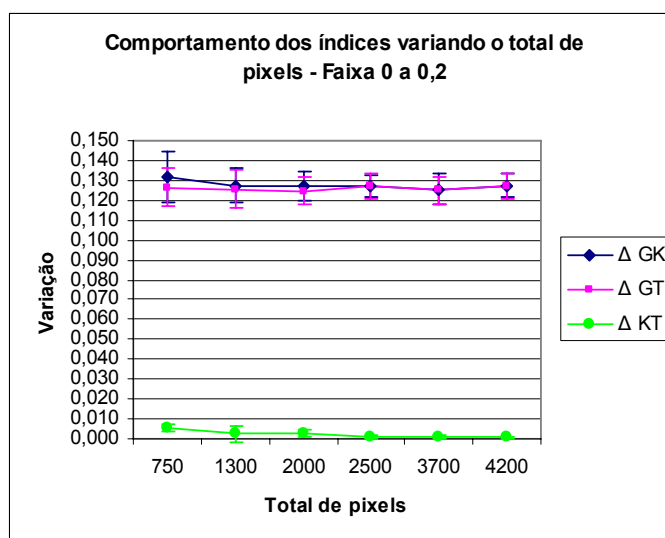


Figura 21: Variando o total de amostras num nível de desempenho ruim.

Variando o número total de pixels com presença de classe informacional predominante, para validação, das Figuras 22 a 26 nota-se que  $\Delta$ GK e  $\Delta$ GT foram mais discrepantes um em relação ao outro, sendo que  $\Delta$ GT tende a ser constante, e também o valor da variação aumentou a medida que caiu o nível de desempenho da classificação. A diferença entre  $\Delta$ KT foi mais alta quando ocorreu a presença de classe informacional predominante nos valores amostrados.

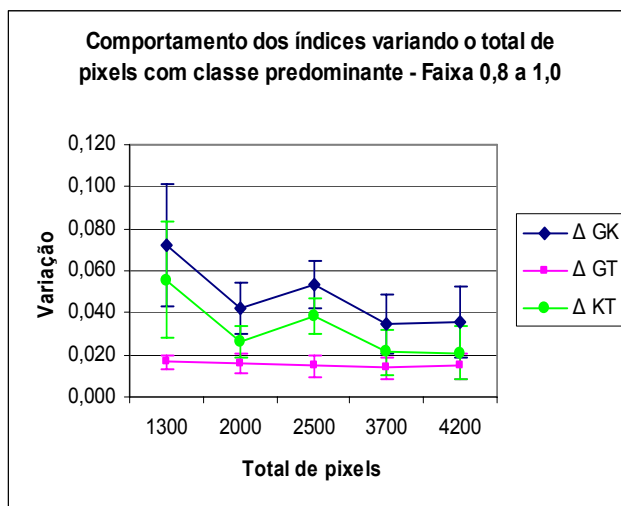


Figura 22: Variando o total de pixels com classe predominante num nível de desempenho excelente.

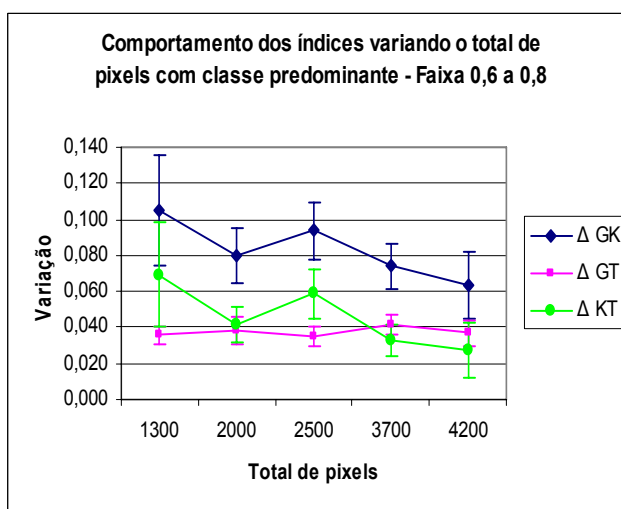


Figura 23: Variando o total de pixels com classe predominante num nível de desempenho muito bom.

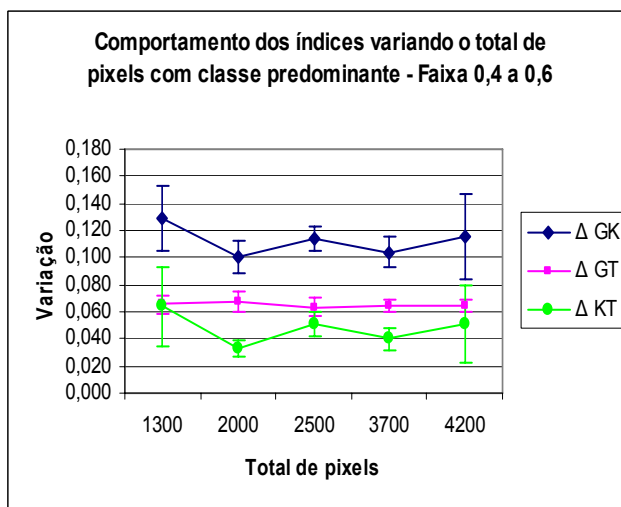


Figura 24: Variando o total de pixels com classe predominante num nível de desempenho bom.

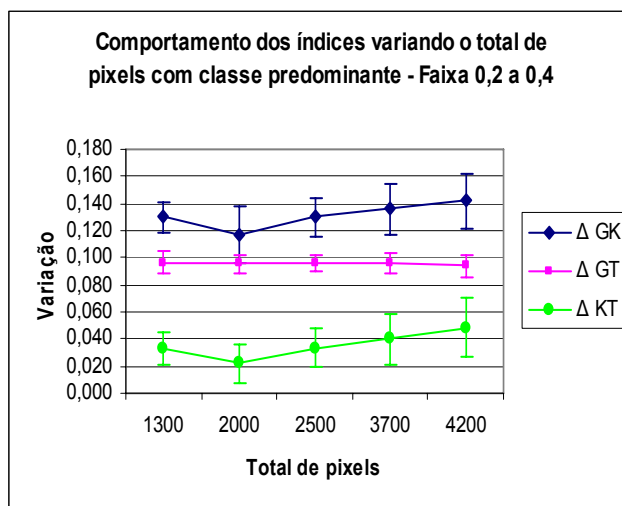


Figura 25: Variando o total de pixels com classe predominante num nível de desempenho razoável.

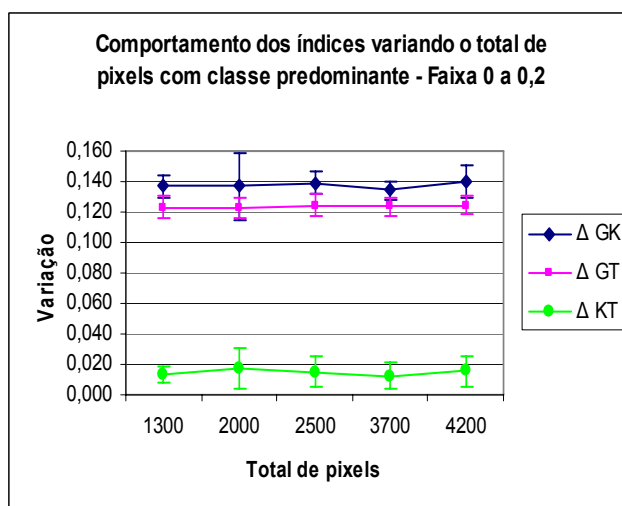


Figura 26: Variando o total de pixels com classe predominante num nível de desempenho ruim.

Diante dos resultados encontrados, pode-se observar que os índices de exatidão apresentam comportamentos diferentes quando são submetidos a variação do número de classes informacionais e/ou do número de padrões de validação; entretanto, o comportamento deles é mais alterado quando ocorre a presença de classe informacional predominante entre os valores amostrados, para a validação. Este fato confirma a importância de se realizar uma amostragem bem distribuída para todas as classes, evitando que a validação da classificação seja comprometida.

### 5.2.2 Análise espacial da exatidão temática

Para caracterizar a distribuição espacial do erro, foram realizadas duas classificações (MaxVer e RNA) na imagem Ikonos da região da Serra do Salitre - MG. A Tabela 21 apresenta a matriz de confusão resultante da classificação pelo algoritmo da Máxima Verossimilhança (MaxVer) e os índices derivados dessa matriz para verificar erros oriundos do processo de classificação. As formulações desses índices se encontram na seção 4.1.1, sendo que o índice Tau foi calculado a partir da Equação (17) tanto para classificação MaxVer como para a RNA.

Tabela 21 – Matriz de confusão resultante da classificação MaxVer e os índices de exatidão.

Classes	Dados de referência							total	Usuário (%)
	1	2	3	4	5	6	7		
1	92	0	0	0	0	43	0	135	68,1
2	1	117	0	0	1	0	0	119	98,3
3	0	0	116	0	0	0	0	116	100
4	0	0	0	120	0	2	0	122	98,4
5	0	3	4	0	119	0	0	126	94,4
6	27	0	0	0	0	72	0	99	72,7
7	0	0	0	0	0	3	120	123	97,6
<b>total</b>	120	120	120	120	120	120	120	840	
<b>Produtor (%)</b>	76,7	97,5	96,7	100	99,2	60,0	100		
<b>Exatidão global</b>	90%								
<b>Limite inferior (s)</b>	87,9%								
<b>Kappa (K)</b>	88,3%	Var <sub>K</sub> = 0,000146		Z = 73,227					
<b>Tau (T)</b>	88,3%	Var <sub>T</sub> = 0,000146		Z = 73,147					

A Tabela 22 apresenta os resultados da matriz de confusão obtidos no processo de classificação por redes neurais (RNA) e os índices derivados da a partir da mesma.

Tabela 22 – Matriz de confusão resultante da classificação RNA e os índices de exatidão.

Classes	Dados de referência							total	Usuário (%)
	1	2	3	4	5	6	7		
1	91	0	0	0	0	5	1	97	93,8
2	4	112	1	0	0	1	0	118	94,9
3	0	0	114	2	0	0	0	116	98,3
4	0	0	0	117	1	0	0	118	99,2
5	0	3	5	0	118	1	0	127	92,9
6	25	5	0	1	1	107	0	139	77,0
7	0	0	0	0	0	6	119	125	95,2
<b>total</b>	120	120	120	120	120	120	120	840	
<b>Produtor (%)</b>	75,8	93,3	95	97,5	98,3	89,2	99,2		
<b>Exatidão global</b>	92,6%								
<b>Limite inferior (s)</b>	91,1%								
<b>Kappa (K)</b>	91,4%	Var <sub>K</sub> = 0,000111		Z = 86,900					
<b>Tau (T)</b>	91,4%	Var <sub>T</sub> = 0,000111		Z = 86,708					

O resultado dos índices Kappa e Tau foram iguais para as duas classificações, porém na classificação RNA apresentou um resultado melhor. As variâncias foram praticamente iguais, diferindo em algumas casas decimais. O limite de confiança inferior nos indica que 95% das amostras de treinamento com precisões observadas de 90% (MaxVer) e 92,6% (RNA) tiveram verdadeiras precisões de 87,9% e 91,1% respectivamente ou maior. O teste estatístico Z apresentou um valor maior que o valor de  $Z_{\alpha/2} = 1,96$  ( $\alpha=0,95$ ) para ambos os índices Kappa e Tau, significando que o nível de acerto da classificação e os dados de referência foram significativos.

A partir da imagem temática, gerada pela classificação MaxVer, gerou-se a imagem de distância, em escala de cinza (0-255), apresentada na Figura 27. Esta imagem nos permite visualizar as discrepâncias de *pixels* individuais dos meios das classes para as quais foram atribuídos, onde os *pixels* mais escuros indicam que são mais prováveis de serem classificados corretamente por estarem próximos ao centróide da classe, e já os *pixels* em tons de cinza altos são mais prováveis de serem classificados erroneamente.

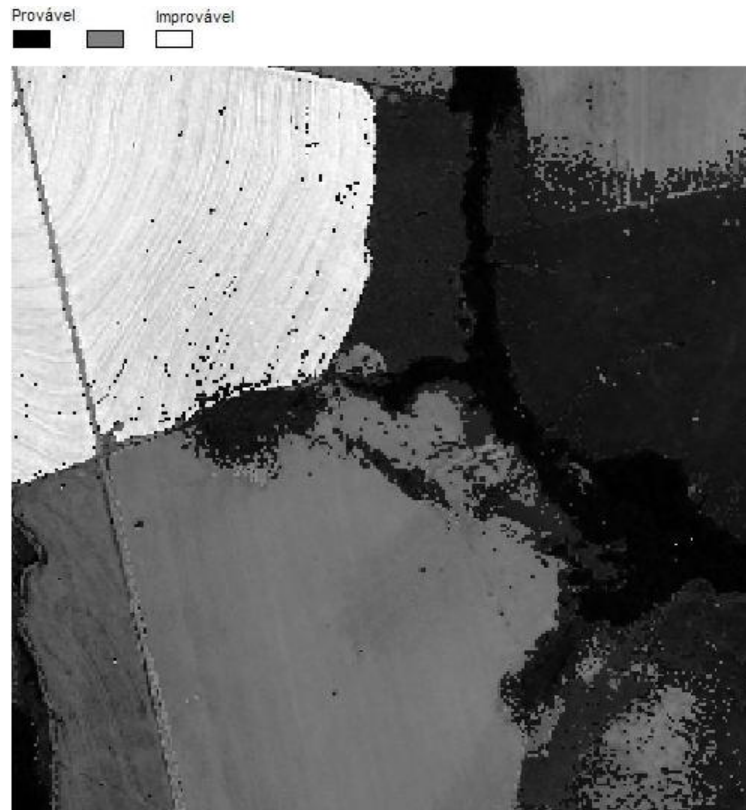


Figura 27: Imagem de distância.

A Figura 28 apresenta uma outra forma de espacializar os erros de uma classificação através da comparação da imagem temática com o respectivo mapa de referência, resultando em uma imagem binária do erro, onde o *pixel* que foi rotulado corretamente recebeu o valor "0" e o rotulado erroneamente recebeu o valor "1". Podemos notar que, apesar dos *pixels* da área correspondente ao plantio de café terem apresentado uma grande probabilidade de serem mal classificados, pelo resultado da imagem de distância, eles foram quase que totalmente classificados corretamente. É possível notar também que as áreas que aparecem concentradas na imagem binária do erro, tinham uma probabilidade média de serem classificadas erroneamente como nota-se na imagem de distância. Percebe-se que essa concentração do erro (confusão) ocorreu entre as classes campo limpo e pasto. Para as demais classes da imagem o resultado foi muito bom.

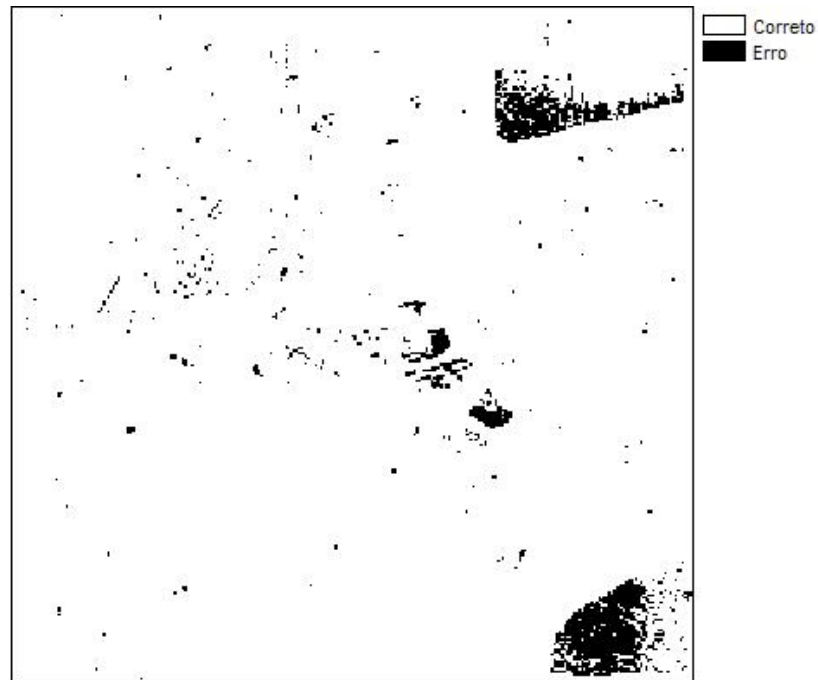


Figura 28: Imagem binária do erro.

Da comparação da imagem temática com o respectivo mapa de referência, gerou-se uma imagem de diferença, onde duas imagens temáticas diferentes (uma gerada pelo algoritmo MaxVer e outra por RNA) foram comparadas com o mapa de referência, como é apresentado na Figura 29. Dessa imagem de diferença pode-se notar que a maior confusão na classificação, para ambos os métodos, ocorre entre as classes campo limpo e pastagem.

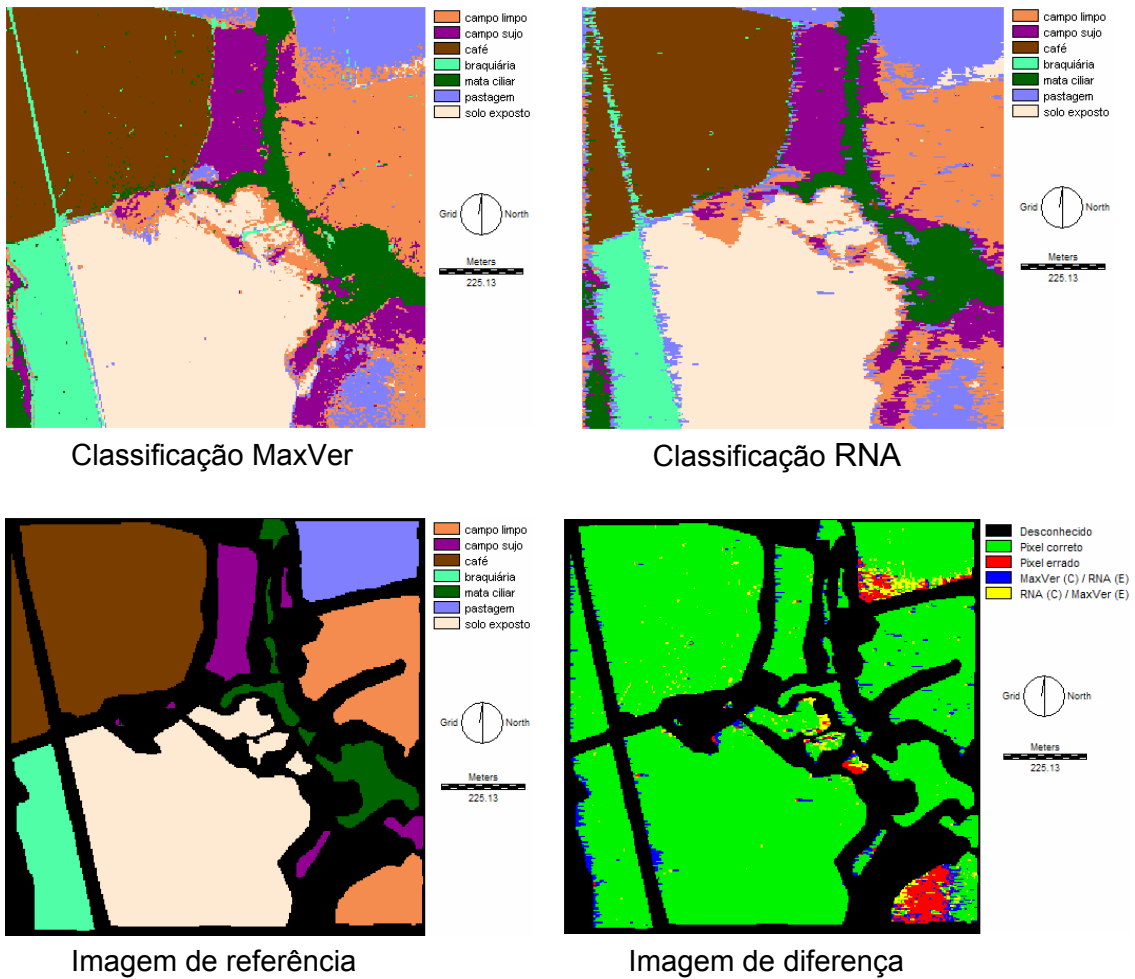


Figura 29: Imagem de diferença entre as classificações MaxVer e RNA.

Os *pixels* correspondentes que foram corretamente rotulados nas duas imagens temáticas receberam o valor 1 (verde) na imagem de diferença. Já os *pixels* correspondentes, para ambas as imagens, que foram erroneamente rotulados receberam o valor 2 (vermelho). O valor 3 (azul) foi atribuído aos *pixels* que foram corretamente rotulados (C) na classificação MaxVer mas foram erroneamente rotulados (E) na classificação RNA. Entretanto, o valor 4 (amarelo) foi atribuído aos *pixels* que foram corretamente rotulados (C) na classificação RNA, porém foram erroneamente rotulados (E) na classificação MaxVer.

Apesar dos índices de exatidão temática terem apresentado bons resultados, a espacialização do erro distinguiu melhor onde ocorreram as confusões no processo de classificação, mostrando ser uma forma mais eficiente de avaliar a qualidade da classificação.

## 6. CONCLUSÃO

Este trabalho apresentou aos usuários metodologias para acessar a exatidão posicional de imagens orbitais, além de apresentar formas para expressar espacialmente a exatidão de uma classificação temática.

Em termos cartográficos, conclui-se, a partir dos resultados desta dissertação, que os produtos do sensoriamento remoto possuem exatidão posicional, porém, deve-se observar para qual escala de trabalho a imagem poderá ser utilizada, verificando se a mesma possui precisão para o fim desejado. Neste estudo, utilizando o Padrão de Exatidão Cartográfica (PEC), as imagens Quickbird e Ikonos apresentaram resultados classe A para escala 1:10.000, sendo uma excelente fonte para efetuar atualização, auxiliar o mapeamento sistemático, dar apoio ao cadastro técnico, entre outras finalidades do mapeamento cartográfico. Na escala 1:25.000, as imagens CBERS e Landsat não apresentaram exatidão posicional, porém, para escala de trabalhos menores, essas imagens são bastante utilizadas na representação do uso e ocupação do solo, dando suporte ao trabalho de gestão ambiental e territorial, e ao monitoramento de áreas de preservação. A imagem MODIS, para a escala de 1:400.000, apresentou um resultado classe A, sendo uma imagem confiável para realizar quantificação e detecção das mudanças da cobertura terrestre, além de ser bastante utilizada para realizar a classificação de mapeamento em escala regional.

Com relação aos Métodos das Áreas e dos Retângulos Equivalentes, que também são utilizados para avaliar a exatidão posicional, sugere-se a realização de estudos para estabelecer valores críticos (tabelados) para possibilitar realizar testes estatísticos para avaliar se os resultados obtidos são satisfatório ou não.

Dos resultados da análise não-espacial na exatidão temática, conclui-se que o número de padrões de validação e/ou de classes informacionais influenciam no cálculo dos índices de exatidão temática e os resultados são mais afetados quando ocorre a presença de classe informacional predominante entre os valores amostrados, para validação. Embora a exatidão Global apresente um valor mais alto, os índices de concordância Kappa e Tau são mais consistentes por envolver no valor final todas as células da matriz de confusão. Como os índices Kappa e Tau apresentam valores bem próximos, ao se realizar uma avaliação da exatidão de mapas temáticos pode-se utilizar apenas o índice Tau, pois este possui uma formulação mais simples.

Sugere-se para trabalhos futuros, a realização de outro teste a fim de verificar se há possibilidades de através de um índice de exatidão temática específico obter o(s) valor(es) de outro(s) índice(s) de exatidão. Por exemplo, do resultado desta pesquisa poderia se pensar na possibilidade de se estimar a diferença entre os índices Global e Tau através da diferença entre os índices Global e Kappa, pelo fato de apresentarem comportamento semelhante.

A validação é essencial no processo de mapeamento e uma das formas mais eficientes para representar a exatidão de uma classificação temática está baseada na matriz de confusão, de onde deriva-se as análises não-espaciais. Porém, essas análises não consideram a distribuição espacial do erro, implícita ou explicitamente, na classificação. As metodologias, utilizadas nesta dissertação, para representação e visualização da confiabilidade espacial *pixel a pixel*, forneceram uma boa avaliação quantitativa da exatidão na imagem temática, possibilitando identificar as áreas que possuíam baixa exatidão.

## REFERÊNCIA BIBLIOGRÁFICA

Alves, R. A. L., Vergara, O. R. **Identificação de alvos urbanos em imagens Ikonos, aplicando classificação orientada a segmentos.** Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil. p. 2573-2580, 2005.

Amorim, G. P. **Confiabilidade de rede GPS de referência cadastral municipal; Estudo de caso: rede do município de Vitória-ES.** Escola de Engenharia de São Carlos – Universidade de São Paulo, São Paulo. 2004, 149p.

Arce, E. C. Rudorff, B. F. T. Berka, L. M. S; Pereira, P. F. C. **Imágenes Landsat y CBERS en la identificación del uso del suelo con cultivos agrícolas.** Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil. p. 41-48, 2005.

Banko, G. **A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory.** International Institute for Applied Systems Analysis (IIASA). 1998. Disponível em: <<http://www.iiasa.ac.at/Publications/Documents/IR-98-081.pdf>> Acesso em: 22 jun 2006.

Bernardes, T. **Caracterização do ambiente agrícola do Complexo Serra Negra por meio de sensoriamento remoto e sistemas de informação geográfica.** Dissertação (Mestrado). Universidade Federal de Lavras, Minas Gerais. 2006, 119p. Disponível em: <<http://www.epamig.br/geosolos/MaterialSite/Teses/Tiago/Tiagodissertacao.pdf>>. Acesso em: 13 out. 2006.

Brites, R. S., Soares, V. P., Ribeiro, C. A. A. S. **Verificação da exatidão em classificações de uma imagem orbital mediante a utilização de três índices.** Revista Árvore, Viçosa-MG. v. 20, nº 3, p.415-424, 1996.

Congalton, R. G. **A review of assessing the accuracy of classifications of remotely sensed data.** Remote Sensing of Environment. 49(12), p. 1671-1678, 1991.

Congalton, R. G., and Green, K. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.** New York: Lewis Publishers. 1999, 137p.

Crósta, A. P. **Processamento Digital de Imagens de Sensoriamento Remoto.** Campinas, IG/UNICAMP. 1993, 170p.

Dicks, S., Lo, T. **Evaluation of thematic map accuracy in a land-use and land-cover mapping program.** Photogrammetric Engineering and Remote Sensing. 56(9), p. 1247-1252, 1990.

Ferreira, L. F., Cintra, J. P. **Quantificação de discrepâncias entre feições lineares por retângulos equivalentes.** Revista Brasileira de Cartografia. nº 51, p. 1-8, 1999.

Fonseca, L. M. G. **Processamento digital de imagens.** Instituto Nacional de Pesquisas Espaciais (INPE). 2000, 105p.

Foody, G. M. **On the compensation for chance agreement in image classification accuracy assessment.** Photogrammetric Engineering and Remote Sensing. v. 58, nº 10, p. 1459-1460, 1992.

Foody, G. M. **Status of land cover classification accuracy assessment.** Remote Sensing of Environment. v. 80, p. 185– 201, 2002.

Galo, M., Camargo, P. O. **O uso do GPS no controle de qualidade de cartas.** Anais 1º Congresso Brasileiro de Cadastro Técnico Multifinalitário, Florianópolis – SC. Tomo II, p.41-48, 1994.

Galo, M., Dal Poz, A. P., Ferreira, F. M. **O uso de feições no controle de qualidade em cartografia.** Anais XIX Congresso Brasileiro de Cartografia, Porto Alegre – RS. 2001.

Hay, A. **Sampling designs to test land-use map accuracy.** Photogrammetric Engineering and Remote Sensing. 45(4), p. 529-533, 1979.

IBGE, Instituto Brasileiro de Geografia e Estatística. Disponível em: <<http://www.ibge.gov.br/cidadesat>>. Acesso em: 05 abr. 2007.

Liang, S. **Quantitative Remote Sensing of Land Surfaces.** Canadá, USA. 2004, 534p.

Lingnau, C., Antunes, A. F. B. **Avaliação da Acurácia de Mapas Temáticos Oriundos de Classificação Digital de Imagens**. Centro Integrado de Estudos em Geoprocessamentos (CIEG). Universidade Federal do Paraná, Curitiba. 2003, 10p. Disponível em: <<http://www.cieg.ufpr.br/>>. Acesso em: 19 out. 2006.

Lugnani, J. B. **Estimativa de qualidade de feições digitalizadas - Um novo método**, Revista Brasileira de Cartografia. nº 39, p. 26-29, 1986.

Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jessen, J. H. and McGwire, K. C. **Remote sensing and geographic information system data integration: error sources and research issues**. Photogrammetric Engineering and Remote Sensing. v. 57, nº 6, p. 677-687, 1991.

Ma, Z., Redmond, R. L. **Tau coefficients for accuracy assessment of classification of remote sensing data**. Photogrammetric Engineering and Remote Sensing. v. 61, nº 4, p. 435-439, 1995.

Marotta, G. S., Vieira, C. A. O. **Aplicação do Padrão de Exatidão Cartográfica em imagens orbitais Aster para fins de atualizações de mapeamentos**. Anais XIX Congresso Brasileiro de Cartografia, Macaé – RJ. 2005.

Mather, P. M. **Computer processing of remotely sensed images: an introduction**. School of Geography, the University of Nottingham, UK. 1999, 292p.

Merchant, D. C. **Spatial Accuracy Standards for Large Scale Line Maps**. In Proceedings of the Technical Congress on Surveying and Mapping. v 1, p. 222-231, 1982.

NBR13133. **Execução de Levantamento Topográfico**. ABNT – Associação Brasileira de Normas Técnicas. Rio de Janeiro. Maio de 1994.

Nero, M. A. **Propostas para o controle de qualidade de bases cartográficas com ênfase na componente posicional**. Tese (Doutorado). Escola Politécnica, USP, São Paulo. 2005, 181p.

Neto, J. O. A. **Análise da precisão e acurácia de pontos georreferenciados com a técnica do código suavizado pela fase da portadora utilizando GPS de simples frequência**. Dissertação (Mestrado). Escola de Engenharia de São Carlos – Universidade de São Paulo, São Paulo. 2006, 207p.

Normas Técnicas da Cartografia Nacional. **Decreto nº 89.817** de 20 de Junho de 1984. Disponível em: <<http://www.concar.ibge.gov.br/indexf7a0.html?q=node/41>> Acesso em: 14 jul. 2006.

Rosefield, G. **Analysis of variance of thematic mapping experiment data.** Photogrammetric Engineering and Remote Sensing. 47(12), p. 1685-1692, 1982.

Rosefield, G., Fitzpatrick-Lins, K., Ling, H. **Sampling for thematic accuracy testing.** Photogrammetric Engineering and Remote Sensing. 48(1), p. 131-137, 1982.

Tisot, D. A., Formaggio, A. R., Galvão, L. S., Rennó, C. D. **Dados Hyperion e ETM+ na classificação de alvos agrícolas com diferentes níveis de distinção espectral.** Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil. p. 439-446, 2005.

Vieira, C. A. O. **Accuracy of remotely sensing classification of agricultural crops: a comparative study.** 2000, 327p. Thesis (Doctor of Philosophy). University of Nottingham. Cap. 5, p. 128-175, 2000.

Vieira, C. A. O., Mather, P. M.; Borges, P. A. F. **Assessing the Positional Accuracy of Remotely Sensed Products.** In. 5th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Melbourne, Australia, 2002.

Vieira, C. A. O., Mather, P. M. **Techniques for estimating the positional and thematic accuracy of remotely sensed products.** Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil. p. 4351-4359, 2005.

Zell, A., Mamier, G., Vogt, M., Mache, N., Hubner, R.; Doring, S., Herrmann, K., Soyez, T., Schmalzl, M., Sommer, T., Hatzigeorgiou, A., Posselt, D., Schreiner, T., Kett, B., Clemente, G., Wieland, J., Gatter, J. **Stuttgart Neural Network Simulator v4.2.** University of Stuttgart: Institute for Parallel and Distributed High Performance Systems/University of Tübingen - Wilhelm-Schickard-Institute for Computer Science – Department of Computer Architecture. 1996.

## APÊNDICE A

Valores dos índices Kappa, Global e Tau, obtidos das matrizes de confusão utilizadas para o experimento no nível de desempenho excelente (0,8 a 1,0).

### A1. Mantendo o total de pixels fixo e variando o número de classes.

Número de classes = 2					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,98522	0,96957	0,95130	0,94696	0,94522
K	0,96996	0,93833	0,89479	0,88768	0,88146
T	0,97043	0,93913	0,90261	0,89391	0,89043

Número de classes = 3					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,97391	0,95304	0,96261	0,91130	0,94783
K	0,95926	0,92699	0,94353	0,86499	0,91865
T	0,96087	0,92957	0,94391	0,86696	0,92174

Número de classes = 4					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,96783	0,94696	0,92609	0,90870	0,91826
K	0,95549	0,92671	0,89969	0,87654	0,88709
T	0,95710	0,92928	0,90145	0,87826	0,89101

Número de classes = 5					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,96435	0,95739	0,94174	0,89130	0,92870
K	0,95465	0,94588	0,92600	0,86135	0,90965
T	0,95543	0,94674	0,92717	0,86413	0,91087

Número de classes = 7					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,92696	0,91304	0,89652	0,83913	0,82870
K	0,91269	0,89608	0,87752	0,80369	0,79796
T	0,91478	0,89855	0,87928	0,81232	0,80014

Número de classes = 9					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,90783	0,86522	0,83913	0,83043	0,91739
K	0,89400	0,84539	0,81558	0,80640	0,90565
T	0,89630	0,84837	0,81902	0,80924	0,90707

Número de classes = 11					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,85130	0,82783	0,91739	0,85043	0,82696
K	0,83171	0,80706	0,90714	0,83232	0,80845
T	0,83643	0,81061	0,90913	0,83548	0,80965

Número de classes = 13					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,87826	0,84174	0,82174	0,91913	0,85391
K	0,86728	0,82780	0,80526	0,91185	0,84049
T	0,86812	0,82855	0,80688	0,91239	0,84174

Número de classes = 15					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,88696	0,86000	0,85130	0,92174	0,83043
K	0,87842	0,84942	0,83962	0,91591	0,81745
T	0,87888	0,85000	0,84068	0,91615	0,81832

**A2.** Mantendo o número de classes fixo e variando o número total de pixels.

Total de pixels = 750					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,92533	0,90267	0,83600	0,87733	0,85600
K	0,91230	0,88570	0,80541	0,85449	0,82675
T	0,91289	0,88644	0,80867	0,85689	0,83200

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,91385	0,88615	0,83538	0,85385	0,84769
K	0,89868	0,86543	0,80408	0,82239	0,81451
T	0,89949	0,86718	0,80795	0,82949	0,82231

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,92650	0,89050	0,86600	0,90450	0,83600
K	0,91324	0,87134	0,83813	0,88377	0,80229
T	0,91425	0,87225	0,84367	0,88858	0,80867

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,93480	0,95160	0,86600	0,88200	0,84160
K	0,92276	0,94263	0,83852	0,85742	0,80693
T	0,92393	0,94353	0,84367	0,86233	0,81520

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,96081	0,93649	0,89351	0,86351	0,91135
K	0,95326	0,92429	0,87130	0,83407	0,89368
T	0,95428	0,92590	0,87577	0,84077	0,89658

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,95286	0,92429	0,97000	0,90857	0,87881
K	0,94352	0,90801	0,96402	0,88955	0,85362
T	0,94500	0,91167	0,96500	0,89333	0,85861

**A3.** Mantendo o número de classes fixo, variando o número total de pixels e com presença de classe informacional predominante, para validação.

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,92462	0,89154	0,87538	0,91385	0,89923
K	0,85863	0,77566	0,79096	0,86745	0,85254
T	0,91205	0,87346	0,85462	0,89949	0,88244

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,92650	0,90600	0,88800	0,93850	0,86700
K	0,89667	0,85971	0,83836	0,90908	0,81055
T	0,91425	0,89033	0,86933	0,92825	0,84483

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,95040	0,91600	0,88720	0,87600	0,92760
K	0,90754	0,85375	0,83804	0,80849	0,88273
T	0,94213	0,90200	0,86840	0,85533	0,91553

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,96081	0,93216	0,89459	0,90973	0,88784
K	0,94077	0,90984	0,86022	0,86065	0,83828
T	0,95428	0,92086	0,87703	0,89468	0,86914

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,96119	0,93143	0,90667	0,88952	0,87238
K	0,94032	0,90622	0,86655	0,86034	0,80971
T	0,95472	0,92000	0,89111	0,87111	0,85111

## APÊNDICE B

Valores dos índices Kappa, Global e Tau, obtidos das matrizes de confusão utilizadas para o experimento no nível de desempenho muito bom (0,6 a 0,8).

**B1.** Mantendo o total de pixels fixo e variando o número de classes.

Número de classes = 2					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,80087	0,80348	0,82522	0,87565	0,88087
K	0,60067	0,61575	0,63176	0,74452	0,75346
T	0,60174	0,60696	0,65043	0,75130	0,76174

Número de classes = 3					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,86087	0,83739	0,78870	0,75043	0,80435
K	0,78634	0,75160	0,67952	0,62110	0,70153
T	0,79130	0,75609	0,68304	0,62565	0,70652

Número de classes = 4					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,79478	0,84957	0,75913	0,72522	0,78870
K	0,72197	0,79378	0,67735	0,62720	0,70989
T	0,72638	0,79942	0,67884	0,63362	0,71826

Número de classes = 5					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,79565	0,83913	0,81130	0,72000	0,74522
K	0,73787	0,79437	0,75748	0,64173	0,67248
T	0,74457	0,79891	0,76413	0,65000	0,68152

Número de classes = 7					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,69130	0,75217	0,80261	0,79043	0,70609
K	0,63317	0,70462	0,76659	0,74395	0,65221
T	0,63986	0,71087	0,76971	0,75551	0,65710

Número de classes = 9					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,78609	0,70000	0,75130	0,66435	0,81217
K	0,75273	0,65558	0,71388	0,61877	0,78375
T	0,75935	0,66250	0,72022	0,62239	0,78870

Número de classes = 11					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,78870	0,66435	0,71739	0,75826	0,71391
K	0,76069	0,62523	0,68535	0,72725	0,68364
T	0,76757	0,63078	0,68913	0,73409	0,68530

Número de classes = 13					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,81130	0,69304	0,71565	0,76957	0,66696
K	0,79377	0,66422	0,69040	0,74886	0,63649
T	0,79558	0,66746	0,69196	0,75036	0,63920

Número de classes = 15					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,72783	0,80435	0,69739	0,67217	0,73826
K	0,70743	0,78918	0,67439	0,64637	0,71820
T	0,70839	0,79037	0,67578	0,64876	0,71957

**B2.** Mantendo o número de classes fixo e variando o número total de pixels.

Total de pixels = 750					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,81467	0,76133	0,72267	0,80533	0,67333
K	0,78250	0,72072	0,67397	0,77148	0,61512
T	0,78378	0,72156	0,67644	0,77289	0,61889

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,71154	0,72615	0,74692	0,69154	0,74000
K	0,65692	0,67692	0,69716	0,63116	0,68913
T	0,66346	0,68051	0,70474	0,64013	0,69667

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,82250	0,73700	0,71350	0,80450	0,81100
K	0,79071	0,69144	0,65255	0,76633	0,77267
T	0,79292	0,69317	0,66575	0,77192	0,77950

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,72000	0,75960	0,72880	0,67880	0,77360
K	0,66696	0,71719	0,67495	0,60983	0,72817
T	0,67333	0,71953	0,68360	0,62527	0,73587

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,81027	0,74865	0,67676	0,75973	0,72514
K	0,77600	0,70164	0,61160	0,71251	0,67603
T	0,77865	0,70676	0,62288	0,71968	0,67932

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,73381	0,82667	0,76667	0,69476	0,81452
K	0,68720	0,78999	0,72501	0,63318	0,77756
T	0,68944	0,79778	0,72778	0,64389	0,78361

**B3.** Mantendo o número de classes fixo, variando o número total de pixels e com presença de classe informacional predominante, para validação.

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,78615	0,76615	0,81846	0,80846	0,74538
K	0,66892	0,61319	0,72930	0,73350	0,65446
T	0,75051	0,72718	0,78821	0,77654	0,70295

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,82650	0,79100	0,70300	0,77550	0,76750
K	0,77269	0,71065	0,60774	0,69380	0,68109
T	0,79758	0,75617	0,65350	0,73808	0,72875

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,82600	0,75480	0,79920	0,75600	0,81720
K	0,72608	0,64060	0,72211	0,65630	0,73981
T	0,79700	0,71393	0,76573	0,71533	0,78673

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,77162	0,73027	0,78730	0,70486	0,76622
K	0,70177	0,65858	0,72865	0,61249	0,68808
T	0,73356	0,68532	0,75185	0,65568	0,72725

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,72310	0,75595	0,81333	0,82571	0,77690
K	0,65356	0,68950	0,76027	0,78552	0,68747
T	0,67694	0,71528	0,78222	0,79667	0,73972

## APÊNDICE C

Valores dos índices Kappa, Global e Tau, obtidos das matrizes de confusão utilizadas para o experimento no nível de desempenho bom (0,4 a 0,6).

**C1.** Mantendo o total de pixels fixo e variando o número de classes.

Número de classes = 2					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,70696	0,73391	0,76348	0,79652	0,78435
K	0,40918	0,46497	0,52089	0,58082	0,56124
T	0,41391	0,46783	0,52696	0,59304	0,56870

Número de classes = 3					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,73043	0,62696	0,70174	0,66348	0,73130
K	0,59042	0,43899	0,54539	0,48695	0,58271
T	0,59565	0,44043	0,55261	0,49522	0,59696

Número de classes = 4					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,69913	0,67391	0,63739	0,61217	0,57565
K	0,59468	0,55329	0,51414	0,48166	0,42345
T	0,59884	0,56522	0,51652	0,48290	0,43420

Número de classes = 5					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,67826	0,64522	0,61652	0,59391	0,55217
K	0,58931	0,55016	0,51090	0,48356	0,43163
T	0,59783	0,55652	0,52065	0,49239	0,44022

Número de classes = 7					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,62522	0,65043	0,58957	0,55565	0,51652
K	0,55702	0,58455	0,51900	0,47449	0,43010
T	0,56275	0,59217	0,52116	0,48159	0,43594

Número de classes = 9					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,61913	0,58609	0,55478	0,56783	0,50609
K	0,56780	0,53267	0,49826	0,51192	0,44386
T	0,57152	0,53435	0,49913	0,51380	0,44435

Número de classes = 11					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,56435	0,62870	0,58609	0,53565	0,50174
K	0,51859	0,58756	0,54360	0,48777	0,45058
T	0,52078	0,59157	0,54470	0,48922	0,45191

Número de classes = 13					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,60087	0,62348	0,56174	0,52261	0,48174
K	0,56642	0,58946	0,52353	0,48216	0,43769
T	0,56761	0,59210	0,52522	0,48283	0,43855

Número de classes = 15					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,60957	0,57913	0,54348	0,52261	0,47304
K	0,58106	0,54753	0,50919	0,48676	0,43247
T	0,58168	0,54907	0,51087	0,48851	0,43540

**C2.** Mantendo o número de classes fixo e variando o número total de pixels.

Total de pixels = 750					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,63733	0,65067	0,59867	0,56000	0,51333
K	0,57705	0,59141	0,52599	0,48682	0,43284
T	0,57689	0,59244	0,53178	0,48667	0,43222

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,58769	0,54923	0,64308	0,52462	0,50538
K	0,51410	0,47300	0,57943	0,44444	0,41839
T	0,51897	0,47410	0,58359	0,44538	0,42295

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,64250	0,57950	0,60300	0,54400	0,52550
K	0,58360	0,50848	0,52565	0,45769	0,44057
T	0,58292	0,50942	0,53683	0,46800	0,44642

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,60760	0,62880	0,56120	0,65120	0,52240
K	0,53808	0,56702	0,48199	0,58645	0,44460
T	0,54220	0,56693	0,48807	0,59307	0,44280

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,62351	0,65027	0,59027	0,55541	0,53000
K	0,55964	0,58924	0,51491	0,47780	0,45110
T	0,56077	0,59198	0,52198	0,48131	0,45167

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,63857	0,65190	0,60000	0,55190	0,53167
K	0,57714	0,58928	0,53175	0,46836	0,45234
T	0,57833	0,59389	0,53333	0,47722	0,45361

**C3.** Mantendo o número de classes fixo, variando o número total de pixels e com presença de classe informacional predominante, para validação.

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,63231	0,65308	0,61846	0,55308	0,58231
K	0,51139	0,48965	0,47856	0,43133	0,48212
T	0,57103	0,59526	0,55487	0,47859	0,51269

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,65300	0,62950	0,55300	0,57550	0,57000
K	0,57050	0,53600	0,44376	0,46313	0,46568
T	0,59517	0,56775	0,47850	0,50475	0,49833

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,62600	0,63680	0,62760	0,65160	0,55320
K	0,50195	0,52266	0,52786	0,53935	0,43280
T	0,56367	0,57627	0,56553	0,59353	0,47873

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,64811	0,62216	0,62514	0,59135	0,57865
K	0,55527	0,52965	0,51908	0,48179	0,45840
T	0,58946	0,55919	0,56266	0,52324	0,50842

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,64929	0,62095	0,61786	0,58429	0,58714
K	0,56561	0,51832	0,49690	0,47966	0,42084
T	0,59083	0,55778	0,55417	0,51500	0,51833

## APÊNDICE D

Valores dos índices Kappa, Global e Tau, obtidos das matrizes de confusão utilizadas para o experimento no nível de desempenho razoável (0,2 a 0,4).

**D1.** Mantendo o total de pixels fixo e variando o número de classes.

Número de classes = 2					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,60435	0,63826	0,69478	0,68609	0,67565
K	0,20359	0,27621	0,37229	0,36795	0,34851
T	0,20870	0,27652	0,38957	0,37217	0,35130

Número de classes = 3					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,54609	0,58348	0,59739	0,48609	0,51826
K	0,30181	0,37718	0,39423	0,23026	0,27374
T	0,31913	0,37522	0,39609	0,22913	0,27739

Número de classes = 4					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,53043	0,54435	0,44609	0,41478	0,42870
K	0,37447	0,38773	0,26135	0,22214	0,23020
T	0,37391	0,39246	0,26145	0,21971	0,23826

Número de classes = 5					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,51739	0,48261	0,46000	0,41826	0,37043
K	0,39528	0,35035	0,31962	0,26918	0,21174
T	0,39674	0,35326	0,32500	0,27283	0,21304

Número de classes = 7					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,47739	0,46435	0,43652	0,36696	0,35495
K	0,39171	0,37205	0,34104	0,26205	0,24116
T	0,39029	0,37507	0,34261	0,26145	0,24745

Número de classes = 9					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,45217	0,42087	0,40435	0,37217	0,36087
K	0,38425	0,35023	0,32897	0,29310	0,27999
T	0,38370	0,34848	0,32989	0,29370	0,28098

Número de classes = 11					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,44609	0,43217	0,38609	0,35565	0,33304
K	0,38897	0,37299	0,32388	0,28832	0,26405
T	0,39070	0,37539	0,32470	0,29122	0,26635

Número de classes = 13					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,43391	0,39913	0,40087	0,35913	0,31739
K	0,38541	0,34461	0,35060	0,30569	0,25935
T	0,38674	0,34906	0,35094	0,30572	0,26051

Número de classes = 15					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,41043	0,43130	0,35217	0,32609	0,29217
K	0,36737	0,38994	0,30385	0,27583	0,24001
T	0,36832	0,39068	0,30590	0,27795	0,24161

**D2.** Mantendo o número de classes fixo e variando o número total de pixels.

Total de pixels = 750					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,47733	0,43733	0,41200	0,36267	0,33733
K	0,39036	0,34095	0,30475	0,25235	0,22327
T	0,39022	0,34356	0,31400	0,25644	0,22689

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,42692	0,39538	0,34231	0,44769	0,47462
K	0,32970	0,29395	0,22865	0,35491	0,38461
T	0,33141	0,29462	0,23269	0,35564	0,38705

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,47900	0,38800	0,33200	0,46000	0,44500
K	0,39400	0,28247	0,21921	0,36660	0,34857
T	0,39217	0,28600	0,22067	0,37000	0,35250

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,48320	0,46000	0,44240	0,39840	0,37640
K	0,39531	0,36936	0,34830	0,29543	0,27619
T	0,39707	0,37000	0,34947	0,29813	0,27247

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,48027	0,42459	0,40162	0,38108	0,37108
K	0,39425	0,32472	0,29874	0,27666	0,26590
T	0,39365	0,32869	0,30189	0,27793	0,26626

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,48357	0,46143	0,41905	0,36714	0,38405
K	0,39722	0,36986	0,32279	0,25733	0,28212
T	0,39750	0,37167	0,32222	0,26167	0,28139

**D3.** Mantendo o número de classes fixo, variando o número total de pixels e com presença de classe informacional predominante, para validação.

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,48385	0,45769	0,41846	0,38385	0,35769
K	0,36695	0,32518	0,27515	0,24932	0,23613
T	0,39782	0,36731	0,32154	0,28115	0,25064

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,45800	0,48050	0,42300	0,40050	0,37950
K	0,36204	0,38411	0,30266	0,26032	0,24654
T	0,36767	0,39392	0,32683	0,30058	0,27608

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,45000	0,47280	0,39560	0,41560	0,37880
K	0,30710	0,35535	0,24871	0,29835	0,25389
T	0,35833	0,38493	0,29487	0,31820	0,27527

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,46432	0,47784	0,41243	0,39135	0,38135
K	0,30624	0,36619	0,28516	0,24111	0,24877
T	0,37505	0,39081	0,31450	0,28991	0,27824

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,47952	0,47571	0,46310	0,38905	0,37857
K	0,33636	0,35866	0,29054	0,25285	0,23691
T	0,39278	0,38833	0,37361	0,28722	0,27500

## APÊNDICE E

Valores dos índices Kappa, Global e Tau, obtidos das matrizes de confusão utilizadas para o experimento no nível de desempenho ruim (0 a 0,2).

### E1. Mantendo o total de pixels fixo e variando o número de classes.

Número de classes = 2					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,56087	0,59652	0,52087	0,58174	0,54435
K	0,12158	0,19199	0,03263	0,16355	0,09115
T	0,12174	0,19304	0,04174	0,16348	0,08870

Número de classes = 3					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,46609	0,44435	0,43217	0,39739	0,38087
K	0,19587	0,16066	0,13716	0,09458	0,07202
T	0,19913	0,16652	0,14826	0,09609	0,07130

Número de classes = 4					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,39739	0,37304	0,33043	0,31565	0,28957
K	0,19660	0,16286	0,10938	0,08727	0,05270
T	0,19652	0,16406	0,10725	0,08754	0,05275

Número de classes = 5					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,35652	0,33478	0,27739	0,30609	0,23130
K	0,19895	0,16879	0,09061	0,12835	0,03412
T	0,19565	0,16848	0,09674	0,13261	0,03913

Número de classes = 7					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,23913	0,23217	0,30087	0,21913	0,18435
K	0,12242	0,10803	0,18299	0,09039	0,04773
T	0,11232	0,10420	0,18435	0,08899	0,04841

Número de classes = 9					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,23565	0,20957	0,19652	0,28261	0,18000
K	0,14176	0,11485	0,09523	0,19210	0,07657
T	0,14011	0,11076	0,09609	0,19293	0,07750

Número de classes = 11					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,22087	0,20696	0,17478	0,16348	0,26609
K	0,14147	0,12700	0,09134	0,07595	0,19196
T	0,14296	0,12765	0,09226	0,07983	0,19270

Número de classes = 13					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,25739	0,21304	0,18696	0,16261	0,13826
K	0,19595	0,14578	0,11865	0,09267	0,06681
T	0,19551	0,14746	0,11920	0,09283	0,06645

Número de classes = 15					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,21043	0,24957	0,18435	0,14870	0,13130
K	0,15538	0,19635	0,12571	0,08727	0,06899
T	0,15404	0,19596	0,12609	0,08789	0,06925

**E2.** Mantendo o número de classes fixo e variando o número total de pixels.

Total de pixels = 750					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,31200	0,29067	0,23333	0,18933	0,17867
K	0,19936	0,16580	0,09963	0,04768	0,03501
T	0,19733	0,17244	0,10556	0,05422	0,04178

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,25000	0,19538	0,18077	0,30923	0,29538
K	0,12512	0,06277	0,04343	0,19473	0,16848
T	0,12500	0,06128	0,04423	0,19410	0,17795

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,30400	0,19500	0,24950	0,23000	0,27750
K	0,18588	0,05894	0,12321	0,09650	0,15427
T	0,18800	0,06083	0,12442	0,10167	0,15708

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,24120	0,22960	0,22120	0,30240	0,20040
K	0,11500	0,10052	0,09018	0,18409	0,06831
T	0,11473	0,10120	0,09140	0,18613	0,06713

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,29811	0,25054	0,21703	0,19865	0,28108
K	0,18249	0,12630	0,08393	0,06544	0,16007
T	0,18113	0,12563	0,08653	0,06509	0,16126

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,30262	0,23048	0,22857	0,21262	0,20762
K	0,18519	0,10203	0,09988	0,08096	0,07619
T	0,18639	0,10222	0,10000	0,08139	0,07556

**E3.** Mantendo o número de classes fixo, variando o número total de pixels e com presença de classe informacional predominante, para validação.

Total de pixels = 1300					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,29154	0,31231	0,25692	0,23769	0,20462
K	0,15891	0,18257	0,12361	0,08929	0,06499
T	0,17346	0,19769	0,13308	0,11064	0,07205

Total de pixels = 2000					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,29800	0,29250	0,27800	0,23500	0,21250
K	0,18878	0,17313	0,12994	0,07349	0,06636
T	0,18100	0,17458	0,15767	0,10750	0,08125

Total de pixels = 2500					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,24200	0,29520	0,30760	0,22080	0,19880
K	0,09600	0,14880	0,17802	0,08744	0,05774
T	0,11567	0,17773	0,19220	0,09093	0,06527

Total de pixels = 3700					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,30676	0,24432	0,28270	0,22649	0,23000
K	0,16553	0,11167	0,14834	0,10169	0,09246
T	0,19122	0,11838	0,16315	0,09757	0,10167

Total de pixels = 4200					
	Matriz 1	Matriz 2	Matriz 3	Matriz 4	Matriz 5
G	0,30000	0,26619	0,25429	0,23190	0,21119
K	0,16289	0,14054	0,10286	0,08334	0,07295
T	0,18333	0,14389	0,13000	0,10389	0,07972