

5 – MÉTODOS

Para atingir os objectivos propostos de análise multitemporal do coberto do solo, foram utilizados os seguintes passos principais:

- 1 – Construção de uma base de dados cartográfica digital com base na cartografia em suporte de papel preexistente;
- 2 – Construção da legenda para os mapas de coberto do solo;
- 3 – Pré-Processamento das imagens de satélite;
- 4 – Caracterização espectral das classes do coberto, verificação da separabilidade espectral entre classes e cálculo de índices espectrais de vegetação;
- 5 – Classificação das imagens de satélite para obtenção de mapas de coberto do solo para os anos 1973, 1986, 1994, 2000 e 2003;
- 6 – Quantificação de alterações do coberto do solo por aplicação de álgebra de mapas num sistema de informação geográfica.

5.1 – CONSTRUÇÃO DE UMA BASE DE DADOS CARTOGRÁFICA DIGITAL

As cartas 1:50000 da JIU foram digitalizadas com um *scanner Contex Chroma Tx 40*, com uma resolução de 300x300 *dpi* e foram georreferenciadas com a atribuição das coordenadas dos cantos. Procedeu-se de seguida à construção de um mosaico da área em estudo integrando a parte correspondente de cada uma das cartas acima referidas.

Com base no mosaico procedeu-se à vectorização das curvas de nível, pontos cotados, estradas, rios e principais manchas de vegetação. Por aplicação de módulos específicos em SIG, construíram-se sequencialmente as seguintes camadas de informação:

- a carta da altimetria, que engloba os pontos cotados e curvas de nível;
- um modelo digital do terreno tipo matricial, que é uma malha rectangular e regular de pontos cotados;

Finalmente todos os mapas vectoriais (com excepção das estradas) foram rasterizados em células de 30m x 30m, de maneira a torná-los sobreponíveis para análise em formato *raster* com os dados obtidos por processamento das imagens de satélite. Assim, ficaram disponíveis as seguintes camadas de informação na base de dados geográfica:

Raster 30m x 30m – modelo digital do terreno e coberto do solo de 1953.

Vectorial – pontos cotados, curvas de nível, contorno da área de estudo, estradas, rios e manchas de vegetação.

5.2 – CONSTRUÇÃO DA LEGENDA PARA OS MAPAS DE COBERTO DO SOLO

5.2.1 - Definição de classes de coberto e das classes da legenda (LCCS)

Considerou-se que existem na área em estudo três grandes tipos de coberto: a) áreas naturais sem vegetação (massas de água, vasas, areias), b) áreas naturais e semi-naturais com vegetação (floresta densa, floresta aberta, savana arborizada, savana herbácea (lalas) e mangal) e c) áreas cultivadas com culturas arvenses (bolanhas e arroz de pam pam), ou com culturas arbóreas (pomares de caju, consociados e tradicionais). Algumas das modalidades do coberto do solo, sobretudo savana arborizada e lala, podem sofrer queimadas regularmente.

Para construção da legenda a empregar neste estudo, usou-se o *Land Cover Classification System* (LCCS) que é um sistema de classificação *a priori* do coberto do solo desenvolvido com o propósito de ser uma base uniformizada para uma cartografia da ocupação e utilização de solo numa escala global. Este método foi inicialmente desenvolvido a partir do projecto Africover (*East Africa Project*) e utiliza os critérios de classificação adoptados pelo *Food and Agriculture Organization* (FAO) e pelo *United Nations Environment Programme* (UNEP). Segundo Di Gregório e Jensen (2000), este sistema de classificação tem como objectivo harmonizar os diversos sistemas de classificação do coberto do solo adoptados pelas diferentes equipas técnicas e baseia-se nas características estruturais e funcionais da vegetação, permitindo uma definição hierárquica de classes de coberto do solo.

A definição do 1º nível hierárquico da legenda foi feita com base nas características estruturais da vegetação – percentagem de coberto, se tem um ou mais estratos de árvores, arbustos ou herbáceas, culturas, etc. Na fase seguinte definiu-se o 2º nível hierárquico com base na introdução de parâmetros específicos como a sazonalidade (folha persistente ou caduca), regime hídrico, etc. Os níveis hierárquicos foram estabelecidos para as classes tendo em conta as informações auxiliares existentes e assim na tabela 6 discriminam-se as classes de legenda obtidas com utilização do programa LCCS.

Tabela 6 – Classes de Legenda LCCS definida para o mapa de coberto de solo (LC – Land Cover; LCCLLevel – Sequência de passos na classificação dicotómico-hirárquica; LCCOwnLabel – Rótulo do utilizador e LCCLLabel – Rótulo do tipo de coberto atribuído pelo programa)

LC	LCCLLevel	LCCOwnLabel	LCCLLabel
Forest	A3A10B2C1D1 E2F2F5F7G2F1- B5E4F9G5	Floresta densa	Multi-Layered Semi-Deciduous High Forest (With Second Layer Of High Trees)
Woodland	A3A11B2C1D1 E1F2F6F7G3F1- A12B5E4F9G8	Palmar misto	Semi Evergreen ((70-60)-40%) Woodland With Open High Shrubs
Woodland	A3A11B2C1D1 E2F2F5F10G2F2 F6F10G3- A12B6E4G5G9	Floresta aberta	Semi-Deciduous ((70-60) - 40%) Woodland With High Emergents And Sparse Medium High Shrubs
Savanna Woodland	A3A11B2C1D1 E2F2F4F7G4F2F 6F10G3A13B6E4 F8G11G9	Savana arborizada	Semi-Deciduous (40(20-10%)) Woodland With Closed Medium To Tall Herbaceous Layer And Sparse Medium High Shrubs
Open Shrubs	A4A13B3C1D1 E1F2F5F10G2- A14B8C5G7	Mangal	Broadleaved Evergreen Open ((70- 60) - 40%) High Shrubs With Low Emergents On Permanently Flooded Land (With Daily Variations)
Grasslands	A6A12B4C1E5 F2F5F10G2- B11C4G6	Savana herbácea (Lala)	Mixed Closed Tall Grassland With Medium High Emergents On Permanently Flooded Land (Persistent)
Tree Crops	A1XXB5C1D1 D9-A7	Pomares de Caju	Permanently Cropped Area With Rainfed Broadleaved Tree Crop(s)
Tree Crops	A1B1B5C2-A7B4	Pomares Consociados	Continuous Medium Sized Field(s) Of Broadleaved Tree Crop(s) .
Graminoid Crops	A4B1B5C1D1 D7	Arroz de pam pam	Shifting Cultivation Of Graminoid Crop(s)
Aquatic Or Regularly Flooded Graminoid Crops	A1B1B5C1-B4	Bolanha	Continuous Medium Sized Field(s) Of Graminoid Crops On Permanently Flooded Land
Natural Waterbodies	A1B3-A4B5	Vasas	Tidal Area (Flowing) (Surface Aspect: Bare Soil)
Built Up Áreas	A4-A13	Terreno urbano	Urban Area(s)
Natural Waterbodies	A1B3-A4	Água	Tidal Area (Flowing)
Natural Waterbodies	A1B3-A4B6	Areias	Tidal Area (Flowing) (Surface Aspect: Sand)

5.2.2 - Trabalho de campo para recolha de amostras no terreno

Segundo Chuvieco (2002), para a interpretação de uma imagem é necessário um conhecimento prévio do terreno a que esta imagem pertence, a fim de permitir identificar cada uma das classes de interesse. O objectivo é seleccionar com rigor uma

série de *pixels* da imagem que represente cada uma das classes que se pretende discriminar, tendo em conta a sua variabilidade na área em estudo, a fim de obter resultados mais precisos a menor custo. A distribuição, tão homogénea quanto possível, do conjunto de dados de referência (dados de treino e validação) sobre a imagem aumenta a probabilidade deste conjunto de dados ser representativo de toda a variabilidade inerente às classes de ocupação do solo presentes.

Assim, foi realizada uma missão de trabalho de campo em Março de 2005 com o objectivo principal de amostrar um conjunto de pontos suficientemente numerosos de cada classe de coberto para a classificação supervisionada do coberto de solo e validação desta classificação. Para a selecção dos locais de amostragem foi previamente elaborada uma grelha de 3000 pontos distribuídos aleatoriamente sobre uma carta da área em estudo, numerados aleatória e sequencialmente, para garantir a representatividade estatística da amostra recolhida. Na utilização prevista para esta grelha seriam amostrados para cada classe de coberto observada os pontos com menor número de ordem até perfazer uma quantidade considerada suficiente para o fim em vista.

No entanto, esta metodologia revelou-se impossível de aplicar no terreno dada a dificuldade de acesso à maior parte dos pontos aleatórios previamente delineados. Assim, atendendo às condições no terreno e à duração limitada da missão de trabalho de campo, teve que se adoptar outra metodologia de selecção dos locais de amostragem. Optou-se por fazer a amostragem nos locais de acesso viável, nomeadamente ao longo das vias (estradas, caminhos, etc.) e cobrir toda a área possível. Nestas condições foram amostrados 251 pontos (fig. 17) em que foi colhida a seguinte informação:

- coordenadas geográficas (com um GPS *Garmin Etrex Euro*);
- tipo de coberto;
- dados sobre a estrutura da vegetação (% Cobertura Total, % Árvores, % Arbustos, % herbáceas, % Trepadoras, bem como as respectivas alturas) e
- dados sobre variáveis biofísicas (anexo A).

Foi igualmente feito um registo fotográfico dos pontos amostrados.

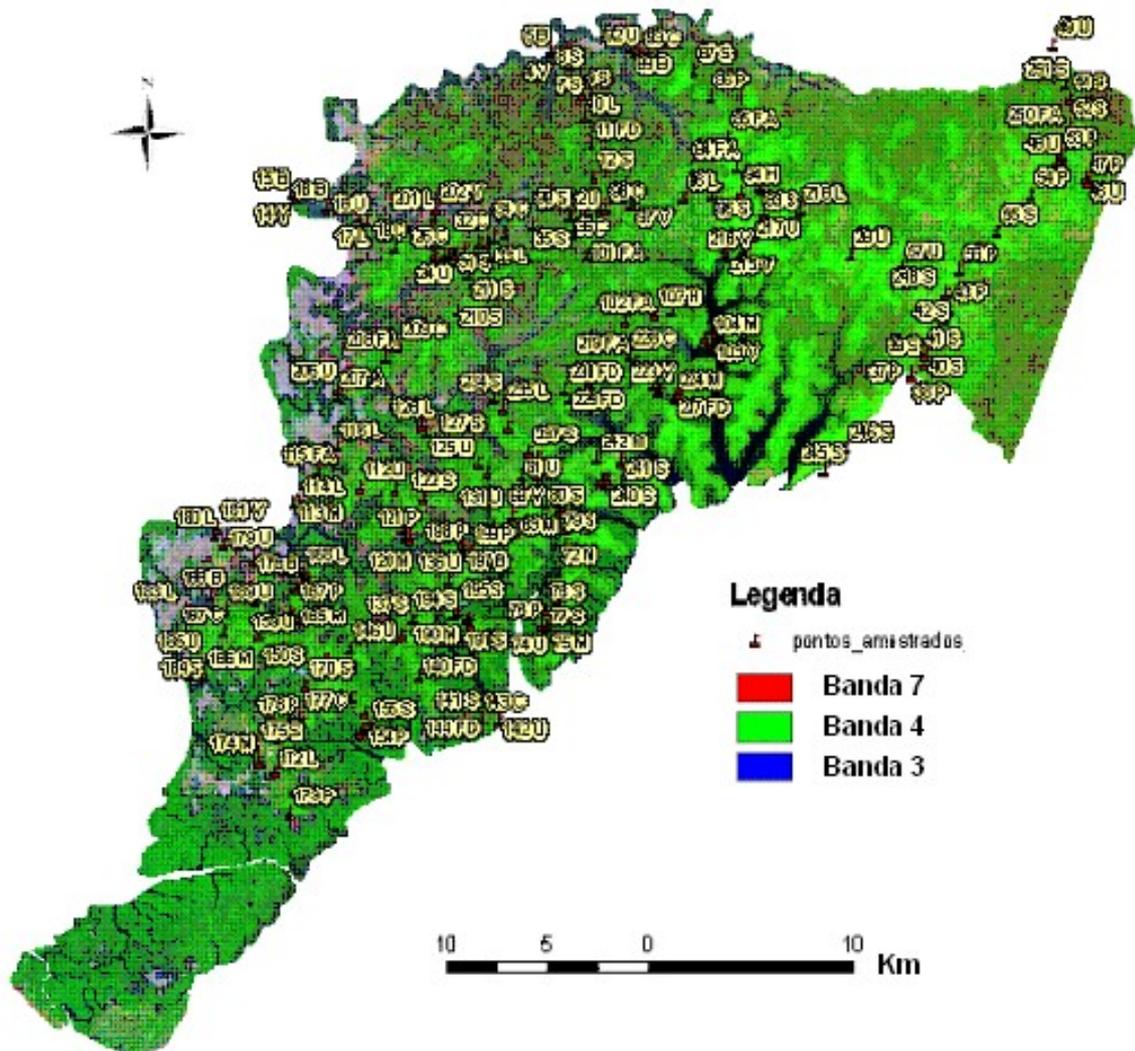


Figura 17 – Pontos amostrados em trabalho de campo e classes de coberto a que pertencem (U - urbano, V - vasas, B - bolanha, FD - floresta densa, FA - floresta aberta, C – pomar de caju, H – pomar consociado, S - savana arborizada, L - savana herbácea (lala), P - palmar, M - mangal e A – massas de água). Combinação de bandas RGB-743.

Identificaram-se no terreno 12 classes da legenda pré-estipulada – urbano, vasas, bolanha, floresta densa, floresta aberta, pomar de caju, pomar consociado, savana arborizada, savana herbácea, palmar, mangal e massas de água, como se pode constatar na figura 17.

5.2.3 - Ajuste na nomenclatura da legenda constituição de conjunto de referências

Embora no terreno todas as classes de coberto constantes na secção 5.2.2 sejam visualmente distinguíveis, quatro delas (*pomares consociados*, *de caju*, *palmar*, e *urbano*) não foram tidas em conta na classificação. Este facto pode ser explicado devido a semelhança espectral entre essas classes de coberto (constatada na verificação quantitativa através do cálculo da distância de Jeffries-Matusita - ver 5.2.5),

o que não possibilitou separar com um nível de rigor aceitável *pomares consociados*, *de caju e palmar de floresta*, algumas *culturas diversas de savana arborizada*. A classe *urbano* foi excluída, por ocupar pequenas áreas e ter assinatura espectral sobreponível a várias outras classes.

Devido às dificuldades anteriormente descritas e à necessidade de estabelecer uma legenda compatível com todas as fontes de informação utilizadas, as classes de coberto foram agregadas numa única legenda comum (tabela 7). Nas figuras 18 e 19 estão ilustradas as classes de coberto *bolanha* e *vasas*, respectivamente.



Figura 18 – Bolanha – Cafal Balanta.
P^{to} de Amostr. n^o 160



Figura 19 – Vasas – Canenque.
P^{to} de Amostr. n^o 239

Para cada classe de coberto constante da tabela 7, foram delineadas áreas de treino representativas sob a forma de polígonos definidos em torno dos pontos amostrados na área de estudo. No delineamento destes polígonos, cujo objectivo é de reunir um conjunto de *pixels* representativos que descrevam o comportamento espectral para cada classe de ocupação de solo, foram seleccionadas manchas espectralmente homogéneas e evitados *pixels* situados nas fronteiras de transição entre manchas diferentes (fig. 20). Para a constituição da amostragem referente a datas anteriores a 2003 procedeu-se a uma análise visual comparativa entre as

imagens dessas datas e as de 2003 e eliminaram-se os pontos de amostragem recolhidos no terreno que recaíram sobre manchas com alterações evidentes.

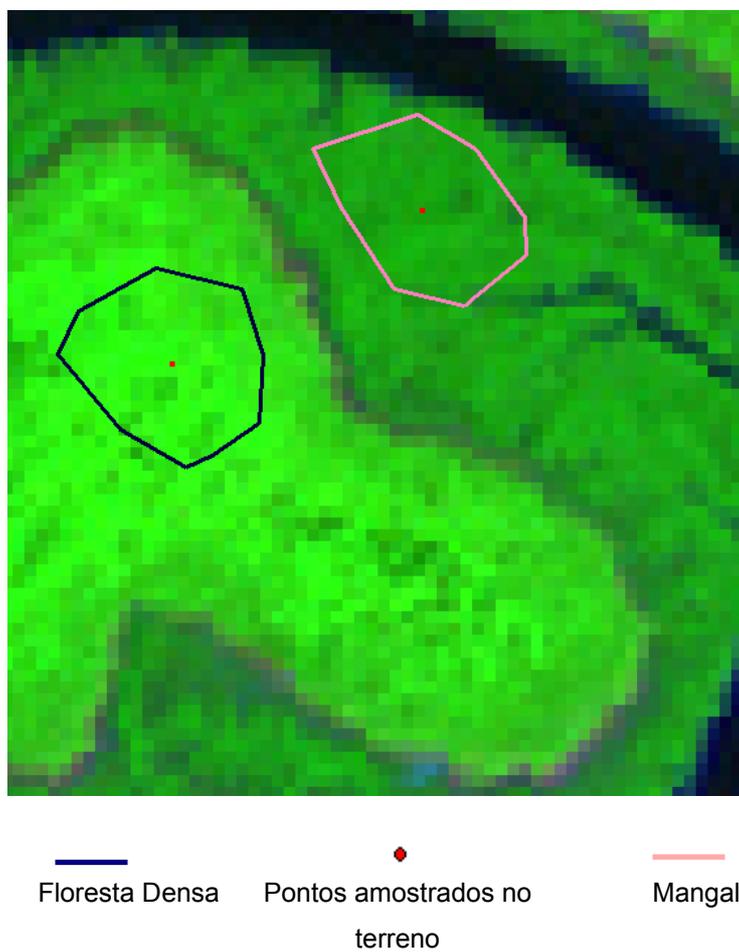


Figura 20 – Conjunto de dados de referência dentro de áreas homogêneas, sob forma de polígonos em torno de pontos amostrados no terreno

Tabela 7 – Agregação das classes numa única legenda e a respectiva correspondência do sistema LCCS e nas cartas 1:50000 - LC – Land Cover

Classes de legenda	LC	LCCS	Características Estruturais	Correspondência no terreno	Legenda nas cartas 1:50000
<i>Vegetação Lenhosa Densa (VLD)</i>	Forest	Multi-Layered Semi-Deciduous High Forest (With Second Layer Of High Trees)	Vegetação lenhosa com 90-100% árvores e 0-20% arbustos.	Floresta densa	Florestas sem vestígios de culturas actuais
	Woodland	Semi Evergreen ((70-60)-40%) Woodland With Open High Shrubs	Vegetação lenhosa densa dominada por <i>Elaeis guineensis</i> .	Palmar misto	Palmeiras de azeite em maciços consideráveis e não desbravadas
	Tree Crops	Permanently Cropped Area With Rainfed Broadleaved Tree Crop(s)	Vegetação lenhosa monoespecífica, baixa e densa.	Pomar de Caju bem desenvolvido	–
<i>Vegetação Lenhosa Esparsa (VLE)</i>	Woodland	Semi-Deciduous ((70-60) - 40%) Woodland With High Emergents And Sparse Medium High Shrubs	Vegetação lenhosa 40-80% de árvores e arbustos e 10-40% herbáceas.	Floresta aberta	(parte da classe Savanas arbustivas e florestas derrubadas para culturas) (1)
	Tree Crops	Permanently Cropped Area With Rainfed Broadleaved Tree Crop(s)	Vegetação lenhosa monoespecífica, baixa e esparsa.	Pomar de Caju em desenvolvimento	–
	Tree Crops	Continuous Medium Sized Field(s) Of Broadleaved Tree Crop(s) .	Vegetação lenhosa mista (plantada e natural).	Pomares tradicionais densos (por exemplo de cola)	Palmares desbravados com culturas diversas
<i>Vegetação Herbácea com Lenhosas Dispersas (VLHD)</i>	Savanna Woodland	Semi-Deciduous (40(20-10%) Woodland With Closed Medium To Tall Herbaceous Layer And Sparse Medium High Shrubs	Vegetação predominantemente herbácea: 60-100% de herbáceas e 10-40% de árvores e arbustos.	Savana arborizada Pousio	Savanas arbustivas e florestas derrubadas para culturas (1)
	Tree Crops	Permanently Cropped Area With Rainfed Broadleaved Tree Crop(s)	Vegetação lenhosa monoespecífica, recentemente plantada.	Pomar de Caju recentemente plantado	–
	Tree Crops	Continuous Medium Sized Field(s) Of Broadleaved Tree Crop(s)	Vegetação lenhosa mista plantada.	Pomares Consociados (plantações de fruteiras)	Culturas diversas “Pontas”, “hortas”, granjas e outras culturas

Tabela 7 – conclusão

Classes de legenda	LC	LCCS	Características Estruturais	Correspondência no terreno	Legenda nas cartas 1:50000
<i>Vegetação Herbácea (VH)</i>	Grasslands	Tall Grassland With Medium High Trees And Medium High Shrubs	Vegetação herbácea: 100% de herbáceas e 0-10% de árvores e arbustos.	Lalas	Lala de água doce Lala de água salgada
<i>Vegetação Lenhosa das Vasas Costeiras (VLVC)</i>	Open Shrubs	Broadleaved Evergreen Open ((70-60) - 40%) High Shrubs With Low Emergents On Permanently Flooded Land (With Daily Variations)	Vegetação lenhosa das vasas costeiras 20-100% de árvores e arbustos.	Mangal	Mangal
<i>Culturas arvenses ou solo a descoberto (CA/SD)</i>	Aquatic Or Regularly Flooded Graminoid Crops	Mixed Closed Tall Grassland With Medium High Emergents On Permanently Flooded Land (Persistent)	Culturas de arroz nos solos de zonas baixas alagados junto ao mangal, ou nas lalas. Na sua maior parte encontram-se nas margens do rio Cumbijã.	Bolanha	Arrozal de regadio (bolanha)
<i>Vasas (V)</i>	Natural Waterbodies	Permanently Cropped Area With Rainfed Broadleaved Tree Crop(s)	Depósitos de sedimentos finos nas zonas costeiras.	Vasas	Lodo ou areia
<i>Água (2)</i>	Natural Waterbodies	Tidal Area (Flowing)	As massas de água mais importantes na área de estudo são rios Cumbijã e Cacine e os respectivos afluentes.	Água	–
<i>Areias (2)</i>	Natural Waterbodies	Tidal Area (Flowing) (Surface Aspect: Sand)	Areias depositadas em geral nas zonas baixas por acção do vento ou da água.	Areias	Dunas de areias

(1) esta classe na legenda das cartas 1:50000 pode corresponder a mais que uma classe na legenda final utilizada neste trabalho, nomeadamente *Vegetação Lenhosa Esparsa* e *Vegetação Herbácea com Lenhosas Dispersas*, tendo-se optado pela inclusão nesta última. As consequências desta opção serão discutidas adiante.

(2) classes não consideradas na legenda final.

5.3 – PRÉ-PROCESSAMENTO DE IMAGENS DE SATÉLITE

Os dados obtidos por satélite sofrem diversos tipos de distorções no decorrer da sua aquisição, principalmente a nível geométrico e radiométrico. O pré-processamento é o processo de correcção destas distorções, e é aplicado às imagens antes da sua análise quantitativa ou simplesmente análise visual, cujo objectivo é de extrair informação sobre as imagens. Neste trabalho foram aplicadas correcções geométricas.

Os satélites registam as informações das imagens na superfície terrestre sem referência a um sistema de projecção. A correcção geométrica serve para corrigir as distorções existentes nas imagens atribuindo-lhes um determinado sistema de coordenadas. Este processo é realizado em três fases:

1) Estabelecimento de pontos de controlo – Primeiro identificaram-se as coordenadas na imagem (linha e coluna) de vários pontos distintos, designados por pontos de controlo. Estabelecem-se em locais facilmente identificáveis, como cruzamentos de estradas, edifícios, pontes, etc., na imagem deformada (de satélite), e fazem-se coincidir com as suas verdadeiras posições em coordenadas terrestres (x e y em metros).

Para este procedimento, utilizaram-se como base para a correcção geométrica as imagens ortorectificadas Landsat ETM+ de 04/12/01 e 04/04/02, com projecção *Universal Transverse Mercator* (UTM) nutm 28, isto é, folha 28 norte e *datum* WGS 84, por serem consideradas imagens de referência corrigidas. Foram utilizados como pontos de controlo sobretudo os cruzamentos das estradas e algumas pontes sobre rios, uma vez que as suas localizações não sofreram alterações ao longo do tempo. A qualidade de ajuste depende da precisão com que estes pontos são localizados e de como estão distribuídos na área de estudo.

2) Função de transformação – Calculada a partir das coordenadas da imagem, x e y originais (linhas e colunas) e das coordenadas das imagens, x e y em metros. Tendo em conta a área em estudo optou-se para uma equação polinomial quadrática.

O processo do cálculo da função de transformação envolve o cálculo do erro médio quadrático RMS (*Root Mean Squared Error*). O RMS pode ser definido como a média das diferenças entre o valor estimado e o valor observado, para cada um dos pontos de controlo no terreno, e indica a precisão da marcação dos pontos. O valor limite para um erro aceitável estabelecido é de 1 *pixel*. Ultrapassando este valor, o ponto será excluído.

3) Reamostragem (*resampling*) – serve para determinar os níveis digitais (DN) a atribuir nas novas localizações dos *pixels* da imagem de satélite corrigida. Este

processo calcula os novos valores a partir dos valores dos *pixels* da imagem original. Nesta fase da correcção geométrica utilizou-se o método do vizinho mais próximo (*nearest neighbour*), dado que é o método mais simples e não altera os valores originais, o que não acontece com os outros algoritmos – interpolação bilinear e convolução cúbica (Carvalho, 2004). A malha utilizada é de 30m x 30m para as imagens Landsat TM e ETM+. Também se utilizou a mesma malha para a imagem MSS a fim de se poder compará-la com o mapa de 1953 e com imagens TM e ETM+.

5.4 - CARACTERIZAÇÃO ESPECTRAL DAS CLASSES DE COBERTO, VERIFICAÇÃO DA SEPARABILIDADE ESPECTRAL ENTRE CLASSES E CÁLCULO DE ÍNDICES ESPECTRAIS DE VEGETAÇÃO

5.4.1 - Caracterização espectral das classes de coberto

A caracterização espectral das classes de coberto do solo foi feita para a imagem 204/052 de 2003 com base no conjunto de dados de referência, tendo por finalidade mostrar o comportamento espectral de classes de coberto consideradas.

Nos sensores TM e ETM+ os números digitais (DN) variam entre 0 e 255 e no MSS entre 0 e 63. Segundo o site <http://www.landsat.usgs.gov> e Chuvieco (2002), o processo da conversão de DN em reflectividades realiza-se em duas fases:

1) – conversão de DN em valores de radiância ($L_{sen,k}$), a partir dos coeficientes dos ganhos (*gains*) e perdas (*offset/biases*);

2) - estimacão de valores de reflectividade aparente (ρ^*_k), conhecendo a irradiância solar e a ficha de aquisicão (onde se estima o factor de correcção da distância terra-sol (D) e o ângulo zenital θ_i).

1) conversão de DN em valores de radiância ($L_{sen,k}$)

A conversão de DN em valores de radiância é feita aplicando a seguinte expressão:

$$L_{sen,k} = Gains_k \times DN_k + Offset_k \quad \text{Onde:}$$

$L_{sen,k}$ – é a radiância espectral recebida pelo sensor na banda k, expressa em ($Wm^{-2}sr^{-1}\mu m^{-1}$);

$Gains_k$ – é o ganho reescalado na banda k contido na informacão da imagem, expressa em ($Wm^{-2}sr^{-1}\mu m^{-1}$);

$Offset_k$ – é a perda reescalada na banda k (biases) contida na informacão da imagem, expressa em ($Wm^{-2}sr^{-1}\mu m^{-1}$);

DN_k – é o número digital da imagem na banda k.

Na tabela 8 apresentam-se os números digitais médios das classes de coberto consideradas em cada banda.

Tabela 8 – Números digitais médios das classes do coberto consideradas para a imagem 204/052/2003

Classes	ETM 1	ETM 2	ETM 3	ETM 4	ETM 5	ETM 7
Vasas	81.50	68.50	66.00	35.00	46.50	30.50
VHLD	68.00	52.00	50.00	47.50	93.00	56.00
VLVC	67.50	50.50	42.00	48.00	41.50	23.00
VH	67.00	53.00	49.00	59.50	78.00	39.50
VLD	62.50	48.00	35.50	66.00	56.00	28.00
VLE	67.00	52.00	51.00	58.00	86.00	55.00
CA/SD	81.50	68.50	81.50	54.50	121.50	87.00
Areias	143.00	148.50	173.00	85.00	194.00	190.00
Ardido	71.00	52.50	51.50	31.50	81.50	76.50
Água	93.00	82.50	85.00	34.00	18.00	15.50

2) – estimaco de valores de reflectividade aparente (ρ_k^*)

A reflectividade aparente é definida como a proporo de radiao incidente que é reflectida por uma superfcie bidimensional e recebida pelo sensor. Os valores de reflectividade aparente obtêm-se a partir da seguinte expresso:

$$\rho_k^* = \frac{D\pi L_{sen,k}}{E_{0,K} \cos\theta_i}$$

Onde:

$E_{0,K}$ – é a irradincia solar numa fraco da atmosfera (para a banda k do espectro), expressa em ($Wm^{-2}sr^{-1}\mu m^{-1}$);

ρ_k^* – é a reflectividade aparente do coberto na banda k;

θ_i – é o ângulo zenital solar formado entre a vertical do terreno e raios solares incidentes. Nos modelos simplificados este ângulo é o complementar do ângulo da elevao solar;

D – é o factor de correco da distncia terra-sol, e é calculado por:

$$D = \left[\frac{(1 + 0.01674(\sin(2\pi (J - 93.5))))}{365} \right]^2 \quad \text{Onde: J – indica o dia do calendrio juliano.}$$

Na Tabela 9 apresentam-se os parmetros da irradincia solar, os ganhos e as perdas para o sensor ETM+ numa fraco da atmosfera.

Tabela 9 – Parâmetros da irradiância solar para imagens Landsat ETM+ numa fracção da atmosfera, adaptado do site de Landsat, ganhos e perdas.

Bandas	$E_{0k} (Wm^{-2}sr^{-1}\mu m^{-1})$	Gains	Offsets
ETM1	1969.00	0.770	-6.97
ETM2	1840.00	0.790	-7.19
ETM3	1551.00	0.620	-5.62
ETM4	1044.00	0.960	-6.06
ETM5	225.70	0.120	-1.12
ETM7	82.07	0.043	-0.39
ETM8	1368.00	–	–

Efectuou-se uma correcção atmosférica subtraindo o DN do objecto escuro (águas profundas) em cada banda aos DN das respectivas bandas das classes consideradas e calculou-se a radiância espectral recebida pelo sensor na banda k ($L_{sen,k}$). O factor de correcção da distância terra-sol (D) tem o valor 24, correspondente ao dia juliano à data da aquisição da imagem (24/01/2003). Assumindo que se trata de uma zona plana, o ângulo zenital solar θ_i foi calculado com base no complementar de ângulo de elevação solar (46.5°) que se encontra na informação da imagem.

Assim, obteve-se um valor de 43.5° como sendo o ângulo zenital solar, e calcularam-se as reflectâncias de superfície das classes de coberto consideradas nas k bandas (ρ^*_k). De salientar que a classificação da imagem foi feita com base nos níveis digitais (DN).

5.4.2 – Verificação da Separabilidade das Classes

Para aferir a separabilidade entre as classes utilizaram-se dois métodos, um gráfico e outro numérico.

a) - *Método Gráfico (Scattergrams)*

O *Scattergrams* é um método gráfico que permite verificar a separabilidade das classes. Neste trabalho usou-se o método gráfico a 2D que projecta para cada classe, elipses contendo 95% dos pares de DN's, no espaço espectral definido por bandas emparelhadas duas a duas.

b) - *Método Numérico (Distância de Jeffries-Matusita (JM))*

Segundo Richards (1986), a distância JM (às vezes chamada também distância Bhattacharyya) entre um par de distribuições de probabilidades (classes espectrais neste caso) é definida como sendo uma medida de separabilidade espectral assumindo que estas são representadas por uma distribuição normal multidimensional. É calculada pela seguinte fórmula:

$$JM_{ij} = \sqrt{2(1 - e^{\alpha})}$$

$$\alpha = \frac{1}{8}(\mu_1 - \mu_2)^T \left(\frac{C_i + C_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{\frac{1}{2}|C_i + C_j| + \sqrt{|C_i| \times |C_j|}}{\sqrt{|C_i| \times |C_j|}} \right)$$

Onde:

i, j – assinaturas das classes a comparar;

μ_i – vector médio da assinatura i ;

μ_j – vector médio da assinatura j ;

T – função transposta;

C_i – matriz da covariância da assinatura i ;

C_j – matriz da covariância da assinatura j ;

$|C_i|$ – determinante de C_i ;

$|C_j|$ – determinante de C_j ;

\ln – função logaritmo natural.

O limite máximo da distância JM é de 2 para classes espectrais completamente separáveis e de 0 para as classes espectralmente idênticas.

5.4.3 – Cálculo de Índices Espectrais de Vegetação

A vegetação saudável tem uma reflectância forte na porção do espectro de infravermelho próximo e absorve fortemente na região do visível. O solo e a água apresentam reflectâncias idênticas nas regiões do espectro do infravermelho próximo e no vermelho. Dos vários índices de vegetação existentes, neste trabalho calcularam-se o índice de vegetação simples (VI), o índice de vegetação de diferença normalizada (NDVI) e o índice de vegetação ajustado ao solo (SAVI).

O objectivo destes índices é realçar aspectos que permitam discriminar entre diferentes tipos e estados de vegetação e minimizar efeitos indesejáveis nas radiâncias do mesmo tipo de coberto, resultantes de diferenças de iluminação devidas à topografia. Melhoram a aparência da distribuição espacial da informação das imagens tornando-a de fácil interpretação visual e computacional. Assim, mapas com índices de vegetação podem ser utilizados em conjunto com as bandas originais das imagens na construção de classificadores.

Índice de Vegetação Simples (VI)

O quociente entre a banda 4 e a banda 3 do Landsat TM e ETM+ e entre as bandas 4 e 2 do Landsat MSS, representa o índice de vegetação simples. Estas bandas são as mais usadas, por conterem mais de 90% da variação da resposta espectral da vegetação. A combinação destas bandas toma em atenção a elevada reflectância da vegetação na banda 4 (infravermelho próximo) e a baixa reflectância

da vegetação na banda 3 (vermelho) do Landsat TM e ETM+ e a elevada reflectância da vegetação na banda 4 MSS (infravermelho próximo) e a baixa reflectância da vegetação na banda 2 MSS (vermelho). Resulta em valores muito maiores que 1 para a vegetação e próximos de 1 para solo e água. Assim, conseguir-se-á discriminar a vegetação (saudável e com stress) de outros tipos de coberto do solo (Carvalho, 2004). Este índice tem também desvantagens, dado que diminuem as diferenças de albedo (razão entre a energia reflectida e a energia recebida, que varia entre 0 e 1).

Índice de Vegetação de Diferença Normalizada (NDVI)

O NDVI é o índice mais utilizado para quantificar a vegetação, discriminar os vários tipos de coberto e para detectar as alterações tanto da vegetação, como da superfície terrestre através da análise de séries temporais (Cabral, 2004). Ajuda a compensar as mudanças das condições de iluminação, do declive do terreno e exposição, entre outros factores (Carvalho, 2004). Este índice é expresso como a diferença das reflectâncias entre as bandas 4 e 3 do Landsat TM e ETM+ e as bandas 4 e 2 do Landsat MSS, normalizada pela soma destas bandas.

O intervalo de variação deste índice é de -1 a 1. Os valores mais altos (próximos de 1) encontram-se nas regiões de vegetação densa e verde (madura), devido à elevada reflectância da vegetação madura na região do infravermelho próximo e baixa reflectância na região do vermelho. Zonas com pouca ou nenhuma vegetação apresentam valores próximos ou iguais a zero. As águas e nuvens apresentam reflectância elevadas na região do visível e baixas reflectâncias na região do infravermelho próximo, apresentando assim valores negativos de NDVI.

Índice de Vegetação Ajustado ao Solo (SAVI)

A sensibilidade do NDVI ao solo visível nas zonas vegetadas e aos efeitos atmosféricos, entre outros, tem gerado um crescente interesse no desenvolvimento de novos índices, tal como índice de vegetação ajustado ao solo – SAVI (Rosendo, 2004). As variações no brilho do solo influenciam bastante o NDVI, facto indesejável, principalmente quando se pretende conhecer apenas o grau da cobertura verde. Daí a justificação da criação de um índice que minimizasse essa influência.

A relação entre as reflectâncias do infravermelho próximo e do vermelho, para os solos variando de escuros a claros, é geralmente linear, ou seja, o solo cai sempre ao longo de uma linha no plano cartesiano definido pelas reflectâncias das duas bandas (banda 4 e banda 3 do Landsat TM e ETM+ e banda 4 e 2 do Landsat MSS). À medida que qualquer destes solos é coberto com vegetação, a posição do *pixel*

afasta-se perpendicularmente da linha do solo. Segundo Huete (1988), o SAVI está entre um grupo de índices de vegetação que medem ou aproximam a distância entre o pixel e a linha do solo, assim retirando o efeito de solos claros ou escuros. Na tabela 10 apresentam-se as fórmulas referentes ao Landsat TM e ETM+ para o cálculo dos índices considerados.

Tabela 10 – Fórmulas de cálculo dos índices VI, NDVI e SAVI

Fórmulas de Índices de Vegetação	Onde:
$VI = \frac{B4}{B3}$	B3 – reflectância do vermelho;
$NDVI = \frac{B4 - B3}{B4 + B3}$	B4 – reflectância do infravermelho próximo
$SAVI = (1 + L) \frac{B4 - B3}{B4 + B3 + L}$	L – parâmetro relacionado com a reflectividade do solo, que varia entre 0 e 1.

O valor do coeficiente L escolhido deve ser tanto menor quanto mais densa for a vegetação (Lima & Nelson, 2003; Huete, 1988). Para efeitos práticos indicaram-se três factores de ajuste:

L = 1 para a vegetação de muito baixa densidade de cobertura;

L = 0.5 para a vegetação com densidade intermédia;

L = 0.25 para altas densidades de vegetação.

No caso da nossa área de estudo utilizou-se L = 0.25 por se tratar de região com vegetação de alta densidade.

5.5 - CLASSIFICAÇÃO DAS IMAGENS

A classificação de imagens pode ser feita recorrendo a três métodos: i) automática ou não supervisionada, ii) assistida ou supervisionada e iii) classificação híbrida.

Os método não supervisionado baseia-se na procura automática de grupos naturais ou “clusters” das propriedades espectrais dos *pixels*, e atribui cada *pixel* a uma classe com base nos parâmetros de agrupamento definidos inicialmente. Isto é, agrupam-se os *pixels* em classes espectrais com base apenas na similitude dos dados numéricos neles contidos. Depois de identificados estes grupos ou classes epectrais, é atribuída uma classe temática a cada um deles.

Os classificadores supervisionados baseiam-se em exemplos da resposta espectral correspondentes a cada uma das classes de ocupação do terreno pré-definidas (treino) e constantes da legenda. Os mais comuns empregam métodos

estatísticos paramétricos que assumem que as observações recolhidas durante a fase de treino para cada classe de coberto e em cada banda espectral são de natureza Gaussiana, isto é seguem a distribuição normal (Jensen, 1996). A fase do treino é entendida como a fase em que se deve reunir um conjunto de estatísticos que descrevem o comportamento espectral de cada classe de ocupação de solo definida na legenda. Todas as classificações supervisionadas constroem classificadores com base nesta amostra de treino e a qualidade desta amostra determina o sucesso da fase de classificação e o valor da informação por ela gerada.

O classificador de Máxima Verosimilhança Standard é um dos métodos de classificação supervisionada mais comuns e é o mais utilizado na classificação de dados de detecção remota devido à sua robustez e por se ajustar com mais rigor à disposição original dos dados (Chuvieco, 2002). Este autor descreveu também outros classificadores como o de mínima distância e o de paralelepípedo. A classificação híbrida engloba os dois métodos – supervisionado e não supervisionado – anteriormente descritos.

No presente trabalho foram utilizados os dois primeiros métodos. Foi realizada uma classificação não supervisionada para a imagem de 2003, com a finalidade de identificar manchas uniformes do ponto vista espectral e ajudar na pré definição das classes de coberto durante o trabalho do campo e respectiva legenda.

No que concerne ao método supervisionado, utilizou-se um algoritmo de classificação não paramétrico, a indução de árvores de decisão de Breiman *et al.* (1984). Este é um método que não requer a verificação de pressupostos relativos à distribuição dos dados, e que tem como objectivo a detecção de padrões no conjunto de dados, com complexidade variável permitindo reduzir significativamente o tempo dispendido na análise directa dos dados. Assim, foi construído um classificador para cada uma das imagens, cuja finalidade é de produzir mapas de coberto do solo.

5.5.1 - Classificação Não Supervisada

Para a classificação não supervisionada utilizou-se o algoritmo ISODATA, que consiste em vários passos. Estes passos estão sucintamente descritos em Chuvieco (1996). Utilizaram-se os seguintes parâmetros:

- número de bandas: 6 bandas (1,2,3,4,5 e 7)
- número de classes pré-estabelecidas: 30 classes;
- percentagem máxima de *pixels* cujos valores são considerados imutáveis entre duas iterações: 98% (o algoritmo pára ao atingir as 30 classes e este número);
- percentagem mínima de *pixels* numa classe: 0.01 (se o grupo gerado tiver uma percentagem inferior à definida é apagado e os *pixels* serão agrupados a um grupo alternativo);
- desvio padrão máximo: 4.5 (quando atinge um valor máximo definido para um grupo e o número de *pixels* numa classe for duas vezes maior que o número mínimo de *pixels* especificado, o grupo é dividido em dois);
- distância mínima entre os centros dos grupos: 3.2 (grupos com distância menor que este valor são agrupados);
- intervalo de amostragem tanto para linha como para a coluna: 1.

Após a interpretação visual, as 30 classes pré-estabelecidas foram agregadas em 8 classes, a partir das quais se gerou um mapa de coberto de solo. Com base neste mapa gerado e no conhecimento prévio do terreno, foi possível definir as classes a considerar para o trabalho de campo.

5.5.2 - Classificação Supervisada

No presente trabalho a classificação supervisionada foi efectuada depois de definidas as classes de legenda, e consistiu em três passos i) definição de conjunto de dados de referência e ajuste na nomenclatura da legenda (ver 5.2.3); ii) construção e aplicação do algoritmo de classificação em árvore e iii) verificação e validação dos resultados da classificação.

5.5.2.1 - Construção e aplicação do algoritmo de classificação em árvore

O modelo de árvore de classificação foi desenvolvido por Breiman *et al.* (1984) e implementado no software CART desenvolvido pela Salford Systems (Steinberg & Colla, 1997). É uma função $f(x)$ definida pelas covariáveis no espaço X em que cada vector x , $f(x)$ corresponde a K classes (Breiman *et al.*, 1984). Este algoritmo ajusta

modelos através de partições binárias recursivas, e utiliza este algoritmo recursivo para subdividir o conjunto de dados de treino até que uma partição seja composta inteiramente ou predominantemente de casos pertencentes a uma classe, ou até que algum critério de paragem seja alcançado (nós terminais) (Almentero *et al.*, 2004).

Segundo Safavin e Landgrebe (1991), as regras de divisão de nós pressupõe que cada nó descendente deve ser mais puro que o ascendente. Devido à sua simplicidade e à crescente eficiência computacional em lidar com grandes volumes de dados, este algoritmo tem vindo a ser muito usado (Sá *et al.*, 2000). Apresenta-se mais algumas vantagens que estão na base da crescente utilização da classificação em árvore:

- baseia-se em métodos estatísticos não paramétricos e como tal não requerem a verificação de pressupostos relativos à distribuição dos dados utilizados;
- não pressupõe a atribuição da igualdade de pesos às observações verificadas, o que confere uma maior robustez ao método em casos de observações *outlier*;
- informa sobre os erros que são cometidos durante a fase de classificação a nível global, em cada uma das classe da variável resposta bem como do erro em cada um dos nós da árvore;
- as informações mais importantes durante a fase de crescimento da árvore são extraídas, reduzindo assim a complexidade dos dados;
- a falta de valores observados nalgumas variáveis não impede a classificação.

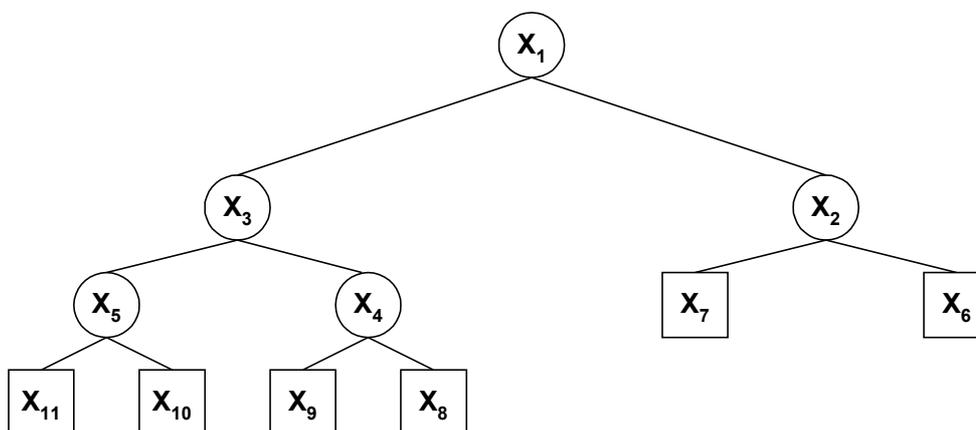


Figura 21 – Esquema ilustrativo da árvore de classificação binária. Os quadrados representam nós terminais e os círculos nós de decisão.

A figura 21 representa uma árvore de classificação binária, dado que, os dois subconjuntos resultantes (quadrados) resultam da divisão de cada conjunto de indivíduos (círculos), isto é, cada nó resulta da divisão do nó que o antecede em exactamente dois nós. Segundo (Sá, 1999), as árvores de decisão podem também ter mais do que dois subconjuntos de indivíduos, isto é, mais do que dois nós intermédios

e terminais, resultantes das divisões das observações do nó ancestral. O algoritmo da árvore de classificação utilizado no programa CART só produz árvores de decisão binárias.

Breiman *et al.* (1984) apontaram três principais etapas para a construção de árvore de classificação. A primeira etapa baseia-se na selecção das covariáveis e dos respectivos valores para a divisão de cada nó. A segunda etapa na decisão de quando considerar um nó como terminal ou continuar a divisão e a terceira na atribuição de uma classe a cada um dos nós terminais.

Ainda estes mesmos autores consideram dois tipos de modelos na definição de regras de decisão e separação de nós: 1) árvores de separação univariadas (que consideram uma única variável na definição dos limites de decisão de todos os nós) e 2) árvores de decisão multivariadas (que consideram variáveis múltiplas e/ou combinações lineares entre elas). Os pressupostos matemáticos que estão associados aos critérios de divisão de nós estão explicados em Breiman *et al.* (1984). No presente trabalho foi usado o segundo modelo na definição de regras de decisão e separação de nós (árvores de classificação multivariadas) para a construção de uma árvore de classificação para cada imagem a classificar.

Muitos critérios de divisão dos nós definem uma função de impureza à qual está associada uma medida da heterogeneidade do nó, através do cálculo da proporção da cada classe da variável resposta em cada nó. Segundo estes mesmos autores a impureza de um nó é uma função não negativa da proporção de cada uma das classes existentes nesse nó, sendo máxima quando todas as classes surgem com igual probabilidade e mínima quando o nó contém uma única classe. Assim, pode dizer-se que a qualidade de uma regra de divisão de um dado nó é dada pelo valor do decréscimo da função impureza, quando aquele é subdividido em dois nós. A regra escolhida em cada nó é aquela que maximiza a redução do valor de variância de nó parental para os dois nós descendentes (Sá, 1999). Breiman *et al.* (1984) sugerem dois critérios de divisão dos nós: o índice de *Gini* e o critério de *Twoing*, se a variável resposta apresenta mais do que dois valores discretos.

Segundo estes mesmos autores o índice de *Gini* representa um valor estimado para o erro de classificação, ou seja, corresponde a uma medida da diferença existente entre indivíduos com probabilidades diferentes, podendo deste modo ser utilizado como uma medida de impureza do nó, e então como um critério para avaliar o desempenho de uma dada regra na obtenção de nós descendentes mais puros que o ancestral. Traduz-se na procura de uma regra que maximiza a variação de uma função de impureza em cada nó.

O critério de *Twoing* separa em cada nó o conjunto de todas as classes em duas subparcelas de modo a obter o máximo decréscimo na impureza desse nó. Fornece informação acerca da semelhança entre classes, ou seja, em cada nó as classes são separadas nos dois grupos de classes que apresentam a maior dissemelhança. Este critério agrupa perto do topo da árvore um grande número de classes que são semelhantes em alguma característica, separando-as perto dos nós terminais.

Para Breiman *et al.* (1984) a escolha do melhor critério de divisão dos nós depende do problema e do tipo de informação que se pretende extrair. O índice de *Gini* tende a produzir um nó descendente mais puro do que o outro. Pelo contrário o critério *Twoing* que tende a favorecer as regras que procuram equilibrar a distribuição das observações por ambos os nós descendentes, sendo mais utilizados em problemas de classificação em que o número de classes da variável resposta é elevado. No entanto, neste trabalho aplica-se o critério de *Gini*, por ser aquele que produziu árvores de melhor qualidade.

Segundo estes mesmos autores para a avaliação da precisão do classificador existem três métodos: 1) para uma amostra pequena de dados, podem usar-se as mesmas observações utilizadas na construção do classificador para o cálculo do erro de classificação. O resultado obtido é designado por estimativa de resubstituição. Este método subestima o erro cometido pelo classificador; 2) pode ser usada uma amostra de dados independente da utilizada na construção do classificador para estimar o erro de classificação. Reduz-se neste caso o tamanho da amostra usado na construção do classificador e na estimativa do erro; 3) para uma amostra com dimensões grandes pode utilizar-se o método de validação cruzada (*V-fold Cross-Validation*) para estimar o erro de classificação. Os pressupostos matemáticos que estão associados aos critérios de avaliação do classificador estão explicados em Breiman *et al.* (1984). Estes mesmos autores afirmam que a regra que minimiza o erro de classificação é aquela que atribui a cada nó terminal uma dada classe.

A produção de uma árvore de classificação no CART baseia-se num processo no qual é gerado um conjunto de árvores de decisão aplicáveis aos dados de treino em análise. Este conjunto de árvores inclui desde a árvore mais simples de todas com poucos ramos e regras e em que os nós terminais são impuros (isto é, a precisão do classificador é baixa), até à árvore mais precisa de todas com nós terminais mais puros, mas que corresponde sempre a árvores mais complexas, com mais regras de decisão, e mais nós intermédios e terminais. A árvore de decisão óptima a aplicar em cada caso é pois seleccionada de um conjunto de árvores com precisão e complexidade crescentes.

A Árvore de Classificação Ótima

Subdividiram-se os dados brutos do conjunto de referência em dois grupos: o conjunto de treino contendo uma amostra aleatória com 75% das observações disponíveis para cada classe e o conjunto de validação com os restantes 25% das observações disponíveis para cada classe. Este último conjunto foi reservado para a fase de validação do mapa de coberto do solo gerado por aplicação do classificador à totalidade da imagem, sobre a qual foi recolhido o conjunto de referência. Para treino e construção de um conjunto de árvores de classificação no software CART foram utilizados os dados do conjunto de treino. Utilizaram-se os seguintes parâmetros para construção das árvores:

- a variável classes de coberto foi definida como variável dependente
- foram utilizadas como variáveis predictoras as bandas;
- para a separação dos nós usou-se primeiramente o critério *Twoing* e só depois o critério de *Gini*, usando combinações lineares com a finalidade de obter a separabilidade espectral dos padrões espaciais;
- probabilidades *a priori* iguais para cada classe;
- erros de custo de classificação para cada classe iguais;
- nós terminais com pelo menos 10 observações;

Inicialmente ensaiou-se a utilização dos três índices de vegetação (VI, SAVI e NDVI) como variáveis predictoras em conjunto com as bandas da imagem a classificar. No entanto, a opção de utilização de combinações lineares entre as bandas da imagem, automaticamente determinadas durante o processo de construção da árvore de classificação, revelou-se mais eficaz.

A Percentagem de Concordância (*Prediction Success*) é a percentagem da classificação correcta global e por classe sobre o conjunto de treino, isto é sobre o mesmo conjunto de valores que foram utilizados para a construção da árvore de classificação. A árvore de classificação óptima é determinada através de um compromisso entre a árvore com o menor erro de classificação, avaliado através da percentagem de concordância, e que simultaneamente não tenha um número excessivo de nós terminais. Isto consegue-se estabelecendo no algoritmo um valor para o parâmetro de custo de complexidade, que penaliza árvores muito grandes (Breiman *et al.*, 1984).

Neste estudo, para cada imagem a classificar, produziu-se uma árvore de classificação com base nos respectivos dados de treino. A árvore de classificação pode ser interpretada como uma série de regras "IF THEN", em que cada regra pode ser uma combinação linear das variáveis predictoras. Estas regras são posteriormente

implementadas sob a forma de algoritmo no programa ER Mapper 6.4 e aplicadas às imagens para gerar os mapas de coberto de solo.

5.5.2.2 - Verificação e validação dos resultados - Avaliação da Fiabilidade dos Mapas do Coberto Gerados

O objectivo deste processo é verificar a qualidade de cada classificação, permitindo a identificação e correcção dos erros cometidos ao longo do processo. Segundo Lillesand & Kieffer (1994), este processo de validação de classificação é feito comparando os resultados obtidos na classificação com a verdade do terreno, neste caso, com o conjunto de 25% dos dados de referência que não foram utilizados na construção do classificador e que constituem o conjunto de validação.

Para a validação da classificação utilizou-se o método das matrizes de confusão ou tabelas de contingência. Esta matriz é uma matriz quadrada, onde as linhas representam as classes reais (a verdade de terreno) e as colunas representam as classes obtidas na classificação. A diagonal principal da matriz representa o número de *pixels* correctamente classificados, isto é, *pixels* que apresentam concordância entre o tipo de coberto atribuído pelo classificador e o que foi verificado no terreno.

Os erros de comissão ocorrem quando na classificação se atribui um *pixel* da imagem a uma classe à qual ele não pertence (inclusão) e os de omissão, que se traduz em não atribuir a determinada classe um *pixel* que a ela pertence (exclusão) Na matriz os erros de omissão correspondem aos elementos não diagonais de cada linha e os erros de comissão aos elementos não diagonais de cada coluna.

Outro método de validação mais exigente é o do cálculo do coeficiente *Kappa*. Este teste analisa as relações múltiplas entre as distintas classes de treino. Segundo Chuvieco (2002), o coeficiente *Kappa* é um estatístico que mede a diferença entre a realidade observada e o que se pode esperar simplesmente por acaso. Delimita o grau de ajuste devido a exactidão da classificação, prescindindo do causado por factores aleatórios. Este estatístico tem valores entre 0 e 1. Para $k=1$ indica um acordo pleno entre a realidade e o mapa, e para k próximo de zero indica que o acordo observado é puramente devido ao acaso. É representado pela expressão:

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+i})}$$

N – número total de observações incluídas na matriz;
 r – número de linhas ou matriz de erro
 x_{ii} - número de observações na linha i e na coluna i ;
 x_{i+} - número total de observações na linha i ;
 x_{+i} - número total de observações na coluna i

Calculou-se também a precisão da classificação (*overall classification accuracy*) que consiste em calcular a fiabilidade global do mapa, relacionando os elementos da diagonal principal com o número total de observações incluídas na matriz, que é definido por:

$$CA = \frac{\sum_{i=1}^r X_{ii}}{N}$$

CA - precisão de classificação;

X_{ii} - número de observações na linha i e na coluna i ;

N - número total de observações incluídas na matriz.

Após a validação da classificação, para suavizar os ruídos verificados nas imagens classificadas, foi aplicado um filtro de moda (*majority.ker*) com uma matriz de dimensão 3 x 3 a cada imagem de satélite classificada. Este filtro suaviza a imagem original e atribui aos *pixels* isolados à classe mais frequente na vizinhança.

5.6 - QUANTIFICAÇÃO DE ALTERAÇÕES DO COBERTO DO SOLO

5.6.1 - Criação de Máscara

O objectivo deste procedimento é o de garantir que os *pixels* em análise sejam os mesmos para todas as datas pois a análise multitemporal das alterações do coberto do solo requer que os *pixels* em análise sejam sempre os mesmos em todas as datas. Para isso é necessário eliminar da análise situações que resultem em alterações de coberto artificiais associadas a aspectos circunstanciais de cada uma das imagens classificadas e de cada um dos mapas de coberto delas gerado.

Inicialmente criou-se uma máscara para a área de estudo e para a zonagem funcional de Cubucaré com base nos mapas de coberto do solo gerados para 1986, 1994, 2000 e 2003 a fim de criar uma série temporal de mapas com exactamente a mesma geometria e número de *pixels*, e assim estudar as transições de classes entre datas. Os *pixels* não elegíveis para análise numa determinada data (*água*, *ardido* e *areias*) são eliminados em todas as datas por reclassificação. Aos *pixels* de classe *água*, *areias* e *ardido* em qualquer data foi atribuído o valor 0 e aos restantes o valor 1 em todas as imagens. A *água* foi retirada da análise porque, devido ao já referido efeito das marés, poder-se-iam observar transições de coberto do solo artificiais entre *água* e outras classes em diferentes datas. Do mesmo modo, o *ardido* não é propriamente uma classe de coberto, mas sim um estado de coberto transitório que iria gerar uma série de transições artificiais entre este e classes de vegetação. As *areias* apenas se encontraram em duas datas e numa percentagem ínfima de *pixels* sendo portanto uma classe com representatividade negligenciável. Assim, estas

classes passaram a pertencer ao conjunto de *pixels* não classificados. Os mapas binários obtidos foram cruzados um a um para a obtenção de uma máscara única para estas datas. Esta máscara foi cruzada com os mapas multitemporais, obtendo os mapas finais só com as classes elegíveis.

Numa segunda fase criou-se uma máscara diferente para o conjunto total de mapas (1953, 1973, 1986, 1994, 2000 e 2003) cuja finalidade é criar uma série temporal de mapas de coberto de solo sobre a área em estudo e estimar as áreas totais por classe e por ano.

5.6.2 – Melhoria da classificação da imagem MSS

Tendo sido necessário aplicar ajustes à classificação obtida com base na imagem MSS; a que resulta da imagem mais grosseira do ponto de vista espacial, espectral e radiométrico foram utilizados operadores SIG com o objectivo de corrigir os erros tratáveis com base na localização (*spatial query*). Assim, dadas as confusões espectrais entre *mangal* e *floresta*, reclassificaram-se para *mangal* os *pixels* classificados como *floresta* ou como *savana arborizada* em zonas baixas e perto das linhas de água (onde é sabido haver apenas mangal), e reclassificaram-se para *floresta* ou *savana arborizada*, de acordo com a vizinhança mais próxima, as manchas de *mangal* com área inferior a 16 ha, acima da cota 5 m (onde é sabido não existir mangal).

Para o fim em vista utilizaram-se o mapa resultante da classificação da imagem MSS, o modelo digital do terreno, e operadores de reclassificação, de cálculo de distâncias, de cálculo de áreas das manchas, e de atribuição de legenda de acordo com a vizinhança mais próxima. As regras e os operadores do IDRISI utilizados foram os seguintes.

1- Se *mangal* e acima de 5 m de altitude e área da mancha inferior a 16 ha reclassifica para *floresta* ou *savana arborizada* conforme o atributo do vizinho mais próximo (operadores: *group*, *area*, *distance*, *reclass*, *allocate*)

2- Se *floresta* ou *savana arborizada* e abaixo de 5 m de altitude e a uma distância inferior a 400 m de água, reclassifica para *mangal* (operadores: *reclass*, *distance*).

5.6.3 - Criação de uma série temporal de mapas de coberto de solo sobre a área em estudo e sobre a zonagem funcional de Cubucaré e quantificação das alterações

Segundo Chuvieco (2002), a detecção de alterações a partir de técnicas de classificação pode ser abordada comparando imagens previamente classificadas, desde que tenham as mesmas classes de legenda, obtendo-se uma série temporal de cartas de coberto do solo. Neste contexto a análise multitemporal, que consiste na extracção de informações provenientes de um mesmo local obtido em momentos distintos, permite detectar as mudanças pelas variações presentes na série analisada (Langley et al., 2001). Este tipo de análise é de imprescindível importância para a monitorização de áreas desflorestadas (Moselli, 2004).

Neste trabalho, o estudo comparativo das cartas de coberto do solo em diferentes datas permite seguir a evolução e quantificar as principais alterações ocorridas na área em análise nas últimas cinco décadas. A legenda utilizada para comparação de todas as datas resulta de uma agregação da legenda utilizada para os mapas de 1986 a 2003. As classes *vegetação lenhosa densa* (VLD) e *vegetação lenhosa esparsa* (VLE) que não se puderam distinguir nas imagens de 1953 e 1973 foram fundidas numa única classe a que se chamou *vegetação lenhosa terrestre* (VLT) a fim de quantificar a evolução das estimativas de áreas totais por classe e por ano. O mesmo processo foi utilizado para as três zonas funcionais definidas por Anginot (1988) com a legenda desagregada.

A análise de transições foi efectuada exclusivamente para os mapas de 1986 até 2003, considerando as classes de legenda menos agregadas. Estes mapas servem de “input” para a criação de tabelas de contingência (*Cross-Tabulation*), dois a dois. As tabelas resultantes permitem avaliar as transições ocorridas na área em estudo entre 1986 e 1994, 1994 e 2000, e 2000 e 2003. Também se procedeu à análise de transições entre os mapas de 1953 e 2003.