

# Análise Exploratória de Dados Espectrais e Tabulares para o Município de Simão Dias (SE)

*Valdinete Vieira Nunes, Shalana Santos Carvalho<sup>2</sup>, Lauro Rodrigues Nogueira<sup>3</sup>, Márcia Helena Galina<sup>4</sup>*

## Resumo

O objetivo do presente trabalho foi a elaboração de uma análise exploratória de dados espectrais e tabulares, a partir de imagens orbitais advindas do Programa Landsat (L5, L7 e L8), da USGS, e de dados tabulares referente à produção agropecuária municipal, por meio de métodos e técnicas vinculadas às geotecnologias. A área de estudo compreendeu o município de Simão Dias (SE), em função de sua posição de destaque quanto à produção de grãos no estado e, portanto área de interesse para avaliação das modificações da cobertura da terra. Também foram utilizados dados vetoriais provindos do IBGE e da Secretaria de Estado do Meio Ambiente e dos Recursos Hídricos de Sergipe (SEMARH). No processo de classificação das cenas orbitais foi empregado o algoritmo "Iterative Self-Organizing Data Analysis Technique", com o intuito de identificar dos padrões dos dados espectrais.

**Palavras-chave:** imagens orbitais, Landsat, dados espectrais, ISODATA, uso e ocupação do solo, Simão Dias (SE).

## Introdução

O uso do Sensoriamento Remoto para análise e acompanhamento de culturas agrícolas tem sido crescente ao longo dos anos. Segundo Yi et al.(2007), para que esta ferramenta possa contribuir num sistema com objetivo de acompanhamento e de previsão das safras agrícolas, é fundamental o

<sup>1</sup> Graduanda em Engenharia Florestal, Universidade Federal de Sergipe (UFS), São Cristóvão, SE, val.ufs@gmail.com.

<sup>2</sup> Graduanda em Engenharia Ambiental, Universidade Federal de Sergipe (UFS), São Cristóvão, SE, shalanacarvalho@hotmail.com.

<sup>3</sup> Engenheiro-agrônomo, doutor em Recursos Florestais, pesquisador da Embrapa Tabuleiros Costeiros, Aracaju, SE, lauro.nogueira@embrapa.br.

<sup>4</sup> Tecnologia em Informática, Bacharel Geografia, doutora em Geografia, pesquisadora da Embrapa Tabuleiros Costeiros, Aracaju, SE, marcia.galina@embrapa.br.

conhecimento destas áreas e de suas dinâmicas. Considerando que os dados sobre a vegetação, provenientes do Sensoriamento Remoto estão fortemente relacionados com as características intrínsecas da vegetação estudada, Jensen (2009), evidencia a importância de entender a fisiologia das plantas e especialmente as suas características de pigmentação, de forma que seja possível identificar uma determinada planta mesmo que apresente uma diminuição da absorção pela clorofila devido à senescência ou por algum tipo de estresse ambiental.

Vale salientar que a vegetação verde e sadia apresenta uma assinatura espectral característica, sendo esta, influenciada por fatores como os pigmentos existentes no mesófilo paliçádico, atuantes na região do visível do espectro eletromagnético (aproximadamente 0,4 - 0,65  $\mu\text{m}$ ), o espalhamento da energia infravermelha no mesófilo esponjoso, região infravermelho próximo (aproximadamente 0,65 - 1,4  $\mu\text{m}$ ), e a quantidade de água presente na folha, região do infravermelho médio (aproximadamente 1,4 - 2,6  $\mu\text{m}$ ), (JENSEN, 2009). São estas características intrínsecas das plantas que nos permite a identificação e diferenciação entre as culturas e entre a vegetação natural.

Assim como o Sensoriamento Remoto, o SIG também vem sendo utilizado constantemente como ferramenta para a análise de uso e cobertura do solo. Na maioria dos estudos desenvolvidos em áreas de agricultura, tem-se aplicado uma combinação destas duas ferramentas das geotecnologias buscando gerar resultados mais completos sobre a área de estudo, como foi observado em: Galvêncio et al. (2007); Baboo e Devi (2010); Rozenstein e Karnieli (2011); e Mallupattu e Reddy (2013). O escopo do presente trabalho foi a elaboração de uma análise exploratória de dados espectrais, a partir de imagens orbitais advindas do Programa Landsat, escala regional, por meio de dados e técnicas vinculadas às geotecnologias. Como área de estudo tem-se o município de Simão Dias (SE), em função de seu destaque na produção de grãos e uma das áreas alvo para a análise das mudanças de uso e cobertura do solo do projeto maior constante do Portfólio de Geotecnologias da Embrapa.

## Material e Métodos

Os dados matriciais foram obtidos do acervo pertencente ao Programa Landsat (L5, L7 e L8), a partir do órgão norte-americano denominado "United States

Geological Survey" (USGS); foram baixadas as cenas 215/67 e 215/68, de acordo as especificações da Tabela 1, vale destacar que os sensores OLI possuem algumas diferenciações do TM e ETM, conforme consta, na Tabela 2. Os dados vetoriais – divisões políticas e hidrografia – e tabulares foram obtidos junto a fontes oficiais, como o IBGE e a Secretaria de Estado do Meio Ambiente e dos Recursos Hídricos de Sergipe (SEMARH).

**Tabela 1.** Dados sobre o satélite/sensor e datas das cenas utilizadas.

Satélite/Imageador	215/67	215/68
Landsat7 Sensor ETM	02/11/2002	02/11/2002
Landsat 5 Sensor TM	07/12/2006	07/12/2006
Landsat 8 Sensor OLI/TIRS	14/04/2013	10/12/2013

**Tabela 2.** Dados espectrais sobre as cenas dos diferentes sensores e resolução espacial.

Bandas Landsat 5 e 7	Correspondência das Bandas entre os Sensores L5,7 / L8	Faixa espectral L5, 7	Faixa espectral L8	Resolução Espacial (m)
<b>Blue</b>	1 / 2	0,45 - 0,52 $\mu\text{m}$	0.45 – 0.51 $\mu\text{m}$	30
<b>Green</b>	2 / 3	0,52 - 0,60 $\mu\text{m}$	0.53 – 0.59 $\mu\text{m}$	30
<b>Red</b>	3 / 4	0,63 - 0,69 $\mu\text{m}$	0.64 – 0.67 $\mu\text{m}$	30
<b>NIR</b>	4 / 5	0,76 - 0,90 $\mu\text{m}$	0.85 – 0.88 $\mu\text{m}$	30
<b>SWIR</b>	5 / 6	1,55 - 1,75 $\mu\text{m}$	1.57 – 1.65 $\mu\text{m}$	30

Os softwares de processamento das imagens, tratamento, análise e preparação dos layouts foram o ENVI 5 e o ArcGis 10. Procedeu-se com a catalogação, seleção e coleta das cenas orbitais, com atenção às questões que envolvem cobertura de nuvens, estação do ano (umidade) e ruídos. Houve uma série de dificuldades nessa seleção dada a grande cobertura de nuvens encontradas nas imagens. Quanto aos dados tabulares, eles foram acessados por meio do Sistema IBGE de Recuperação Automática e associados às feições geométricas municipais no banco de dados espaciais.

Os pares das imagens orbitais brutas foram mosaicados e recortados com base numa máscara do estado de Sergipe com um buffer de 50 Km, passaram

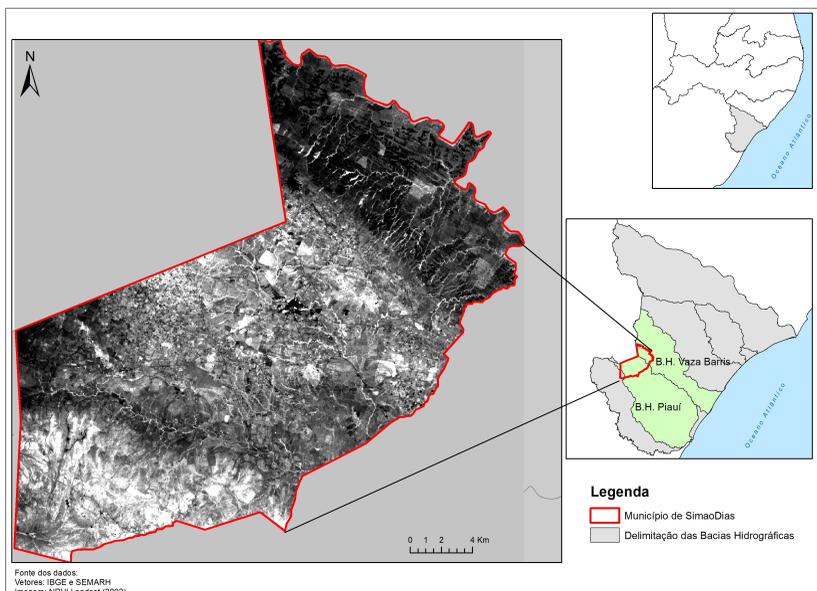
inicialmente por uma correção geométrica, no software envi 5.0, por meio da transformação polinomial de primeiro grau, utilizando interpolação por vizinho mais próximo, com pontos amostrais que foram distribuídos por toda a imagem, tendo como base as imagens do Projeto Geocover da Nasa. Esta correção permitiu a remoção das distorções sistemáticas inseridas durante o processo de aquisição das imagens. Após a correção geométrica, foram elaborados os recortes para os três conjuntos de dados, utilizando como máscara o contorno do município de Simão Dias acrescido de um buffer de 10 km, a fim de não se garantir a integridade do entorno da área de estudo. Fez-se então a correção atmosférica destes recortes com a finalidade de reduzir os efeitos de espalhamento atmosférico sobre as imagens, esta correção pode ser feita por vários métodos, entre estes o FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes), Dark Subtraction (subtração do escuro), ATmospheric REMoval (ATREM). Para o estudo em questão utilizou-se o método Dark Subtraction, que segundo Meneses e Almeida, (2012), é um método bem aceito, que consiste na análise do histograma individual de cada banda da imagem, permitindo a identificação e subtração do valor adicional de brilho de cada banda.

Prosseguiu-se com o processo de classificação, por meio do algoritmo “Iterative Self-Organizing Data Analysis Technique” (ISODATA), com o intuito de identificar dos padrões dos dados. A técnica citada congrega os pixels segundo as características espectrais, organizando-os em agrupamentos, cujos parâmetros mais comuns de definição da proximidade espectral no espaço multidimensional são a média, variância e covariância. Há o fornecimento de parâmetros como o número mínimo e máximo de classes desejadas e o número de iterações. A cada iteração, recalculam-se e reclassificam-se os pixels, considerando os novos valores médios. Assim a classificação é realizada em três etapas: ocorre a classificação iterativamente dos pixels; a partir desta classificação definem-se os critérios para cada classe; por fim a imagem é reclassificada até que os padrões de distância espectral nos dados apareçam (TOU, 1974; MENESES e ALMEIDA, 2012).

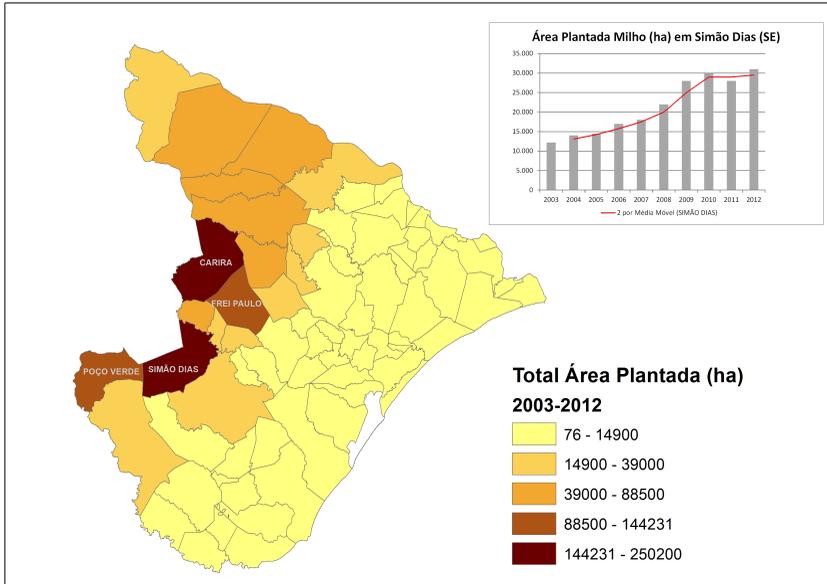
## Resultados e Discussão

O Município de Simão Dias (SE) possui a maior parte de sua extensão territorial pertencente à Bacia Hidrográfica do Rio Piauí, a oeste do estado de Sergipe, na divisa com o estado da Bahia (Figura 1). O município tem se destacado

por apresentar grande crescimento de área plantada, ao longo dos últimos 10 anos, mesmo considerando as adversidades climáticas características do Sertão Ocidental, juntamente com os municípios de Carira, Frei Paulo e Poço Verde, como mostra os dados coletados junto ao IBGE e espacializados para o Estado de Sergipe (Figura 2).



**Figura 1.** Localização do Município de Simão Dias (SE) e sua representação por meio do recorte do produto NDVI (Landsat7, 2002).

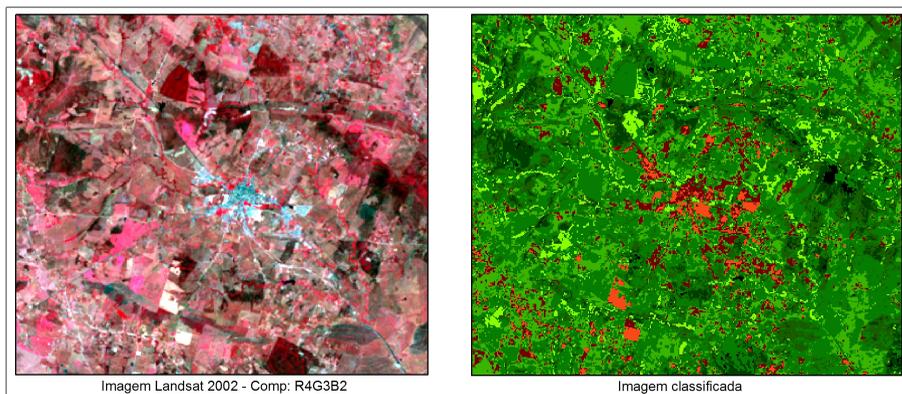


**Figura 2.** Espacialização da área total plantada para o cultivo de milho entre 2003-2012, em SE, detalhe do gráfico para Simão Dias. **Fonte:** IBGE.

Assim, a área pertencente ao município de Simão Dias foi definida para o processo de reconhecimento dos padrões espectrais. Os pares das cenas orbitais dos anos de 2002 e 2006 foram coletados nos meses de novembro e dezembro, respectivamente, portanto estação seca para o NE do Brasil. Já os pares das cenas do sensor OLI para 2013 foram coletadas nos meses de abril/13 (215/67) e dezembro/13 (215/68), portanto a diferença das estações sazonais certamente são fatores adicionais que influenciarão tanto na interação da radiação eletromagnética com os alvos terrestres quanto com a atmosfera (espalhamentos). Nesse sentido, atenção especial deve ser dada para esse último conjunto de dados, assim a análise exploratória inicial de reconhecimento de padrões passa a ser essencial para o tratamento.

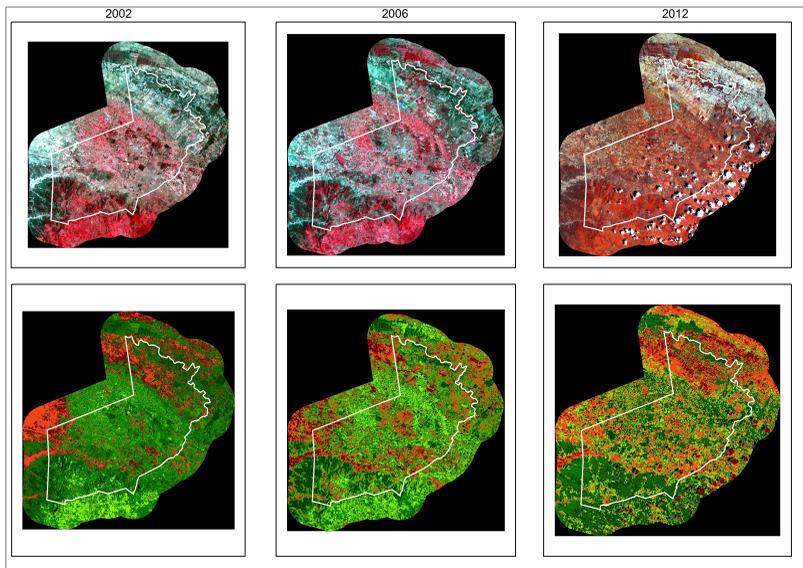
O processo de classificação “Iterative Self-Organizing Data Analysis Technique” emprega métodos de agrupamentos particionais no reconhecimento de padrões, o centro do agrupamento é randomicamente alocado e os pixels são assinalados com base na distância mínima para o centro; o desvio padrão é calculado dentro de cada agrupamento assim como a distância entre os centros

desses agrupamentos. Um agrupamento é partido se o desvio padrão é maior que aquele definido pelo usuário, assim como os agrupamentos são juntados se a distância entre eles é menor que aquele definido. Uma segunda iteração é rodada com novos centros, e outras mais até que i) a média da distância entre os centros atinge aquela definida pelo usuário; ii) a variação média da distância dos centros entre as iterações é inferior a um limiar, ou ii) o número máximo de iterações é atingido. A partir dos limiares inseridos para o primeiro conjunto de dados (2002), procurou-se conservá-los para os demais (2006 e 2013) para fins de comparação (Figura 3).



**Figura 3.** Ilustração da imagem orbital com o centro urbano e a aplicação do algoritmo.

A partir da aplicação dos procedimentos aos conjuntos de dados, houve a identificação dos principais padrões de ocupação, conforme mostra a Figura 4. Notou-se uma maior intensidade de modificações na área com menor interferência orográfica pertencente à Bacia Hidrográfica do Rio Piauí, fato que condiz com as técnicas exigidas para o cultivo de grãos. Da mesma forma, notou-se o aumento de áreas com solo exposto em detrimento de áreas com ocorrência de vegetação densa.



**Figura 4.** Processamentos de dados espectrais por meio do processo de classificação ISODATA.

## Conclusões

A análise exploratória dos conjuntos de dados espectrais na região do visível e do infravermelho a partir de métodos não-supervisionados é uma fase necessária no processo de tratamento de imagens orbitais, sobretudo nos casos que envolvam produtos de média e baixa resolução espacial, em áreas com grande heterogeneidade na cobertura do solo e sobretudo quando há o conhecimento insipiente da cobertura de vastas áreas.

Os aspectos positivos do emprego do algoritmo ISODATA dizem respeito principalmente ao fato de que ele permite um agrupamento não geograficamente polarizado para qualquer porção particular da imagem, além disso, é altamente eficiente na identificação de conjuntos espectrais inerentes.

Por outro lado, uma vez que o referido algoritmo parte do pressuposto de que o número de agrupamentos é conhecido a priori, caso haja a escolha de um valor inadequado, o método irá impor o uso de técnicas de otimização, como o agrupamento de dados. O usuário também deve especificar manualmente vários

parâmetros para controlar o processo de agrupamento, entre eles: os centróides iniciais de cada agrupamento, o número máximo de iterações, limiares para realizar a divisão, fusão ou exclusão de agrupamentos.

Por fim, estudos recentes apontam que para aplicações que apresentam agrupamentos com sobreposições e formatos arbitrários, os critérios tradicionais de validação (variância, densidade e separação) não são suficientes, uma vez que apenas um protótipo (centróide) é utilizado para representar um agrupamento, assim são mais adequados para analisar agrupamentos que apresentam formatos hipersféricos e dados compactos.

## Referências

BADOO, S.; DEVI, R. Integrations of Remote Sensing and GIS to Land Use and Land Cover Change Detection of Coimbatore District. (IJCSSE) **International Journal on Computer Science and Engineering**, v. 2, n. 9, p. 3085-3088, 2010.

EXELIS. ISODATA. Disponível em: <<http://www.exelisvis.com/docs/ISODATAClassification.html>>. Acesso em 01 jan. 2014.

GALVÍNCIO, J. D.; SÁ, I. I. S.; MOURA, M. S. B. de; RIBEIRO, J. G. Determinação das características físicas, climáticas e da paisagem da bacia hidrográfica do rio Brígida com o auxílio de técnicas de geoprocessamento e sensoriamento remoto. **Revista de Geografia**, Recife, v. 24, n. 2, p. 293- 306, mai/ago., 2007.

GALVÍNCIO, J. D.; SÁ, I. I. S.; MOURA, M. S. B. de; RIBEIRO, J. G. Determinação das características físicas, climáticas e da paisagem da bacia hidrográfica do rio Brígida com o auxílio de técnicas de geoprocessamento e sensoriamento remoto. **Revista de Geografia**, Recife, v. 24, n. 2, p. 293- 306, mai/ago., 2007.

IBGE. **Sistema IBGE de Recuperação Automática – SIDRA**. Disponível em: <<http://www.sidra.ibge.gov.br/>>. Acesso em: 03 mar. 2014.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente**. Tradução de José Carlo neves Epiphanyo (Coord.). São José dos Campos, SP: Parêntese, 2009.

MALLUPATTU, P. K.; REDDY, J. R. S. Analysis of Land Use/Land Cover Changes Using Remote Sensing Data and GIS at an Urban Area, Tirupati, India. **The Scientific World Journal**, p. 6., 2013.

MENESES, P. R.; ALMEIDA, T. de. (Org.). **Introdução ao Processamento de Imagens de Sensoriamento Remoto**. Brasília: UNB, CNPq, 2012. Disponível em: < <http://www.cnpq.br/documents/10157/56b578c4-0fd5-4b9f-b82a-e9693e4f69d8>>. Acesso em: 14 jun. 2014.

ROZENSTEIN, O.; KARNIELI, A. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. **Applied Geography**, v. 31, p. 533-544, 2011.

SERGIPE. Secretaria de Estado do Meio Ambiente e dos Recursos Hídricos- SEMARH. Atlas Digital sobre Recursos Hídricos de Sergipe. Aracaju/SE, 2012. CD-ROM.

TOU, J. T.; GONZALEZ, R. C. **Pattern Recognition Principles**. Massachusetts: Addison-Wesley Publishing Company, Reading, United States Geological Survey" (USGS), 1974. Disponível em: <<https://landsat.usgs.gov>>. Acesso em 03 mar. 2014.

Yi, J. L. R.; SHIMABUKURO, Y. E.; QUINTANILHA, J. A. Identificação e mapeamento de áreas de milho na região sul do Brasil utilizando imagens MODIS. **Engenharia Agrícola**, Jaboticabal, v. 27, n. 3, p. 753-763, set./dez. 2007.