

APLICAÇÃO DE K-NEAREST NEIGHBOR EM IMAGENS MULTISPECTRAIS PARA A ESTIMATIVA DE PARÂMETROS FLORESTAIS

Marcos Vinicius Giongo Alves¹, Ugo Chiavetta¹, Henrique Soares Koehler²,
Sebastião do Amaral Machado³, Flávio Felipe Kirchner³

¹Eng. Florestal, Dr., Università Del Molise, Pesche (IS), Itália - marcos@world-forestry.com; ugo.chiavetta@unimol.it

²Eng. Florestal, Dr., Depto. de Fitotecnia e Fitossanitarismo, UFPR, Curitiba, PR, Brasil - koehler@ufpr.br

³Eng. Florestal, Dr., Depto. de Ciências Florestais, UFPR, Curitiba, PR, Brasil - samachado@ufpr.br; kirchner@ufpr.br

Recebido para publicação: 23/07/2010 – Aceito para publicação: 03/06/2013

Resumo

A gestão dos recursos naturais requer a estimativa de uma série de parâmetros para o apoio da identificação de alternativas mais adequadas para a gestão e manejo das áreas florestais. Em particular, os ecossistemas florestais exigem um complexo e crescente conjunto de informações, e os inventários florestais fornecem informações preciosas, entretanto, espacialmente, de forma não contínua. Muitos trabalhos científicos vêm direcionando esforços para o desenvolvimento de metodologias que relacionam os dados da terra com informações de imagens multiespectrais. A modelagem dessas relações pode estender as estimativas dos dados de inventário florestal em áreas não amostradas. Neste trabalho, foi avaliado o desempenho de uma análise não paramétrica, com a utilização do algoritmo K-Nearest Neighbor em imagens SPOT5. Foram avaliados os resultados obtidos na espacialização de atributos florestais em uma área em Molise, na Itália. Entre as diversas metodologias para os cálculos das distâncias espaciais, o uso do método baseado nas distâncias multirregressivas não paramétricas apresentaram os melhores resultados. A densidade e número de espécies levantados em campo apresentaram um coeficiente de correlação de Pearson $\rho = 0,58$, comparativamente às informações obtidas nas imagens multispectrais, ligeiramente inferior aos obtidos para a área basal e volume, que obtiveram, respectivamente, $\rho = 0,62$ e $0,71$.

Palavras-chave: Inventário florestal; sensoriamento remoto; área basal; volume.

Abstract

Application of k-nearest neighbor on multispectral images to estimate forest parameters. Natural resources management requires several parameters estimate in order to support the identification of the best alternatives to forest areas management. In particular, forest ecosystems require a complex and increasing set of descriptive information, where forest inventories put up important information, however not in a continuous spatial way. Lately, several scientific researches have been focusing on establishing methodologies to relate data from field to those obtained from multispectral images. Modeling these relations can extend the estimates of forest inventory data to not sampled areas. This research evaluated performance of non-parametric analysis using the K-Nearest Neighbor (k-NN) on SPOT 5 images. It evaluated the results obtained from the spatialization of some forest attributes in a forest area located at Molise, Italy. Among several methodologies for spatial distance calculations, the use of multiregressive non-parametric distances revealed the best results. Density and number of species on the ground revealed a Pearson correlation coefficient of $\rho = 0.58$ as compared to data obtained from multispectral images, lightly lower than the obtained for basal area and volume, which were $\rho = 0.62$ and 0.71 , respectively.

Keywords: Forest inventory; remote sensing; basal area; volume.

INTRODUÇÃO

O planejamento eficaz para a gestão sustentável dos ecossistemas florestais requer uma grande quantidade de informações detalhadas, das quais, até poucos anos atrás, uma pequena parte podia ser obtida com o uso de técnicas de sensoriamento remoto. Com o desenvolvimento de novos sensores de

observação da Terra, integrados com técnicas avançadas de modelagem espacial, vem-se obtendo uma ampla gama de informações.

O desenvolvimento de técnicas para uso sustentável dos recursos florestais requer informações precisas sobre extensão, características e distribuição das áreas florestais. Os parâmetros florestais em nível de população, como diâmetros, alturas, volume e biomassa, são informações muito importantes e necessárias para o desenvolvimento de técnicas de manejo e gestão dos recursos florestais (LU *et al.*, 2004).

Fotografias aéreas têm sido muito utilizadas para mapear áreas florestais, pela delimitação de população homogênea, como a espécie predominante, estrutura e/ou idade. Alguns atributos podem ser atribuídos a cada povoamento florestal por fotointerpretação direta ou por métodos automáticos de classificação supervisionada (ZSILINSZKY, 1962; KLEINN *et al.*, 2005; THOMPSON *et al.*, 2007).

Os inventários tradicionais, baseados em levantamentos de campo, muitas vezes são de difícil realização e caros, principalmente quando realizados em grandes áreas e, em particular, quando as florestas são localizadas em áreas remotas e de difícil acesso. Nesses casos, a utilização de técnicas de sensoriamento remoto pode ser uma ferramenta muito interessante para a estimativa de variáveis florestais, a um custo relativamente baixo, com precisão aceitável e com tempo relativamente reduzido.

As técnicas de sensoriamento remoto vêm sendo utilizadas desde a década de 70, para fornecer informações contínuas dos padrões espaciais dos atributos florestais (CARNEGGIE; LAUER, 1966; CHUVIECO; CONGALTON, 1989; KUSHWAHA, 1990; FRANKLIN, 2001).

Na verdade, o sensoriamento remoto baseado nos sensores passivos não realiza medidas diretas dos atributos florestais, medindo, sim, a reflexão da radiação eletromagnética em diferentes comprimentos de onda. Com o auxílio de fontes de informação secundárias, como informações georreferenciadas (amostras) coletadas em campo, podem ser elaborados modelos estatísticos capazes de correlacionar o valor da refletância com os atributos de interesse a serem estimados nas áreas florestais.

Recentemente, muitos pesquisadores têm colocado grandes esforços na análise da relação entre as respostas espectrais e os parâmetros florestais pelo uso de métodos paramétricos e não paramétricos utilizando imagens de satélite (LANGFORD, 1976; HORLER; AHERN, 1986; DYMOND, 1988; SPANNER *et al.*, 1990; CURRAN *et al.*, 1992; TURNER *et al.*, 1999; FRANCO-LOPEZ *et al.*, 2001; MCROBERTS *et al.*, 2007; MAGNUSSEN *et al.*, 2009).

As técnicas paramétricas baseiam-se em estabelecer uma relação de regressão entre os valores dos atributos de interesse (variáveis dependentes), medidos em campo em correspondência das unidades amostrais e os seus respectivos valores de variável(is) independente(s), correspondente à reflectância detectada em bandas espectrais individuais, ou de um valor de um índice derivado da combinação de bandas. No caso de uma única variável independente, resultará em uma regressão simples; no caso de múltiplas variáveis independentes, irá resultar em uma regressão múltipla. Dessa forma, a equação de regressão elaborada com esse procedimento é utilizada para estimar o valor da variável dependente para cada pixel da região em estudo (CHIRICI *et al.*, 2008).

As técnicas não paramétricas são baseadas em abordagens muito diferentes entre si, como, por exemplo, as redes neurais (INGRAM *et al.*, 2005) e *k*-NN (*k*-Nearest Neighbor) (MCROBERTS, 2008), entretanto elas compartilham uma característica em comum, pelo fato de poderem ser utilizadas em casos onde as relações entre os dados obtidos com os sensores passivos (valores da reflexão da radiação eletromagnética) e os atributos levantados em campo são estatisticamente fracos.

Apesar dos muitos estudos das relações entre dados de sensoriamento remoto e os parâmetros das áreas florestais, em nível de população, eles variam de acordo com as características específicas de cada área objeto de estudo. A influência dessas características sobre as relações das informações espectrais e os parâmetros florestais é normalmente de difícil compreensão (LU *et al.*, 2004). Por essa razão, é um processo muito difícil a identificação do comprimento de onda ou de uma combinação de bandas para descrever a variação de um ou mais parâmetros em uma área florestal.

O objetivo deste trabalho foi a avaliação do uso do algoritmo *k*-Nearest Neighbor, com o uso de dados obtidos por amostragem em campo e imagens multiespectrais de satélite, para a estimativa de variáveis florestais.

MATERIAIS E MÉTODOS

Área de estudo

A área de estudo abrange 42.032 hectares, a qual inclui 12 municípios do extremo norte da província de Isernia (Itália) e engloba aproximadamente 9,5% da área total da região de Molise. A região objeto de estudo encontra-se em sua maior parte constituída de um relevo montanhoso e está localizada na parte mais

interna do Apenino Centro-Meridional, que faz fronteira com as províncias de Chieti e L'Aquila. Na parte norte, o limite da área de estudo faz fronteira com a região de Abruzzo, a noroeste tem como limite o rio Sangro e ao lado leste o curso d'água Sente, um pequeno afluente do rio Trigno (Figura 1).

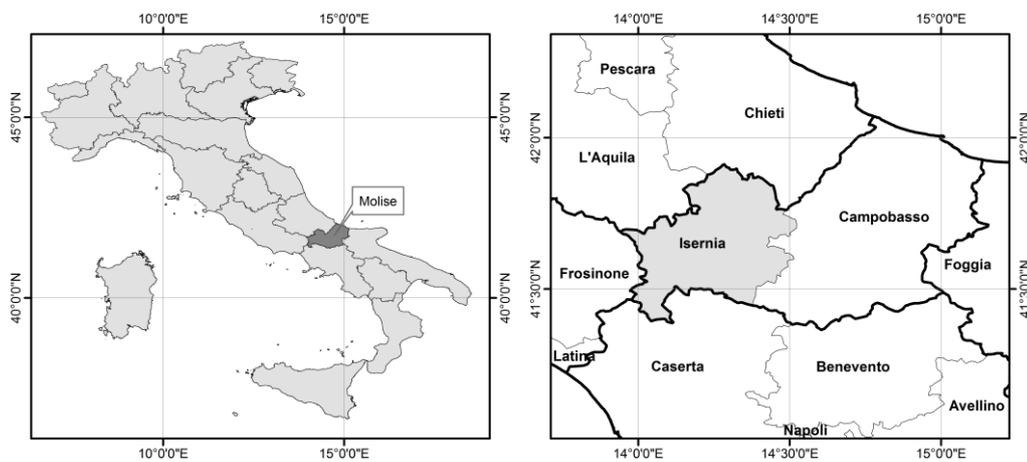


Figura 1. Localização da área de estudo.

Figure 1. Location of the focused area.

Inventário florestal

Na área em estudo, foi realizado um inventário florestal de forma sistemática, com parcelas circulares com um raio de 30 metros, onde foi estabelecido um diâmetro mínimo de amostragem de 3 cm. Internamente na parcela, foi delimitada uma subparcela, com raio de 15 metros, com o mesmo centro da parcela principal, onde também foram medidas as rebrotas, que são muito comuns nessas áreas florestais, resultado das técnicas silviculturais adotadas.

O inventário florestal resultou num total de 84 áreas amostrais, as quais foram inseridas em um banco de dados e sequencialmente processadas com o uso de técnicas convencionais, com as quais inicialmente se obtiveram as informações individuais das árvores, posteriormente agrupadas em nível de amostra e finalmente extrapoladas para as áreas avaliadas. Na tabela 1 observam-se os principais resultados obtidos no inventário florestal.

Tabela 1. Principais resultados obtidos do inventário florestal.

Table 1. Main results from the forest inventory.

| Atributo | Unidade | Mínimo | Máximo | Média | Desvio padrão |
|----------------------|----------------------------------|--------|---------|----------|---------------|
| Densidade de árvores | Nº de árvores | 14 | 5.558 | 1.786,40 | 1.234,27 |
| Número de espécies | Nº de espécies | 1 | 12 | 4,21 | 2,49 |
| Área basal | m ² .ha ⁻¹ | 0,37 | 165,67 | 23,92 | 21,07 |
| Volume | m ³ .ha ⁻¹ | 0,71 | 1350,15 | 191,70 | 191,12 |

Imagens de satélite

Para a elaboração do presente trabalho, utilizou-se uma cena proveniente do sensor SPOT 5 HGRs de 21 de setembro de 2007. Após o georreferenciamento da cena, foram realizados alguns pré-tratamentos na imagem, conforme descrito por Katila e Tomppo (2002). Sequencialmente, realizou-se uma extração das informações referentes aos pixels das diferentes bandas espectrais (variáveis independentes), as quais foram correspondentes aos centros das parcelas georreferenciadas do inventário florestal.

Para a extração dessas informações, utilizaram-se diversas janelas móveis (*kernels*), sendo extraídas as médias e os desvios padrões dos valores presentes. As larguras das janelas utilizadas foram de 1x1, 3x3, 5x5 e 7x7 pixels, que correspondem, respectivamente, a uma largura das janelas de 30, 90, 150 e 210 m.

Os cálculos foram realizados para as 4 bandas da imagem SPOT previamente ortorretificadas. Foi também calculado o Índice de Vegetação por Diferença Normalizada (NDVI) (ROUSE *et al.*, 1974), utilizando-se a equação a seguir, tendo sido também extraídos os valores médios e os desvios padrões.

$$NDVI = \frac{\rho_{IVP} - \rho_V}{\rho_{IVP} + \rho_V}$$

Em que: ρ_{IVP} = refletância do infravermelho próximo;
 ρ_V = refletância do vermelho.

O estudo teve como objeto a avaliação do método não paramétrico *k*-NN, que vem sendo aplicado no campo florestal, utilizando imagens multiespectrais com o apoio de valores pontuais coletados em levantamentos de campo. O uso dessa metodologia não só dá a possibilidade de uma estimativa da quantificação, como também sua espacialização (MUIONEN; TOKOLA, 1990; FAZAKAS; NILSSON, 1996; HOLMSTRÖM, 2002; MCROBERTS *et al.*, 2002; MCINERNEY *et al.*, 2005; GU *et al.*, 2006; FUCHS *et al.*, 2009; TOMPPPO *et al.*, 2009).

O método *k*-NN permite obter os valores dos atributos considerados de interesse para todos os outros pixels da matriz georreferenciada, com base nos dados coletados em amostragem em campo. O método é baseado na hipótese de que, na área em análise, exista uma cobertura de uma imagem multiespectral com resolução geométrica congruente à matriz geográfica utilizada baseada nos dados coletados em campo.

Outro aspecto importante é a existência de áreas florestais com características semelhantes e que as respostas radiométricas dessas áreas devem apresentar dependência dos atributos a serem estimados.

A distância medida no espaço multidimensional definido pelas bandas espectrais do sensor utilizado, $d(p_i, p)$, é calculada entre o pixel da classificação e cada outro pixel p_i para o qual tenha sido levantado em campo seu atributo.

Para cada pixel p da classificação, é selecionado o k pixel espectralmente mais próximo, com base na sua distância espectral $d(p_i, p)$, ..., $d(p_k, p)$ ($d(p_1, p)$... $d(p_k, p)$). A estimativa do valor da variável para cada pixel p é então obtida como uma função dos valores medidos de k pixel e p_i mais próximo, baseado no peso de sua distância espectral em relação ao pixel p .

Dessa forma, a fórmula para o cálculo da distância ponderada é a seguinte:

$$w_{(p_i, p)} = \frac{1}{\sum_{i=1}^k \frac{1}{d_{(p_i, p)}^t}}$$

Em que: $t = 0, 1$ ou 2 , respectivamente, se os pesos são calculados independentemente da distância, em modo inversamente proporcional à distância ou ao quadrado da distância (HOLMSTRÖM, 2001);

$d(p_i, p)$ = distância espectral (obtida utilizando-se diferentes fórmulas) entre o pixel desconhecido p e aquele correspondente a um valor em terra em que foi medido o verdadeiro valor do atributo pesquisado p_i ;

$w(p_i, p)$ = peso atribuído ao pixel incógnito p que corresponde a uma unidade levantada em campo onde foi medido o valor do atributo pesquisado p_i .

A estimativa do atributo m para o pixel é calculada por:

$$m_p = \frac{\sum_{i=1}^k w_{(p_i, p)} \cdot m_{(p_i)}}{\sum_{i=1}^k w_{(p_i, p)}}$$

Em que: $m(p_i)$ (com $i = 1, \dots, k$) é o valor do atributo m na unidade amostral correspondente ao pixel p_i , que é o i -ésimo pixel (entre aqueles correspondentes à unidade amostral onde foi medido o verdadeiro valor em campo) mais próximo, no espaço espectral, ao pixel desconhecido p .

Para a elaboração dessa metodologia, foi utilizado um total de 10 variáveis independentes, obtidas nas imagens multiespectrais, para cada modelo (5 correspondentes às médias do NDVI e das bandas 1, 2, 3 e 4, e outras 5 correspondentes ao desvio padrão do NDVI e às bandas 1, 2, 3 e 4).

Com base em testes por um processo de “*leave-one-out*” sobre o conjunto de dados disponíveis para a área de experimentação, para k foram atribuídos valores de 5 a 9 (dependendo do modelo), os quais encontram-se adequados com base em outros trabalhos citados na literatura (KATILA; TOMPPPO, 2002). Ainda segundo Katila e Tomppo (2002) os valores de k altos geram uma excessiva homogeneidade na estimativa dos atributos a serem espacializados, enquanto os valores baixos do k conduzem a um menor desempenho.

O método utilizado é baseado em uma estratificação em função da carta de cobertura do solo (a busca do k pixel mais próximo é limitada às áreas correspondentes às amostras na região classificada como de cobertura florestal). Os diferentes modelos foram obtidos variando as dimensões das janelas móveis de extração das informações espectrais, bem como as diferentes distâncias calculadas. Para este último parâmetro, foram utilizadas as distâncias mahalanobis, euclidiana, multirregressiva convencional e multirregressiva não paramétrica.

Dessa forma, considerando os quatro tipos de janelas móveis e os cinco métodos de cálculo da distância espectral, foram obtidos 20 modelos para cada variável dependente considerada, com o desempenho dos modelos sendo comparado com o uso do coeficiente de correlação de *Pearson*, sendo sequencialmente selecionados os melhores modelos para a espacialização dos dados.

Para a avaliação dos resultados, foram utilizadas as informações do inventário florestal realizados durante a atualização do Plano de Manejo Florestal de Alto Molise (PMFAM). Dessa forma, as unidades amostrais do PMFAM foram comparadas com o resultado obtido com o uso do algoritmo k -NN, para as variáveis volume e área basal.

RESULTADOS E DISCUSSÃO

Os resultados obtidos com a metodologia não paramétrica k -NN mostram diferentes padrões de desempenho, que estão relacionados às variáveis dependentes consideradas (densidade, número de espécies, área basal e volume), métodos de extração dos valores espectrais (diferentes janelas móveis) e à forma de cálculo da distância espectral das variáveis independentes (4 bandas + NVDI).

A metodologia utilizada resultou na elaboração total de 80 modelos. Na tabela 2, observa-se que para cada variável dependente foram utilizadas cinco formas no cálculo da distância espectral, com quatro tamanhos da janela de extração dos dados. Selecionando os melhores resultados (coeficiente de correlação de *Pearson*) para cada tipo de distância utilizada das diferentes variáveis dependentes, observa-se que a dimensão ideal da janela encontra-se nos valores médios, e os melhores resultados obtidos foram 20% para as janelas de 150 m, 35% para as de 30 m e 45% para as de 90 m, não ocorrendo nenhuma seleção com o uso de janela de tamanho de 210 m (figura 2).

Os valores da constante k , resultantes da otimização pelo processo de “*leave-one-out*”, apresentaram um intervalo de variação de 5 a 9, o que permitiu o uso dos modelos para a espacialização das estimativas. Esses mesmos valores também se encontram entre as indicações encontradas em outros trabalhos descritos na literatura com o uso dessa metodologia (KATILA; TOMPPPO, 2002), tendo em vista que valores altos de k geram uma excessiva homogeneidade na estimativa dos atributos a serem espacializados, enquanto os valores de k baixos conduzem a um desempenho reduzido do sistema.

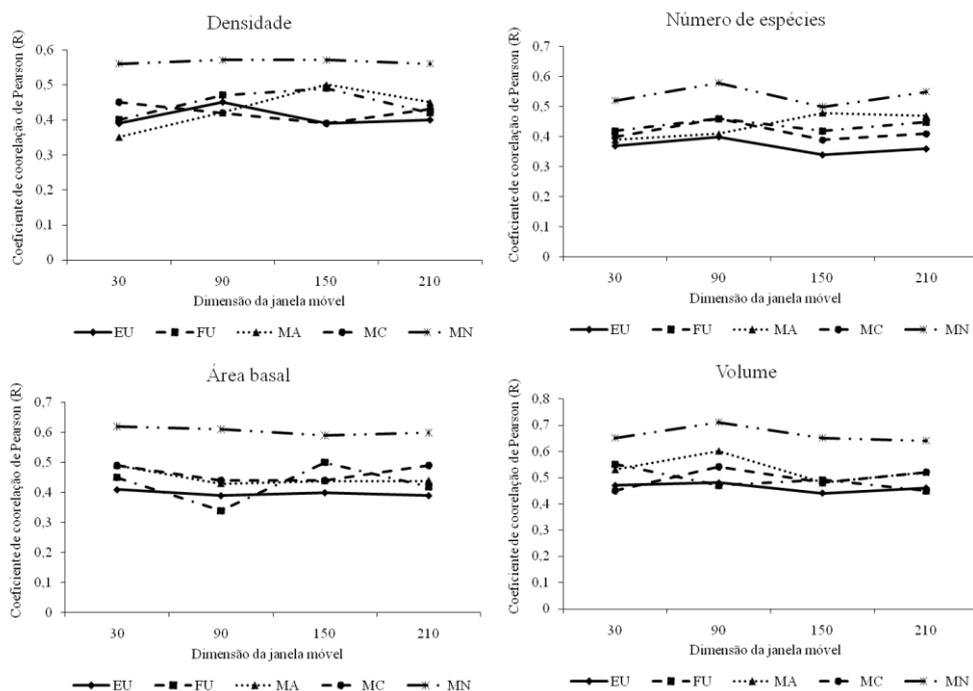
Aproximadamente 9% dos modelos não paramétricos elaborados mostraram um desempenho adequado para utilização no processo de correlação dos atributos espaciais e dados de levantamentos de campo (coeficiente de correlação de *Pearson*) $\rho \geq 0,6$, e cerca de 28% apresentaram $\rho \geq 0,5$.

Tabela 2. Resultados obtidos com a aplicação do algoritmo k-NN.

Table 2. Results obtained using the k-NN algorithm.

| Variável | Tamanho da janela móvel (m) | Distância utilizada | k | Coefficiente de correlação de Pearson (R) |
|----------------------|----------------------------------|-------------------------------|------|---|
| Densidade de árvores | 30 | euclidiana | 6 | 0,39 |
| | 90 | | 6 | 0,45 |
| | 150 | | 6 | 0,39 |
| | 210 | | 8 | 0,40 |
| | 30 | fuzzy | 8 | 0,40 |
| | 90 | | 7 | 0,47 |
| | 150 | | 5 | 0,49 |
| | 210 | | 7 | 0,42 |
| | 30 | mahalanobis | 6 | 0,35 |
| | 90 | | 7 | 0,42 |
| | 150 | | 8 | 0,50 |
| | 210 | | 5 | 0,45 |
| | 30 | multirregressiva convencional | 7 | 0,45 |
| | 90 | | 7 | 0,42 |
| | 150 | | 6 | 0,39 |
| | 210 | | 7 | 0,43 |
| 30 | multirregressiva não paramétrica | 6 | 0,56 | |
| 90 | | 7 | 0,57 | |
| 150 | | 6 | 0,57 | |
| 210 | | 8 | 0,56 | |
| Número de espécies | 30 | euclidiana | 6 | 0,37 |
| | 90 | | 6 | 0,40 |
| | 150 | | 6 | 0,34 |
| | 210 | | 6 | 0,36 |
| | 30 | fuzzy | 7 | 0,42 |
| | 90 | | 6 | 0,46 |
| | 150 | | 7 | 0,42 |
| | 210 | | 7 | 0,45 |
| | 30 | mahalanobis | 7 | 0,39 |
| | 90 | | 6 | 0,41 |
| | 150 | | 5 | 0,48 |
| | 210 | | 7 | 0,47 |
| | 30 | multirregressiva convencional | 7 | 0,40 |
| | 90 | | 7 | 0,46 |
| | 150 | | 7 | 0,39 |
| | 210 | | 7 | 0,41 |
| 30 | multirregressiva não paramétrica | 6 | 0,52 | |
| 90 | | 8 | 0,58 | |
| 150 | | 8 | 0,50 | |
| 210 | | 6 | 0,55 | |
| Área basal | 30 | Euclidiana | 7 | 0,41 |
| | 90 | | 7 | 0,39 |
| | 150 | | 8 | 0,40 |
| | 210 | | 7 | 0,39 |
| | 30 | Fuzzy | 5 | 0,45 |
| | 90 | | 5 | 0,34 |
| | 150 | | 8 | 0,50 |
| | 210 | | 7 | 0,42 |
| | 30 | mahalanobis | 6 | 0,49 |
| | 90 | | 7 | 0,43 |
| | 150 | | 7 | 0,44 |
| | 210 | | 8 | 0,44 |
| | 30 | multirregressiva convencional | 8 | 0,49 |
| | 90 | | 8 | 0,44 |
| | 150 | | 5 | 0,44 |
| | 210 | | 7 | 0,49 |
| 30 | multiregressiva não paramétrica | 7 | 0,62 | |
| 90 | | 7 | 0,61 | |
| 150 | | 7 | 0,59 | |
| 210 | | 8 | 0,60 | |

| | | | | |
|--------|----------------------------------|-------------|------|------|
| Volume | 30 | euclidiana | 7 | 0,47 |
| | 90 | | 6 | 0,48 |
| | 150 | | 8 | 0,44 |
| | 210 | | 6 | 0,46 |
| | 30 | Fuzzy | 7 | 0,55 |
| | 90 | | 7 | 0,47 |
| | 150 | | 8 | 0,49 |
| | 210 | | 6 | 0,45 |
| | 30 | mahalanobis | 7 | 0,53 |
| | 90 | | 8 | 0,60 |
| | 150 | | 7 | 0,48 |
| | 210 | | 6 | 0,52 |
| 30 | multirregressiva convencional | 7 | 0,45 | |
| 90 | | 6 | 0,54 | |
| 150 | | 7 | 0,48 | |
| 210 | | 8 | 0,52 | |
| 30 | multirregressiva não paramétrica | 9 | 0,65 | |
| 90 | | 6 | 0,71 | |
| 150 | | 7 | 0,65 | |
| 210 | | 5 | 0,64 | |



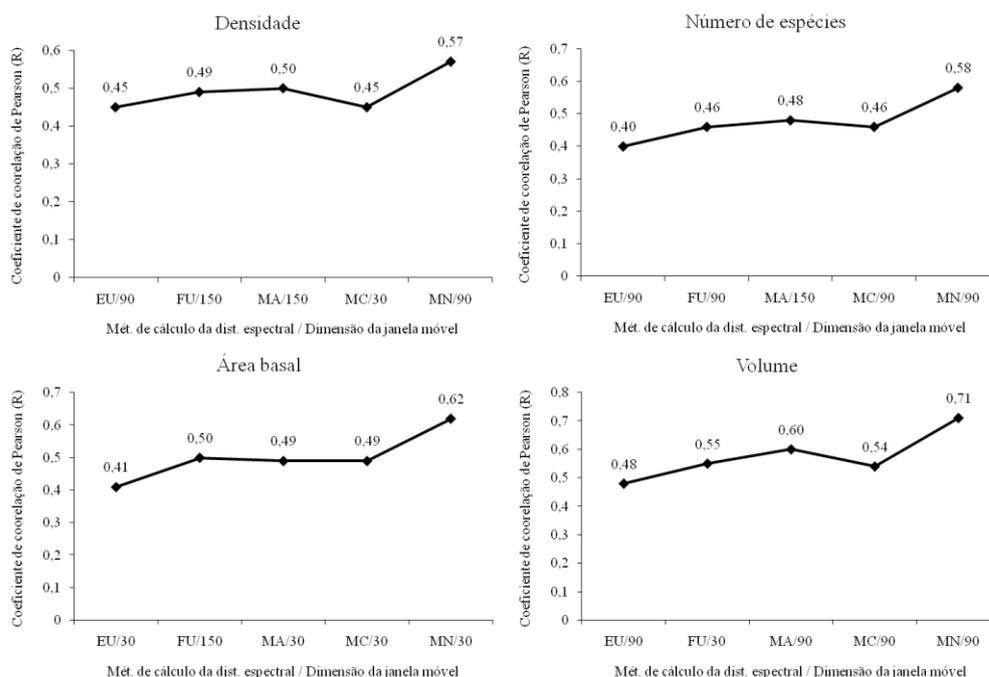
Legenda: EU: Euclidiana; FU: Fuzzy; MA: Mahalanobis; MC: Multirregressiva Convencional; MN: Multirregressiva não Paramétrica.

Figura 2. Desempenho dos modelos não paramétricos usando diferentes dimensões das janelas móveis.
 Figure 2. Performance of nonparametric models using different kernel sizes.

Para todas as variáveis dependentes, os menores valores para o coeficiente de correlação de Person foram obtidos com o uso da distância euclidiana, e os maiores valores obtidos foram com o uso dos cálculos das distâncias multirregressivas não paramétricas (MN), conforme se pode observar na figura 3. A densidade e número de espécies, utilizando a distância multirregressiva não paramétrica, apresentaram ambos um ρ de 0,58, ligeiramente inferior aos obtidos com a área basal e o volume, que obtiveram, respectivamente, ρ de 0,62 e 0,71.

Após as análises, foram selecionados os modelos que apresentaram os melhores resultados para cada uma das variáveis dependentes, dentro das diversas variações utilizadas nas elaborações. Como já

citado anteriormente, os melhores resultados para todas as variáveis dependentes foram obtidos com as análises que utilizam o cálculo da distância multirregressiva não paramétrica. Para a seleção, houve alguma variação apenas em relação ao tamanho das janelas móveis. Para a área basal, a janela com dimensões de 30 m apresentou melhores resultados; já para as demais variáveis (densidade, número de espécies e volume), foram selecionados os modelos baseados no uso de janelas com 90 m.



Legenda: EU = Euclidiana; FU = Fuzzy; MA = Mahalanobis; MC = Multirregressiva Convencional; MN = Multirregressiva não paramétrica.

Figura 3. Desempenho dos modelos não paramétricos nos diferentes cálculos para a distância espectral.

Figure 3. Performance of nonparametric models using different methods for calculation of the spectral distances.

Com base nos modelos desenvolvidos e selecionados, foi possível a espacialização dos dados das variáveis avaliadas, os quais são demonstrados na figura 4.

Em função do reduzido conjunto de dados (amostras) utilizado para a preparação dos modelos, não foi possível definir a exatidão dos dados espacializados, utilizando o mesmo levantamento empregado para a elaboração. No entanto, mesmo sem uma rigorosa validação, as informações obtidas a partir dos dados espaciais apresentam uma relação aos dados do Plano de Manejo Florestal de Alto Molise (PMFAM).

A comparação dos valores das estimativas da área basal obtidas com o algoritmo de k -NN e as informações provenientes do PMFAM indicam a existência de uma relação entre os valores, apresentando um coeficiente de correlação (r^2) de 0,68 e um erro padrão da estimativa ($RMSE$) de 18,1% (Figura 5).

Para os resultados referentes aos volumes, utilizou-se o mesmo procedimento de validação. A comparação dos valores das estimativas de volumes obtidos com o algoritmo k -NN e as informações provenientes do PMFAM indicam a existência de uma relação entre os valores, apresentando um coeficiente de correlação $r^2 = 0,63$ e um erro padrão da estimativa ($RMSE$) de 27,81% (Figura 6). Nesse caso, ocorre a mesma situação dos dados de área basal: uma correlação não muito alta, que pode ser explicada principalmente em função da diferença temporal existente entre a imagem e os dados do PMFAM, podendo também ser explicada pelas atividades de exploração e gestão dessas áreas florestais.

A metodologia, portanto, demonstra um bom potencial do desenvolvimento de modelos para auxílio nas atividades de inventário florestal por meio de imagens multiespectrais.

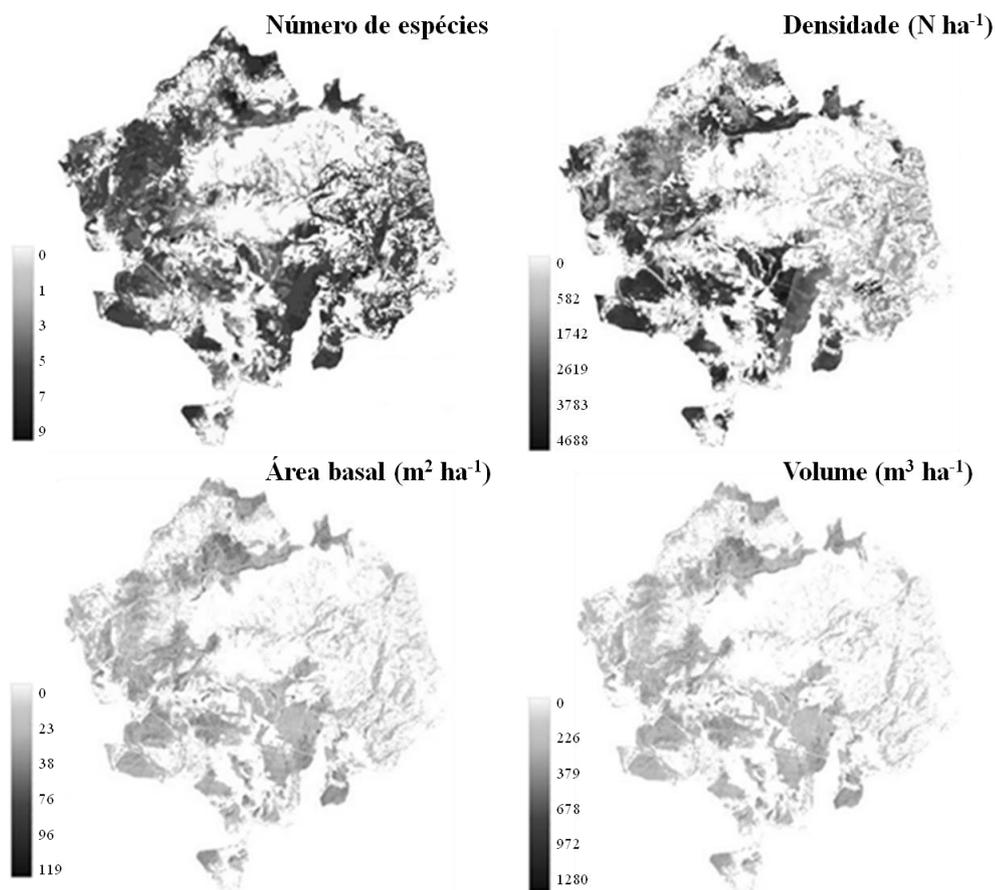
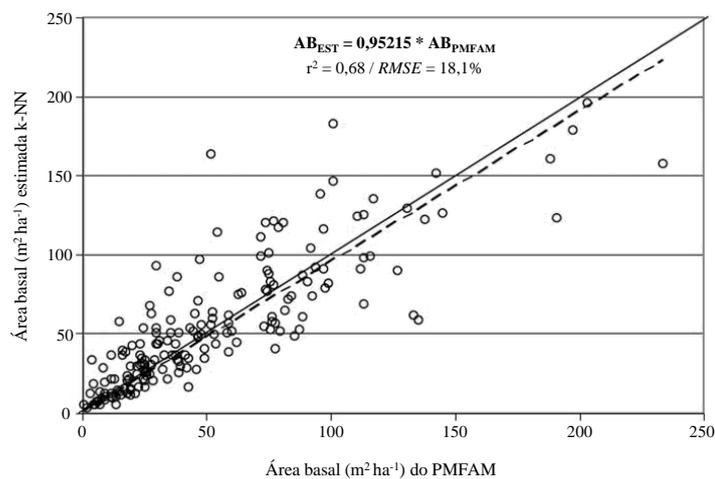
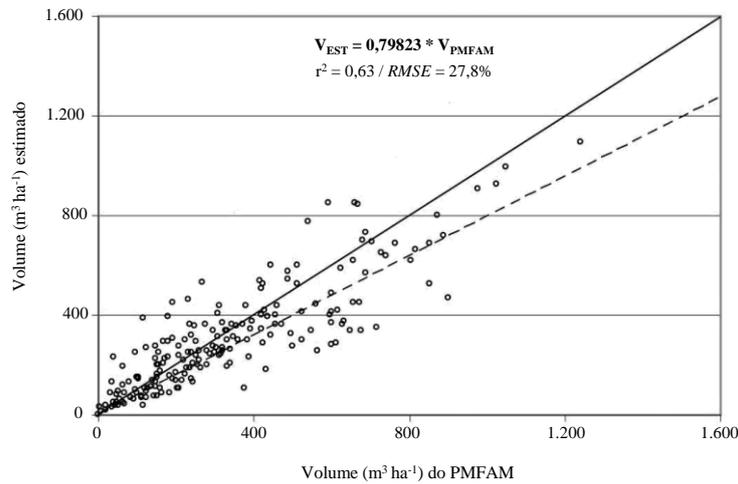


Figura 4. Espacialização dos dados de inventário florestal com a utilização do algoritmo k -NN.
 Figure 4. Spatialization of forest inventory data using the k -NN algorithm.



Legenda: linha tracejada corresponde a uma tendência entre os diferentes valores, enquanto a linha contínua é a bissetriz do gráfico.
 Figura 5. Comparação entre os valores de área basal do PMFAM e as estimativas do algoritmo k -NN.
 Figure 5. Comparison between the basal area from the PMFAM and the estimates from k -NN algorithm.



Legenda: linha tracejada corresponde a uma tendência entre os diferentes valores, enquanto a linha contínua é a bissetriz do gráfico.
 Figura 6. Comparação entre os valores de volume do PMFAM e as estimativas do algoritmo k -NN.
 Figure 6. Comparison between volume from PMFAM and the estimates from k -NN algorithm.

CONCLUSÕES

- Os resultados obtidos indicam a existência de correlações entre os atributos florestais levantados em campo e as informações espectrais das imagens multiespectrais, demonstrando assim a possibilidade de utilização da metodologia para a estimativa e a espacialização de atributos florestais. Outros aspectos a destacar foram as respostas fornecidas pelas diferentes janelas móveis utilizadas para a extração das informações espectrais, que demonstraram ser um parâmetro importante no desempenho dos modelos utilizados. A utilização de grandes dimensões da janela móvel para a extração das variáveis independentes nas imagens possibilitou uma maior probabilidade de incorporar superfícies externas, as quais não representam as amostras que estão sendo relacionadas com os dados levantados em campo. Dessa forma, as dimensões ideais para cada ambiente avaliado deverá apresentar características próprias, em função principalmente da fragmentação e da distribuição da cobertura florestal.
- Apesar de os inventários florestais serem fundamentais para o manejo das florestas, eles apresentam custos elevados, e as informações espaciais não contínuas muitas vezes impedem o seu uso em uma escala inferior à qual foi inicialmente projetado. Dessa forma, a metodologia utilizada demonstrou potencial de aplicação principalmente em função de suas características de espacialização das informações de forma contínua, de potencial de uso de forma integrada aos inventários tradicionais e de aplicação em extensas áreas florestais. Entretanto, a utilização dessa metodologia deve ser sempre acompanhada de uma validação, ou seja, um confronto dos resultados obtidos com o uso do algoritmo com os dados independentes obtidos em levantamento de campo.

REFERÊNCIAS

- CARNEGGIE, D. M.; LAUER, D. T. Uses of multiband remote sensing in forest and range inventory. **Photogrammetria**, v. 21, n. 4, p. 115 - 141, 1966.
- CHIRICI, G.; BARBATI, A.; CORONA, P.; MARCHETTI, M.; TRAVAGLINI, D.; MASELLI, F.; BERTINI, R. Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and Mediterranean forest ecosystems. **Remote Sensing of Environment**, v. 112, n. 5, p. 2686 - 2700, 2008.
- CHUVIECO, E.; CONGALTON, R. G. Application of remote sensing and geographic information systems to forest fire hazard mapping. **Remote Sensing of Environment**, v. 29, n. 2, p. 147 - 159, 1989.

- CURRAN, P. J.; DUNGAN, J. L.; GHOLZ, H. L. Seasonal LAI in slash pine estimated with Landsat TM. **Remote Sensing of Environment**, v. 39, n. 1, p. 3 - 13, 1992.
- DYMOND, J. R. Nonparametric modeling of radiance in hill country. **Remote Sensing of Environment**, v. 25, n. 1, p. 3 - 21, 1988.
- FAZAKAS, Z.; NILSSON, M. Volume and forest cover estimation over southern Sweden using AVHRR data calibrated with TM data. **International Journal of Remote Sensing**, v. 17, n. 9, p. 1701 - 1709, 1996.
- FRANCO-LOPEZ, H.; EK, A. R.; BAUER, M. E. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. **Remote Sensing of Environment**, v. 77, p. 251 - 274, 2001.
- FRANKLIN, S. E. Remote sensing for sustainable forest management. Boca Raton, USA: 2001.
- FUCHS, H.; MAGDON, P.; KLEINN, C.; FLESSA, H. Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory. **Remote Sensing of Environment**, v. 113, n. 3, p. 518 - 531, 2009.
- GU, H.; DAI, L.; WU, G.; XU, D.; WANG, S.; WANG, H. Estimation of forest volumes by integrating Landsat TM imagery and forest inventory data. **Science in China Series E: Technological Sciences**, v. 49, p. 54 - 62, 2006.
- HOLMSTRÖM, H. **Data acquisition for forestry planning by remote sensing based sample plot imputation**. 41f. Doctoral thesis, Department of Forest Resource Management and Geomatics, Swedish University of Agricultural Sciences, Umeå, 2001.
- HOLMSTRÖM, H. Estimation of single-tree characteristics using the kNN method and plotwise aerial photograph interpretations. **Forest Ecology and Management**, v. 167, n. 1-3, p. 303 - 314, 2002.
- HORLER, D. N. H.; AHERN, F. J. Forestry information content of Thematic Mapper data. **International Journal of Remote Sensing**, v. 7, n. 3, p. 405 - 428, 1986.
- INGRAM, J. C.; DAWSON, T. P.; WHITTAKER, R. J. Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks. **Remote Sensing of Environment**, v. 94, n. 4, p. 491 - 507, 2005.
- KATILA, M.; TOMPPO, E. Stratification by ancillary data in multisource forest inventories employing k-nearest neighbor estimation. **Canadian Journal of Forest Research**, v. 32, p. 1548 - 1561, 2002.
- KLEINN, C.; RAMÍREZ, C.; HOLMGREN, P.; VALVERDE, S. L.; CHAVEZ, G. A national forest resources assessment for Costa Rica based on low intensity sampling. **Forest Ecology and Management**, v. 210, n. 1 - 3, p. 9 - 23, 2005.
- KUSHWAHA, S. P. S. Forest-type mapping and change detection from satellite imagery. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 45, n. 3, p. 175 - 181, 1990.
- LANGFORD, K. J. Change in yield of water following a bushfire in a forest of *Eucalyptus regnans*. **Journal of Hydrology**, v. 29, n. 1 - 2, p. 87 - 114, 1976.
- LU, D.; MAUSEL, P.; BRONDÍZIO, E.; MORAN, E. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. **Forest Ecology and Management**, v. 198, n. 1 - 3, p. 149 - 167, 2004.
- MAGNUSSEN, S.; MCROBERTS, R. E.; TOMPPO, E. O. Model-based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. **Remote Sensing of Environment**, v. 113, n. 3, p. 476 - 488, 2009.
- MCINERNEY, D.; PEKKARINEN, A.; HAAKANA, M. Combining Landsat ETM+ with field data for Ireland's National Forest Inventory - a pilot study for County Clare. In: ForestSat 2005, 2005, Borås. **Anais do...** Borås, 2005. p. 12 - 16.

- MCROBERTS, R. E. Using satellite imagery and the k-nearest neighbors technique as a bridge between strategic and management forest inventories. **Remote Sensing of Environment**, v. 112, n. 5, p. 2212 - 2221, 2008.
- MCROBERTS, R. E.; NELSON, M. D.; WENDT, D. G. Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique. **Remote Sensing of Environment**, v. 82, p. 457 - 468, 2002.
- MCROBERTS, R. E.; TOMPPO, E. O.; FINLEY, A. O.; HEIKKINEN, J. Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. **Remote Sensing of Environment**, v. 111, n. 4, p. 466 - 480, 2007.
- MUINONEN, E.; TOKOLA, T. The usability of remote sensing for forest inventory and planning. In: SNS/IUFRO Workshop, 1990. Umeå. **Anais do...** Umeå, 1990. p. 35 - 42.
- ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the Great Plains with ERTS. In: NASA, Third Earth Resources Technology Satellite-1 Symposium, 1974. Greenbelt, MD. **Anais do...** Greenbelt, 1974. p. 301 - 317.
- SPANNER, M. A.; PIERCE, L. L.; PETERSON, D. L.; RUNNING, S. W. Remote sensing of temperate coniferous forest leaf area index the influence of canopy closure, understory vegetation and background reflectance. **International Journal of Remote Sensing**, v. 11, n. 1, p. 95 - 111, 1990.
- THOMPSON, I. D.; MAHER, S. C.; ROUILLARD, D. P.; FRYXELL, J. M.; BAKER, J. A. Accuracy of forest inventory mapping: Some implications for boreal forest management. **Forest Ecology and Management**, v. 252, n. 1 - 3, p. 208 - 221, 2007.
- TOMPPO, E. O.; GAGLIANO, C.; DE NATALE, F.; KATILA, M.; MCROBERTS, R. E. Predicting categorical forest variables using an improved k-Nearest Neighbors estimator and Landsat imagery. **Remote Sensing of Environment**, v. 113, n. 3, p. 500 - 517, 2009.
- TURNER, D. P.; COHEN, W. B.; KENNEDY, R. E.; FASSNACHT, K. S.; BRIGGS, J. M. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. **Remote Sensing of Environment**, v. 70, n. 1, p. 52 - 68, 1999.
- ZSILINSZKY, V. G. The practice of photo interpretation for a forest inventory (Commission VII). **Photogrammetria**, v. 19, p. 192 - 208, 1962.