ALTERNATIVE APPLICATION OF THE K-NN METHOD FOR MAPPING FOREST COVER TYPE

C. Gagliano [a]*, F. De Natale [a], F. Incerti [a], F. Maselli [b]

[a] CRA – Forest and Range Management Research Institute, ISAFA, 38050, Piazza Nicolini 6, Villazzano (Trento), Italy. Phone: +39 0461 381122; Fax: +39 0461 381131; caterina.gagliano@entecra.it
[b] CNR – Institute of Biometeorology, Via G. Caproni 8 – 50145 Firenze, Italy; maselli@ibimet.cnr.it

KEY WORDS: k Nearest Neighbour method, forest types, Landsat TM image classification

ABSTRACT:

The present study explores the performance of a non parametric approach applied to the classification of forest types in an area of Northern Italy (Province of Trentino). The investigation has been carried out by classifying a Landsat 5 TM image on the basis of a reference data set derived from ground surveys and visual interpretation of aerial photos.
The work compares the accuracies obtained with different configurations of a k-NN classifier. In particular the paper focuses on the choice of the best band combination and of the maximum horizontal distance between training field plots and unclassified pixels. The optimal parameter set is evaluated by a leave-one-out cross-validation strategy.
As the conventional configuration of the method leads to a distortion in the estimate of the different class surfaces, a correction methodology is also considered. It consists of computing the error probability from a preliminary classification and applying appropriate prior probabilities to each forest type to influence the pixel assignment to the different classes and consequently the area estimation.

1. INTRODUCTION

Nowadays field works represents the main way to accurately describe the state of forests. However, new technologies can provide alternative tools useful for collecting additional data.
In particular airborne and space borne sensors are currently used in forestry as an auxiliary data source, since they offer a synoptic view of large areas allowing at the same time the fast acquisition of data that can be digitally processed.
Several applications of remotely sensed data have been developed to estimate forest characteristics and to transform the results of field measurements into maps (Hame and Rauste, 1993). This is especially the case for biomass measurements (Ardö, 1992; Anderson et al., 1993), forest productivity (Ahern et al., 1991), tree cover percentage (Duncan et al., 1993), and leaf area index (Curran et al., 1992; Nemani et al., 1993). Correlations of satellite data, mainly based on Landsat 5 TM images, with forest inventory attributes have been found by several authors, e.g. for the discrimination of tree species (De Wulf et al., 1990), for the assessment of basal area (Franklin, 1986), canopy cover (Butera, 1986; Peterson et al.,1986; Oza et al., 1989), leaf area index (Spanner et al., 1990) biomass (Franklin, 1986), stand height (Horler and Ahern, 1986) and stem volume (Poso et al., 1984, 1987; Tomppo, 1990; Maselli et al., 2005; Muinonen and Tokola, 1990).
These studies have demonstrated that the operational use of digital data can be limited by the complexity of the relationships existing between spectral characteristics and forest attributes. For these reasons conventional classification methods, that assume relatively simple relationships between land surfaces and spectral characteristics, have often produced unsatisfactory results (Maselli et al., 2003).
The possibility to minimize these problems is provided by the use of more flexible classification methods as the non-parametric ones. Non-parametric methods, introduced by Fix and Hodges (1951), are discrimination procedures that do not require any assumptions about the statistical distribution of the data unlike the common parametric approaches such as likelihood ratios (Hardin, 1994).
The most common non-parametric method applied to produce classifications and estimates of continuous forest variables is the k-Nearest Neighbour (kNN).
The kNN method is used to generalize information from field plots to pixels for map production. The decision rule assigns each pixel to a class among the closest k labelled training pixels; the k pixels correspond to the field plots where different variables have been observed. Closeness is defined in terms of a metric distance in the multi-spectral space.
Nearest neighbour techniques have been used extensively for forestry estimations in Nordic countries. Tomppo (1990), Nilson (1994) and Trotter et al. (1997) used the technique to estimate forest volume; Muinonen and Tokola (1990), Franco-Lopez et al., (2001) and Bauer (2001) applied the method for satellite images classification.
The method is highly appreciated because it allows to predict all the measured variables at the same time for each pixel or for area units of interest (Tomppo and Halme, 2004).
However there are some problems that should be taken into account when integrating satellite images and field data for mapping and monitoring purposes. Although the non-parametric methods do not require any specific distribution of the data, they assume the sample distribution of field plots to be representative of the spatial variation of the forests in the study area. Moreover, when scaling from field measurements to remotely sensed data, different kinds of error may be involved: incorrect matching between plot area and corresponding pixel area, image noises, timing difference in field data and image acquisition. Finally, to apply the kNN-estimation method, the following assumptions should be done: i. the image pixel values depend only on the forest

* Corresponding author.

condition and not on the geographic location (Killki and Paivinen, 1987); ii. similar kinds of forest exist in the study area and their spectral responses depend on their characteristics. As a consequence of all factors listed above, kNN performances critically depend on a calibration process which aims to properly set the involved parameters.
The aim of this paper is to show the accuracy of kNN forest type classification when plot wise aerial photo-interpretation, combined with field data and other auxiliary information, are used as training data. In the calibration phase different configurations of a k-NN classifier were tested. In particular this study focuses on the choice of the best band combination and the setting of the maximum geographic distance between training field plots and unclassified pixels. Specific procedures to incorporate auxiliary information were also tested. The optimal setting of the parameters was defined by means of a leave-one-out cross-validation strategy.
As the conventional configuration of the method leads to a distortion in the estimate of the different class surfaces, a correction methodology was also considered. This consisted of computing the error probabilities from a preliminary classification to derive prior probabilities for each forest type; these probabilities were then used to guide the pixel classification and consequently to improve the area estimates.

2. STUDY AREA

The study area is located in the Eastern Italian Alps and covers about 40.000 ha of the Trento Province corresponding to the Alta-Valsugana district. This area belongs to the Alpine mountain region and its elevation ranges between 400 and 2400m. The forest composition is therefore strongly influenced by elevation, slope and aspect. Broad-leaved deciduous forests dominate lower and warmest areas. At the highest elevation forests are dominated by mixed conifers. The dominant forest species in the area is *Picea abies*, followed by *Larix decidua* and *Pinus sylvestris*. Among broadleaved species, the more frequent is *Fagus sylvatica*, while other coniferous and broadleaved species are less common.

3. MATERIALS AND METHODS

3.1 Data set and image pre-processing

A ground reference data set, distributed all over the study area, was obtained by means of the integration of field surveys, the forest planning database of Trentino Province and the photo-interpretation of 1:10.000 black and white orthophotos from 1996.
596 plots were identified with a resulting sampling intensity of 0.015 plot/ha.
The spectral signatures of all forest types found in correspondence of the training points were initially considered. Next, because of the poor separability and high spectral similarity among some original classes, some of these were aggregated and a classification scheme with six types was adopted. Table 1 shows the distribution of six different forest types in the reference data set.
A mono-temporal Landsat 5TM (path 192, row 28) dated 19 June 2000 was used, georeferenced and geometrically rectified to ensure correspondence between field plots and corresponding pixels.
No topographic correction was applied to the images because previous studies carried out in the same area showed that these methods don't produce remarkable improvements in the overall classification accuracy when compared to one reached with the use of the original non-corrected images (Puzzolo *et al*, 2005).

Forest types	number of points	%
Forests of *Picea abies/Abies alba*	268	46.3
Forests of other broadleaves	61	10.5
Forests of other conifers	86	14.9
Forests of *Castanea sativa*	12	2.1
Forests of *Fagus sylvatica*	65	11.2
Forests of *Larix deciduas*	87	15.0

Table 1. Forest types identified in the study area

Since the relationship between forest and spectral characteristics may not be constant over a large area (mainly because of the elevation variations which can affect the vegetation structure and composition), a 10m digital elevation model was used as ancillary source of information.

3.2 kNN technique

The kNN algorithm assigns each unknown pixel to the field attributes of the most similar reference pixels for which field data exists. Similarity is expressed as the distance between the query point and the neighbouring ones.
When using kNN, different types of distance metrics can be selected:

- Minkoski distance;
- Euclidean distance;
- Manhattan distance;
- Mahalanobis distance

Let D = $\{p_1, \ldots, p_n\}$ be a set of training pixels *pi*. For each pixel a vector of different attributes *m* is known. Given the training pixel D = $\{p_1, \ldots, p_n\}$, the spectral distance between the pixel *p* to be classified and each training pixel p_i is computed. To this unknown pixel an attribute vector *m* is assigned; m is obtained by calculating the mean or mode over all k-nearest p_i for continuous and categorical attributes respectively. Some applications assign different weights to the training pixels depending on auxiliary variables.

3.3 kNN configuration

In the current test the output variable was categorical and the nearest neighbour classifier assigned the most frequent class to each pixel *p*. In our kNN configuration each of the selected nearest neighbours contributes to the decision in the same way.
The distances between neighbours were calculated using the Euclidean metric, which is the easiest to compute, because the use of other metrics does not generally contribute to a considerable improvement of the estimation accuracy (Maselli *et al.*, 2005).
The Euclidean distance is defined as the shortest distance between two points in the multi-spectral space. It is computed as (1):

$$d_{p(p_i)} = \sqrt{\sum_{j=1}^{nf} (x_{p,j} - x_{(p_i),j})^2} \quad (1)$$

where $x_{p,j}$ = digital number for the feature *j*; *nf* = number of features in the spectral space.

3.4 kNN calibration

Prior to any implementation, the kNN algorithm needs to be calibrated to find its best configuration. This step allows for the identification of the kNN parameters able to produce the best classification accuracy. Following a widely used approach, the current calibration was performed using a leave-one-out strategy. This technique uses the training data sample to estimate the classification accuracy by leaving out one sampling unit and classifying that unit on the basis of the rest of the sample. This omission is repeated for every unit allowing to calculate the classification error. (Gong, 1986).
In order to minimize the effects of possible positional errors, mean filters of 3x3 and 5x5 pixels were applied to the study images. The leave-one-out cross validation results derived from the filtered images were then compared with the ones obtained using the original TM bands.
As already mentioned, the relationship between forest and spectral characteristics may not be constant over a large area. Observations for adjacent pixels are in fact expected to be positively correlated while distant observations are expected to be uncorrelated (McRoberts *et al.*, 2002).
For this reason we applied a method that allowed us defining optimal geographical reference areas, in both vertical and horizontal directions (VRA and HRA), from which the nearest field plots are selected. The area has to be large enough to include all variation of forest cover types and to exclude, at the same time, more distant zones with different characteristics. Different combinations of VRA and HRA were tested and compared in terms of cross validation accuracy.
The last calibration step concerns the band combination. Since all available bands are not always needed and can even reduce the classification accuracy, a subset of bands was selected to be used in the final classification. All 63 possible combinations of the six TM bands were tested to select the subset which gave the highest classification accuracy. The thermal infrared band (TM6) was not considered due to its lower spatial resolution.
After the calibration, the optimal number of neighbours was chosen according to two different criteria: the k value had to be large enough to minimize the misclassification probability and small enough to minimize the bias. When k=1, the prediction incorporates all the variability existing in the observations, whereas the variability is reduced when k>1 because the prediction is based on the mode of multiple observations (McRoberts *et al.*, 2002).
The best configuration of the kNN classification we found was then applied to the study area; the accuracy results, derived through a leave-one-out cross validation method, are presented in terms of global error matrix, overall accuracy, producer's accuracy, user's accuracy, and kappa index. The results of Chi-square (χ^2) tests, which allow to identify systematic errors (i.e. classification bias), are also presented; they were used to check if the true class distribution of the original sample is well represented.

The correspondence between the extent of the ground references and that of the classified areas (Maselli, 1990) was also computed by (2):

$$corr = \frac{N - \sum_{i=1}^{n} abs(c_i - r_i)}{N} \quad (2)$$

corr = correspondence rows/columns
N = total number of pixels
Ci = marginal summary of column i
Ri = marginal summary of row i
In this first trial we assumed that the prior probabilities of a given pixel were equal for all classes.

3.5 Alternative kNN configuration

The values found in the error matrix were used for a discrimination process based on prior probabilities. The study proposes to use the error matrix derived from the first classification as an additional information source in order to reduce the interclass confusion. The matrix, in fact, has been shown to be useful for correcting the statistical weights in discrimination processes and it can be used to improve area estimates of different cover types (Maselli *et al.*, 1990; 1992). The information deriving from the error matrix was incorporated into the classification process changing the prior probability for each class. The prior probabilities were derived from the ratios of the class frequencies in the ground reference data (marginal summary of columns in the error matrix) over those in the classification data (marginal summary of rows in the error matrix) and then used for changing the probability extraction of each class. The main difficulty in following this approach is to calculate prior probabilities appropriate to achieve optimal results.

4. RESULTS AND DISCUSSION

The algorithm was first used with mono-spectral data. In accordance with other studies (Lillesand *et al.*, 2004), the best single band for discriminating forest types was band 4. Other authors found the best results are achieved using visible or middle infrared bands (McRoberts *et al.*, 2002). A reason for differences could be found in the fact that the optimal band selection depends upon both the image and the application (Spanner *et al.*, 1984). In addition, the optimal band combination depends on the characteristics of the training data, forest density and considered season (McRoberts *et al.*, 2002).
The two tested filters produced results comparable with those produced without filter: the accuracy varies between 40% and 50%. On the whole, images filtered using a 5x5 mean filter provided best results in terms of overall accuracy (Figure 1) but the smoothing effect of filtering is questionable because of the masking effect on the spectral differences. The argument for using a multiple pixel window to address plot location error (Franco Lopez, 2001) remains debatable but, in our test, the classification was completed including filter.
Different HRA e VRA sizes were considered. Probably due to the characteristics of training set and of study area, the 2km HRA and the 500m VRA reached the maximum cross validation accuracy (67%).

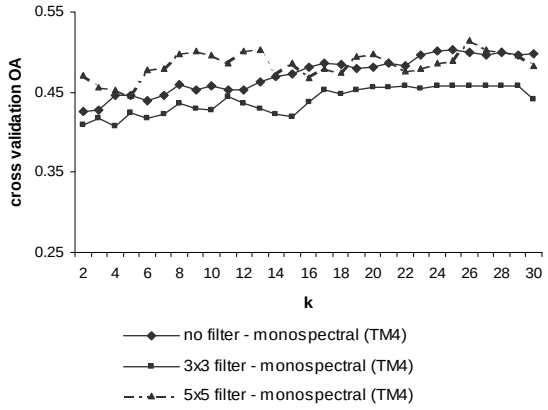

Figure 1. Overall classification accuracy for three different filters at increasing value of k

The geographical space defined by this combination allows representing the spatial variation of the field cover. The accuracy decreased to 62% when considering the maximum geographical window (5km HRA and 700m VRA), probably because a too wide selection area led to include a high number of training plots which were less representative of local conditions.

Comparing the cross validation accuracies reached for the 63 band combinations we tested, we can observe that the overall accuracy did not vary too much and ranged between 0.62 and 0.67. The band combination which produced the most accurate classification was TM2, TM3 and TM7, confirming that the classification accuracy with all six bands was inferior to the accuracy obtained using a band subset. This finding can be explained by the high correlation between bands in the same spectral regions.

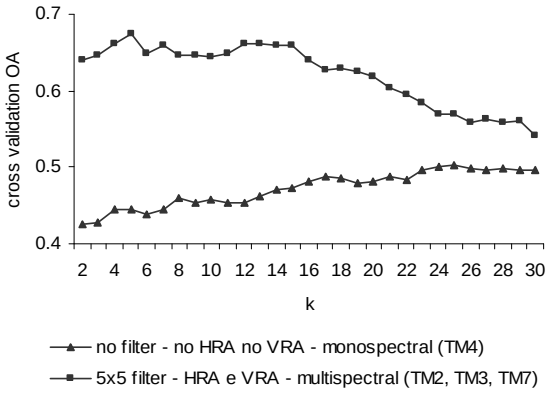

Figure 2. Overall classification accuracy before and after calibration at increasing value of k

Using the Euclidean distance, equal weighting among neighbours, a mean filter of 5x5 pixels, a HRA of 2km, a VRA of 500m and k=5 an overall accuracy of 67% was achieved (Figure 2).

When the number of neighbours was varying from 5 to 1 the accuracy decreased approximately of 4% (Table 2). but the true class distribution of the original sample was preserved ($\chi^2_{est}<\chi^2_{tab}$). This means that the classification accuracy can be improved when k increases, but the estimates will be affected by bias.

The most frequent class, which aggregates *Picea abies* and *Abies alba* forests, showed the higher accuracy. The *Larix decidua* class, even though represented by a high plot number, was not correctly classified (Table 3) being it was confused with other conifers.

k	OA (%)	Chi-square
1	61.66	1.4
2	63.90	67.0
3	64.59	22.1
4	66.15	33.3
5	67.36	23.8

Table 2. Overall classification accuracy and Chi-Square test (χ2tab=11.07; P=0.05;df=5) calculated for different numbers of neighbours

Forest types	PA	UA
Forests of *Picea abies/Abies alba*	88.06	76.87
Forests of other broadleaves	33.33	55.77
Forests of other conifers	53.49	63.89
Forests of *Castanea sativa*	52.31	53.13
Forests of *Fagus sylvatica*	58.33	53.85
Forests of *Larix deciduas*	62.30	53.52
OA	67.36%	
Kappa	0.5315	

Table 3. Confusion matrix at the end of calibration process

Applying the modified prior probabilities (Table 4) the producer's accuracy increased of 10% for larch forests while the mean producer accuracy did not change significantly. Overall accuracy and Kappa coefficient were not suited to appreciate the improvements brought by these prior probabilities but the marginal summaries of the error matrix showed a remarkable improvement. As a result, the overall accuracy and Kappa remained constant, respectively around 65% and 0.5, but the correspondence coefficient passed from 0.82 to 0.95, indicating an improved area estimation of most classes. Moreover the χ^2 statistic was not significant ($\chi^2_{est}<\chi^2_{tab}$, P=0.05), indicating that the classification was not affected by systematic errors (Table 5).

Forest types	PA	UA
Forests of *Picea abies/Abies alba*	85.07	81.14
Forests of other broadleaves	45.98	44.94
Forests of other conifers	53.49	57.50
Forests of *Castanea sativa*	44.62	47.54
Forests of *Fagus sylvatica*	50.00	50.00
Forests of *Larix deciduas*	55.74	60.71
OA	66.15%	
Kappa	0.5227	

Table 4. Confusion matrix with the prior probability use

5. CONCLUSIONS

Since the '80s, the kNN method is commonly used for mapping forest variables. The method requires the selection of a few estimation parameters by a calibration process.

In this study, the use of filtered images, combined with setting a spatial range in which the nearest field plots are chosen and

class	True reference	Classification without priors (k=5)	Classification whit priors
Forests of *Picea abies/Abies alba*	46.29%	53.02%	48.53%
Forests of *Larix deciduas*	10.54%	12.26%	9.67%
Forests of other conifers	14.85%	12.44%	13.82%
Forests of *Fagus sylvatica*	2.07%	2.25%	2.07%
Forests of *Castanea sativa*	11.23%	11.05%	10.54%
Forests of other broadleaves	15.03%	8.98%	15.37%
		OA: 67.36%	OA: 66.15%
		corr: 0.82	corr: 0.95
		$\chi 2$: 23.8	$\chi 2$: 1.75

Table 5. Marginal summaries of confusion matrixes without and with prior probabilities compared to the ground references, correspondence coefficients and Chi-Square tests ($\chi 2$tab=11.07; P=0.05; df=5)

the selection of an appropriate band combination made possible to improve the classification accuracy.

When the calibration process was implemented the accuracy rose from 50% to 70%. One of the most important steps in kNN method is the choice of k. The results of Chi Square test showed that the selection of k=1 is appropriate to maintain the variance of observations and to preserve appropriate variability in predictions. Besides, the leave-one out cross validation showed that the selection of a higher k, equal to 5, is useful to maximize the classification accuracy.

The error matrix derived from the first classification process was used as ancillary data source. The use of this information to set prior probabilities improved the discrimination of some classes but not the overall accuracy of the classification. However the use of external information preserved the class distribution of the original sample and provided better area estimation.

Considering that the method was only tested using a single data set and on a single study area, further investigations are needed. The proposed configuration should be validated in areas with different forest types and topographic characteristics and with training dataset of different sizes.

6. ACKNOWLEGMENTS

The authors would like to thank the RI.SELV.ITALIA project founded by the Italian Ministry of Agriculture and Forest Policies for data providing. The authors would like to thank also Dr. Ciaghi for the support in the photo interpretation and Marco Trombetti for the text revision.

7. REFERENCES

Ahern, F.J., Erdle, T., Maclean, D.A., Kneppek, I.D., 1991. A quantitative relationship between forest growth rates and Thematic Mapper reflectance measurements. *Int.J.Rem.Sens*., 12, pp. 387-400.

Anderson, G.L., Hanson, J.D., Haas, R.H., 1993. Evaluating Landsat Thematic Mapper derived vegetation indices for estimating above-ground biomass on semiarid rangelands. *Rem.Sens.Env*., 45, pp. 165-175.

Ardö, J., 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *Int.J.Rem.Sens*., 13, pp. 1779-1786.

Bauer, M.E., Burk, T.E., Ek, A. R., Coppin, P.R., Lime, S.D., Walsh, T.A., Walters, D.K., Befort, W., Heinzen, D. F., 1994. Satellite inventory of Minnesota forest resources. *Photogrammetric Eng. Remote Sensing*, 60 (3), pp. 287-298.

Butera, K., 1986. A correlation and regression analysis of percent canopy closure versus TMS spectral response for selected forest sites in the San Juan national forest, Colorado. *IEEE Trans. Geosci. Remote Sens.*, GE-24(1), pp. 122-129.

Curran, P.J., Dungan, J.L., Gholz, H.L., 1992. Seasonal LAI in slash pine estimated with Landsat TM. *Rem.Sens.Env*., 39, pp. 3-13.

De Wulf, R.R., Goossens, R.E., De Roover, B.P., Borry, F.C., 1990. Extraction of forest stand parameters from panchromatic and multispectral SPOT-1 data. *Int.J.Rem.Sens*., 11, pp. 1571-1588.

Duncan, J., Stow, D., Fanklin, J., Hope, J., 1993. Assessing the relationship between spectral vegetation indices and shrub cover in the Jornada Basin, New Mexico. *Int.J.Rem.Sens*.,14, pp. 3395-3416.

Fix, E., Hodges, J.L., 1951. Discriminatory analysis, nonparametric discrimination consistency properties. Technical Report 4, Randolph Filed, TX: US Air Force.

Franco-Lopez, H., Ek, A.R., Baue, M.E., 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbours method. *Rem.Sens.Env*, 77, pp. 251-274.

Franklin, J., 1986. Thematic mapper analysis of coniferous forest structure and composition. *Int.J.Rem.Sens*, 7(10), pp. 1287-1301

Gong, G., 1986. Cross-validation, the jack-knife, and the bootstrap: excess error estimation in forward logistic regression. *Journal of the American Statistical Association*, 81(393): pp.108-113.

Hame T. and Rauste Y., 1993. Multitemporal satellite data in forest mapping and fire monitoring. Proceedings to the International Workshop. Satellite technology and GIS for Mediterranean forest mapping and fire management held in Thessaliniki, Greece 4th and 6th November.

Hardin, P.J., 1994. Parametric and Nearest neighbour methods for hybrid classification: a comparison of pixel assignment accuracy. *Photogrammetric Eng. Remote Sensing*, 60(12), pp. 1439-1448.

Horler, D.N.H. and Ahern, F.J., 1986. Forestry information content of Thematic Mapper data. *Int.J.Rem.Sens*., 7(3), pp. 405-428.

Kilkki, P. and Paivinen, R. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. In:Remote Sensing-Aided Forest Inventory. Seminars organised by SNS and Taksaattoriklubi, Hyytiala, Finland, pp. 209-212.

Lillesand, T.M., Kiefer, R.W., Jonathan, W. Chipman, 2004. *Remote sensing and image Interpretation*. 5th ed. New York, Wiley, pp 763

Maselli, F., Conese, C., Zipoli, G., Pittau, M.A., 1990. Use of error probabilities to improve area estimates based on maximum likelihood classification. *Rem.Sens.Env*, 31, pp. 155-160.

Maselli, F., Conese, C., Petkov, L., Resti, R., 1992. Inclusion of prior probabilities derived from a non parametric process into the maximum likelihood classifier, *Photogrammetric Eng. Remote Sensing*, 58, pp. 201-207.

Maselli, F., Bottai, L., Chirici, G., Corona, P., Marchetti, M., Travaglino, D., 2003. Stima di attributi forestali in ambiente mediterraneo tramite integrazione di misure a terra e dati telerilevati. *Italia Forestale e Montana*, 4, pp. 251-263.

Maselli, F., Chirici, G., Bottai, L., Corona, P. Marchetti, M., 2005. Estimation of mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images. *Int.J.Rem.Sens*., 26(17), pp. 3781-3796.

McRoberts, R., Nelson, M., Wendt, D.G., 2002. Stratified estimation of forest area using satellite imagery, inventory data, and the k-NN technique. *Rem.Sens.Env* 82, pp. 457-468.

Muinonen, E., & Tokola, T., 1990. An application of remote sensing for communal forest inventory. The Usability of Remote Sensing For Forest Inventory and Planning. Proceedings from SNS/IUFRO workshop in Umea 26–28 February 1990. Swedish University of Agricultural Sciences, Remote Sensing Laboratory, Report 4. 35-42. Umea, Sweden. ISBN 91-576-4208-7

Nemani, R., Pierce, L., Running, S., 1993. Forest ecosystem processes at the watershed scale: Sensitivity to remotely sensed Leaf Area Index estimates. *Int.J.Rem.Sens*., 14, pp. 2519-2534.

Nilson, T., and Peterson, U., 1994. Age dependence of forest reflectance: analysis of main driving factors. *Rem.Sens.Env*., 48, pp. 319-331.

Oza, M.P., Srivastava, V.K., Pariswad, B.S. Setty, K.R.V., 1989. Relationship between Landsat MSS data and forest tree parameters. *Int.J.Rem.Sens*., 10(11), pp. 1813-1819

Peterson, D.L., Westman, W.E., Stephenson, N.L., Ambrosia, V.G., Brass, J.A., Spanner, M.A., 1986. Analysis of forest structure using thematic mapper data. *IEEE Trans. Geosci. Remote Sens.*, 24, pp. 113-121

Poso, S., Hame, T. Paananen, R., 1984. A method of estimating the stand characteristics of a forest compartment using satellite imagery. *Silva Fennica*, 18, pp. 261-292

Poso, S., Paananen, R. Simila, M., 1987. Forest inventory by compartments using satellite imagery. *Silva Fennica*, 21, pp. 69-94

Puzzolo, V., Panizza, M., De Natale, F., Bruzzone, L., 2005. Correzioni topografiche di immagini Landsat TM e SPOT HRV in aree alpine orograficamente complesse. *Rivista italiana di Telerilevamento*, 32, pp. 67-77.

Spanner, M.A., Brass, J.A., Patterson, D.L., 1984. Feature selection and the information content of thematic mapper simulator data for forest structural assessment, *IEEE Trans. Geosci. Remote Sens.*, 22 (6), pp. 482-489.

Spanner, M.A., Pierce, L.L., Peterson, D. Running, S., 1990. Remote Sensing of temperature coniferous forest leaf area index. The influence of canopy closure, understory vegetation and background reflectance. *Int.J.Rem.Sens*., 11(1), pp. 95-111

Tomppo, E., 1990. Designing a Satellite Image-Aided National Forest Survey in Finland. The Usability of Remote Sensing For Forest Inventory and Planning. Proceedings from SNS/IUFRO workshop in Umea 26–28 February 1990. Swedish University of Agricultural Sciences, Remote Sensing Laboratory, Report 4. 43– 47. Umea, Sweden. ISBN 91-576-4208-7

Tomppo, E., Halme, M., 2004 Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: a genetic algorithm approach. *Rem.Sens.Env*., 92, pp. 1-20.

Trotter, C.M., Dymond, J R., Goulding, C.J., 1997. Estimation of timber volume in a coniferous plantation forest using Landsat TM. *Int.J.Rem.Sens*., 18, pp. 2209-2223.