Strategy to attain high spatial accuracy in Forest Cover Classification

R.K. Gupta, D. Vijayan, T.S. Prasad

National Remote Sensing Agency, Balanagar, Hyderabad- 500 037, India
(training@nrsa.gov.in, rkgupta.ijrs@gmail.com /Fax: +91 40 2387 8648[Training])

Forest cover and its type have primary role in the processes associated with land and global change. Not only the area statistics for the different type of forest covers but also the correctness of their spatial distribution (matching of classified output with GIS overlay) are important for process studies. As maximum likelihood (ML) is widely practiced classification algorithm for extracting thematic information from satellite images, critical evaluation was undertaken using IRS LISS-III image of Antilova tropical moist deciduous forest (bounded by 17˚ 50' to 17˚ 56' N in latitude and 81˚ 45' to 81˚ 54' E in longitude) for which 100 % ground information in the from of GIS overlay was available. GIS overlay has 9 thematic classes i.e., 27.13% dense (DF), 25.60% Semi-evergreen (SE), 29.38% mixed (MF), 0.25% bamboo(BA), 5.70% teak (TK) forests, 5.88% grassland (GL), 4.83% podu / blank (PO), 1.21% Settlements (SET) and water 0.026% (WA).

ML classifier, in general, starts with equal *'a priori'* probability for all the classes (method 'a'). Availability of information on % cover under each thematic class enables assigning of *'a priori'* probability to each thematic class (method 'b'). Method 'b' always gave better results as compared to method 'a'. With the goal to improve classification accuracy (CA), the GL and MF classes that had high standard deviation of 10.29 and 11.29 in NIR band were divided into subclasses. Inclusion of sub-classes in GR improved the area statistics and spatial matching for low % cover classes (teak, grassland and podu). Splitting for MF into subclasses deteriorated the results for major cover classes as MF subclasses were becoming spectrally similar to major forest classes.

Spectrally, bamboo is a subset of MF and has about 20 to 30% spectral overlap with DF and PO besides having a very low *'a priori'* probability of 0.0025. Thus, exclusion of bamboo as a class did not bring any significant change in area statistics, but has dramatically improved the spatial matching in 82 to 88% range for DF, SE, TK and PO; 75%, 71% and 48% for MF, GR and SET, respectively.

In practice, very pure limited training sets are identified for ML classifier (MLp). Thus, MLp based exercise was undertaken for all the approaches discussed in previous paragraph. Spatial matchings were 95%, 95%, 79%, 94%, 75%, and 80% for DF, SE, MF, TK, GL and PO, respectively with marginal degradation in settlement.

Thus, assigning of *'a priori'* probability even with conventional limited pure training areas provides better classification in terms of spatial matching as well as area statistics. Estimate of *'a priori'* probabilities for each thematic class could be obtained from unsupervised classification wherein thematic label to classes could be assigned either using existing maps or limited ground truth.