

ASSESSING DOWNSCALING LIMITS OF SPATIAL RESOLUTION FOR AWiFS AND LANDSAT 8 DATA AS COMPARED TO LISS IV

Chauhan A.^a, Denis. M. D^{b*}, Kumar. M^c

^aResearcher akankshachauhan6392@gmail.com, ^bProfessor (derrickmdenis@gmail.com), ^cAssistant Professor, mukesh_fo@yahoo.co.in, Dept. of Soil, Water, Land Engineering and Management, Vaugh School of Agricultural Engineering and Technology, Sam Higginbottom Institute of Agriculture Technology and Sciences, Allahabad, Uttar Pradesh, India, 211007

KEY WORDS: Downscaling, AWiFS, LANDSAT8, LISS IV, Kappa coefficient

ABSTRACT

Downscaling spatial information has a certain limit. Change in resolution of input image from its original resolution either loses or gains information. To understand this limit accuracy of the downscaled information has to be assessed. This research assesses the downscaled information of AWiFS and LANDSAT 8 as compared and validated by LISS IV. A total of 200 sample points were collected using Systematic Random Sampling. The results showed that the overall accuracy for the supervised classification was 95% for LISS IV where Kappa statistics was 0.92. As the resolution LANDSAT 8 was reduced from 10% to 50% the overall accuracy varied from 84% to 73.5% with 0.79 to 0.65 Kappa coefficient. For AWiFS at the same reducing resolutions, the overall accuracy varied from 80% to 67% with 0.72 to 0.56 kappa coefficients. The results showed that the downscaling limits for AWiFS and LANDSAT8 are only 10% of their actual resolutions. Any downscaling beyond this limit will result in precious information being lost.

1. INTRODUCTION

Downscaling refers to an increase in spatial resolution (Atkinson, 2012). Its affect can be observed in image classification (Althausen, 2002; Barnsley, 1999; Estes & Loveland., 1999; Lefsky & Cohen, 2003) The parameters for considering a particular resolution as input depends upon the user's needs, funds available and its availability as open source. In general, at a local scale, high spatial resolution data such as IKONOS and QuickBird data are helpful but expensive. At a regional scale, medium spatial resolution data such as those from Landsat TM and Terra Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER) are the most frequently used and are available as open source. At a continental or global scale, coarse spatial resolution data such as Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS), and System Pour l'Observation de la Terre (SPOT) vegetation data are preferable. Atmospheric condition such as clouds and vapor are often an obstacle for capturing high-quality optical sensor data.

An image consists of an array of pixels or grid cells and many layers or bands. Each pixel has a Digital Number (DN) representing the intensity of the received signal reflected or emitted by a given area of the earth surface. The size of the area belonging to a pixel is called the spatial resolution. Further bands are created by the sensor that collects energy in specific wavelengths of the electro-magnetic spectrum.

*Corresponding author derrickmdenis@gmail.com

The classification of remotely sensed data has long attracted the attention of the remote sensing community. Scientists and practitioners have made great efforts in developing advanced classification approaches and techniques for improving classification accuracy (Aplin, Atkinson, & Curran, 1999; Blaschke, 2010; Franklin, 2002; G.M.Foody, 1996; Gallego, 2004; Ghimire, Rogan, & Miller, 2010; Gong.P & Howarth, 1992; Kontoes, G.Wilkinson, A.Burriel, S.Goffredo, & Megier, 1993; Lu & Weng, 2007; Pal & Mather, 2003; San Miguel-Ayaz & Biging, 1997; Stuckens, Coppin, & Bauer, 2000) Classification of remotely sensed data generally produces discrete information categories. It is implicitly assumed that spectral class memberships are precisely defined, so that the attribution of a pixel to a land-cover category is always feasible (Curran & Williamson, 1985). Implicit in the traditional classification process is the concept that each feature vector should be mapped into one of the classes of interest.

In reality some classes have very close spectral reflectance values and therefore they overlap in the feature space. Overlap can also indicate mistakes in the definition of the training sites. This produces spectral confusion. Hence it has to go through several iterations of training site adjustment, signature development, and signature evaluation before achieving satisfactory signatures. Scatter Plot is useful for evaluating the quality of one's signatures.

Today, validation or accuracy assessment is an integral component of most mapping projects incorporating remotely sensed data. Other spatial information may not be so stringently evaluated, but at least requires meta-data that documents how the information was generated. In the 1980s, the use of the error matrix became a common tool for representing the accuracy of individual map categories. By

the 1990s, most maps derived from remotely sensed imagery were required to meet some minimum accuracy standard. Accuracy assessment or validation is a key component of any project employing spatial data. This assessment is important to know how well you are doing and to learn from your mistakes, the ability to quantitatively compare methods and the ability to use the information resulting from spatial data analysis in some decision-making process.

A common method for accuracy assessment involves the use of an error matrix, for which the literature has provided the meanings of and calculation methods for overall accuracy, producer's accuracy, user's accuracy, and kappa coefficient (C.P.Smits, Dellepiane, & Schowengerdt, 1999; G. R.

MATERIALS AND METHOD

Datasets used

Topo sheet of Allahabad district (63G15) was used as a raw material of 1:50,000 scales along with three separate datasets i.e. LISS IV, LANDSAT 8 and AWiFS. Each scene was geo referenced and suitable image enhancements are applied to facilitate the delineation and interpretation of different thematic information. The satellite data, Resourcesat-2 (LISS IV) of January 2013, LANDSAT 8 (OLI) of February 2014 and IRS P6 (AWiFS) of February 2010 for Allahabad were used in this study. These image encompasses of large areas of farmlands, residents, wetlands, some of water bodies and more. These datasets consists of different pixels and covers the study area. The advantage of these datasets is availability of the referenced image produced from field survey, which is used for accuracy purpose. Data collection is first step in any assessment procedure, and may be the single most important factor in accuracy assessment. Since an assessment will meaningless if the reference data cannot be trusted. After collecting all the data from various sources we were mosaic all scenes of all satellite data and the study area has been extracted from mosaiced scenes by using the boundary of as area of interest (AOI) tool. All together seven classes were identified and classified for both supervised and unsupervised classification. These classes are shown in Table 1.

Sr. No.	Class	Associate Class
1.	Agriculture land	Cultivated land, Harvested land
2.	Agriculture crop	Agricultural land, Grassland, Turf and grass.
3.	River/Water Bodies	Streams, Lakes, Reservoirs, Ponds and canals.
4.	Barren Land	Exposed soil, Bare exposed rock and Transitional areas.
5.	Settlement	Residential, Industrial, Industrial and commercial complexes, Mixed urban or built-up land and Road.
6.	Forest/Trees	Evergreen forest land, Shrub and bushes
7.	Sand	Bank of river

Table 1 Classes identified in the study area.

Preprocessing of satellite images prior to image classification was done. There after resampling was done to do the accuracy study. This was followed by supervised and unsupervised classification. The classification was done in five classes and given in Table 6. Scatter plots were obtained.

Congalton, 1991; R. G. Congalton & Green, 2008; Foody, 2002) Since multiple sources of sensor data are now readily available, image analysts have more choices to select suitable remotely sensed data for a specific study. Considering these facts three varying resolution remotely sensed data i.e. Linear Imaging Self Scanner (LISS IV) having 5.8 m spatial resolution, LANDSAT 8, 30 m spectral resolution and Advanced Wide Field Sensor (AWiFS), 56 m spectral resolution data were considered for accuracy assessment using error matrix, producer's accuracy, user's accuracy, overall accuracy and kappa coefficient to assess downscaling limits of spatial resolution for AWiFS, LANDSAT 8 as compared to LISS IV

Accuracy Assessment

Accuracy assessments are performed on the thematic information. The percentage of overall accuracy was calculated using following formula:

$$\text{Overall accuracy} = \frac{\text{Total number of correct samples}}{\text{Total number of samples}} \times 100 \quad (1)$$

Besides the overall accuracy, classification accuracy of individual classes was calculated in a similar manner. The two approaches are user's accuracy and producer's accuracy. The producer's accuracy is derived by dividing the number of correct pixels in one class divided by the total number of pixels as derived from reference data. In this study, the producer's accuracy measures how well a certain area has been classified. It includes the error of omission which refers to the proportion of observed features on the ground that is not classified in the map. Meanwhile, user's accuracy is computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category.

Kappa coefficient (K) is another measurement used in this study. It is calculated by multiplying the total number of pixels in all the ground verification classes (N) by the sum of the confusion matrix diagonals ($\sum X_{kk}$), subtracting the sum of the ground verification pixels in a class time the sum of the classified pixels in that class summed over all classes ($\sum X_{k\Sigma} \sum Y_{k\Sigma}$), where $X_{k\Sigma}$ is row total and $Y_{k\Sigma}$ is column total, and dividing by the total number of pixels squared minus the sum of the ground verification pixels in that class times the sum of the classified pixels in that class summed over all classes. The value of Kappa lies between 0 and 1, where 0 represents agreement due to chance only. 1 represents complete agreement between the two data sets. Negative values can occur but they are spurious. It is usually expressed as a percentage (%). Kappa statistic can be a more sophisticated measurement to classifier agreement and thus gives better interclass discrimination than overall accuracy. The calculation of Kappa statistic k is as follows:

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

Where,

x_{ij} = number of counts in the ij th cell of the confusion matrix

N = total number of counts in the confusion matrix

x_{i+} = marginal total of row i

x_{+i} = marginal total of column i

RESULT AND DISCUSSION

Radiometric Normalization of LISS IV data is done to significantly reduce the mean absolute differences between reference scene pixels and corresponding pixels from subject scenes that may be used to fill cloud gaps, as expected. The selection of suitable band combination is essential for LISS IV satellite images for visual interpretation. This enhancement technique was carried out in order to produce a better image classification for digital classification. The images of Allahabad area appear much better after performed enhancement technique. Brightness and contrast level panels were used to adjust the image view to an appropriate level. The image after enhancement clearly showed the separation of the land cover features. Band combinations of 4-3-2 (RGB) for LISS IV was selected as the best combination and later used for digital classification. It has been seen that before histogram matching mean and standard deviation of reference scene is far from corrected scene. But after the histogram matching mean and standard deviation of corrected scene comes very close to mean and standard of reference scene. The statistics of LISS IV scenes before and after radiometric normalization are given in Table 2, Table 3 and Table 4.

Table 2: Statistics of LISS IV scenes before and after radiometric normalization for scene 1 and scene 3

Before Histogram match	Band2	Band3	Band4
Std. dev.(scene1)	38.65	36.48	56.28
Std. dev.(scene3)	37.97	35.14	76.32
Mean (scene1)	72.37	66.08	102.47
Mean (scene3)	69.81	58.30	130.31
After Histogram match			
Std. dev.(scene1)	37.70	34.54	75.75
Mean (scene1)	69.35	57.04	128.61

Table 3: Statistics of LISS IV scenes before and after radiometric normalization for scene 2 and scene 3

Before Histogram match	Band2	Band3	Band4
Std. dev.(scene2)	44.95	39.62	54.62
Std. dev.(scene3)	37.97	35.14	76.32

Mean (scene2)	85.32	74.34	101.11
Mean (scene3)	69.81	58.29	130.31
After Histogram match			
Std. dev.(scene2)	37.62	34.63	75.70
Mean (scene2)	69.19	57.25	128.88

Table 4: Statistics of LISS IV scenes before and after radiometric normalization for scene 4 and scene 3

Before Histogram match	Band2	Band3	Band4
Std. dev.(scene4)	37.61	33.31	76.71
Std. dev.(scene3)	37.97	35.14	76.32
Mean (scene4)	67.42	54.22	132.42
Mean (scene3)	69.81	58.29	130.31
After Histogram match			
Std. dev.(scene4)	37.60	34.70	75.69
Mean (scene4)	69.09	57.44	128.80

The histogram of the difference image with an ample amount of pixels in the tails clearly indicates changes.

Supervised Classification of LISS IV image was done after performing all corrections in the image. The image was classified with classes viz; river/water bodies, urban, agriculture land, agriculture crop, forest/trees, sand and barren land.. Area Statistics of land use land cover classes of supervised classification of LISS IV at 5.6m are given in Table 5.

Table 5 Area Statistics of land use land cover classes of supervised classification of LISS IV

Class	LISS IV @ 5.6 mtr res	LISS IV @ 5.6 mtr res
Particular	Area(ha)	Area%
Barren land	14848	2.924
River / Water bodies	10384	2.045
Settlement	4742.6	0.935
Agriculture land	235724	46.423
Sand	9980.61	1.966
Agriculture crop	133385	26.268
Forest / Trees	98706.6	19.439
Total	507770.81	100%

Scatter Plot of Supervised Image of LISS IV

To compare the supervised classification results for LISS IV with scatter plot is presented, we have found for convince to perform them on the same number of Landcover classes (i.e.7 classes).The method has been implemented on the samples of multispectral images So when we draw a scatter plot for supervised classification it shows different classes as per the classification and the drawn plot shows how much the classification is correct. The closeness of different pixels is the sign of good classification and due to same reflectance of various pixels it may sometimes be overlaps and if scatter plot shows in scatter form and the pixels looks far away from one another it shows the low image quality or imperfect classification. Figure 1,2,3 shows the results of the SCATTER module, which represents the overall layout of reflectance values in the image. The signature plot shows the approximate range reflectance values for each class signature. The most problematic classes are the two overlapping ovals in the bottom of the plot. These represent build-up (the large green oval) and roads (the small blue oval), which are difficult to separate because of reasons mentioned above.

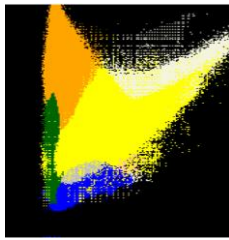


Figure 1 Scatter plot for Supervised classification of LISS IV for band 2 and 3

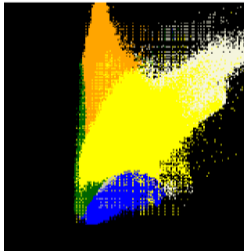


Figure 2 Scatter plot for Supervised classification of LISS IV for band 1 and 3

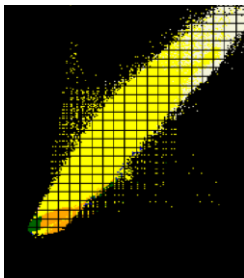


Figure 3 Scatter plot for Supervised classification of LISS IV for band 1 and 2

Accuracy Assessment

In this study, the number of reference points used for the accuracy assessment of classification were 200, most of which were taken from the field visit, and the remaining from the topographic maps. It can be shown in the table 4.20 that the overall classification accuracy for all three satellite images for both classification. Extraction of agriculture crops may be extracted more accurately. Very small indistinctive scattered urbanized areas couldn't be detected in unsupervised classification, and they were omitted from table of classification because of their insignificant effects. The accuracy of the land cover classification from supervised and unsupervised techniques were evaluated and presented as an error or confusion in the form of matrix table.

Evaluating the accuracy of the classification was done here by applying the accuracy assessment methods. A set of reference pixels is usually used where points on the classified image for which actual data are (or will be) known. The relationship between these two compared information is commonly summarized in an error matrix in appendix. The number of rows and columns in the error matrix should be equal to the number of categories whose classification accuracy is being assessed (R.M.Lillisand & R.M.Kiefer, 2000)

The accuracy assessment generated from the supervised classification technique showed an overall classification accuracy was 95% with Kappa Statistic of 0.92 for LISS IV, which indicates a good agreement between thematic maps generated from image and the reference data. This amount of agreement is generally considered a good statistical return.

For instance, in unsupervised classification technique, firstly we were performed unsupervised classification and then we find out the accuracy for LANDSAT 8 and AWiFS at their different resolutions which were reduced at 10% from their original resolution.

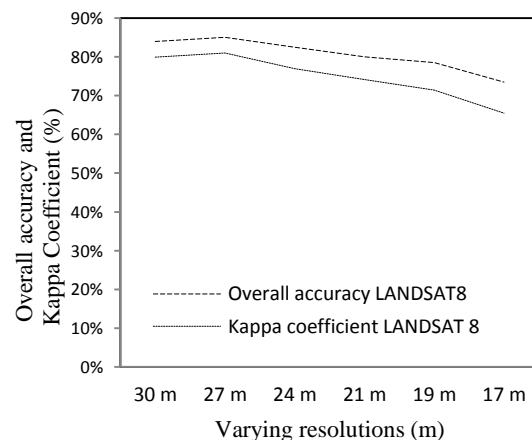


Figure 4 Change in Overall accuracy and Kappa Coefficient with respect to decrease in resolution for LANDSAT8

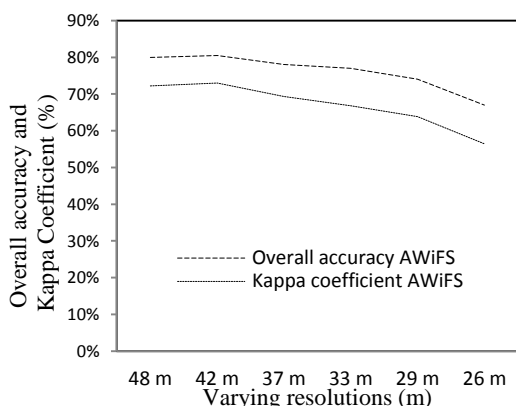


Figure 5 Change in Overall accuracy and Kappa Coefficient with respect to decrease in resolution for AWiFS

Likewise LANDSAT 8 with 30 m resolution, and a reduced resolution at 10% at 30, 28, 24, 22, 19, 17m respectively. At these resolutions the overall accuracy was 84%, 85%, 82.5%, 80%, 78.5%, 73.5% with 0.79, 0.80, 0.76, 0.74, 0.71, 0.65 Kappa statistic respectively. Accordingly, AWiFS with 48m resolution and a reduced resolution 10% the over all accuracy was 80%, 80.5%, 78%, 77%, 74%, 67% with 0.72, 0.72, 0.69, 0.66, 0.63, 0.37 Kappa statistic respectively.

Table 6 presents the comparison of overall accuracy and Kappa statistic between the supervised and unsupervised classification of LISS IV, LANDSAT 8 and AWiFS at their original resolutions. It shows that the accuracy was increase at the reduction of 10% resolution, but after that the accuracy was get increased. The statistical approach of the accuracy assessment consists of different multivariate statistical analysis. A used measure is KAPPA (Cohen, 1960). KAPPA is designed to compare results from different regions or different classifications.

Table 6 Comparison of accuracy for supervised and unsupervised classification

Method	Overall accuracy	Kappa
Supervised (LISS IV)	95 %	0.9256
Unsupervised (LANDSAT 8)	84 %	0.7990
Unsupervised (AWiFS)	80 %	0.7219

REFERENCES

Althausen, J. D. (2002). What remote sensing system should be used to collect the data? In *Manual of Geospatial Science and Technology*, ed. J. D. Bossler, J. R. Jensen, R. B. McMaster and C. Rizos, 276. New York: Taylor & Francis.

Aplin, P., Atkinson, P. M., & Curran, P. J. (1999). Per-field classification of land use using the forthcoming very fine spatial resolution satellite sensors: problems and potential solutions. *Advances in Remote Sensing and GIS Analysis*

Atkinson, P. M. (2012). Downscaling in remote sensing. *Int. J. Appl. Earth Observ. Geoinf.* <http://dx.doi.org/10.1016/j.jag.2012.04.012>.

Barnsley, M. J. (1999). Digital remote sensing data and their characteristics. I. In *Geographical Information Systems: Principles, Techniques, Applications, and Management*, 2nd New York: John Wiley & Sons.

In Table 6 the Kappa statistic illustrated that kappa accuracy and coefficient was better by using supervised compare to unsupervised, where kappa coefficient is slightly higher for supervised with 0.92 and unsupervised with 0.79 and 0.72.

The figures 4 and 5 shows the graphs between overall accuracy and resolutions for LANDSAT 8 and AWiFS shows the increase and decrease in the accuracy at different resolutions.

Discussion

This discussion summaries the extent of change in spatial resolution that can be accepted to produce classification maps with high accuracies. Overall the supervised classifications are the most appropriate to create land use land cover map of the study area. As the pixel merge or disaggregate, their spatial characteristic also change. Such as sand and settlement.

Here KAPPA (Cohen, 1960) is used very successfully to validate the results. It shows that that accuracy of a resolution is increased by 10 % reduction. Any further increase will result in decreasing accuracy. Hence downscaling limits of spatial resolution for AWiFS and LANDSAT 8 as compared to LISS IV is up to 10% of their original resolutions, any resampling beyond this will lead to precious information being lost.

CONCLUSION

This paper has reviewed the factors and techniques to be considered when understanding the limits of resampling of remotely sensed data. Here in this study again it is established that the supervised classification is more appropriate technique to classify an image or satellite data. Unsupervised classification gives lesser results due to merging of pixel reflectance. The techniques and considerations discussed here need to be applied over all spatial data. Techniques developed for other spatial data need to be tested for use with remotely sensed data. Scatter plot justify the result of supervised classification. Hence we concluded that we can reduce the resolution of any LANDSAT8 and AWiFS at only 10% from their original resolution with acceptable accuracy.

Acknowledgement : The auther acknowledges the funding and support received from the Sam Higginbottom Institute of Agriculture Science and Technology, Deemed University towards presentation of this research paper.

- Blaschke, T. (2010). Object based image analysis for remote sensing *Journal Photogrammetric Remote Sensing*, 65(2).
- C.P.Smits, Dellepiane, S. G., & Schowengerdt, R. A. (1999). Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *Int J Remote Sensing*, 20, 1461.
- Cohen, J. (1960). A coefficient of agreement for nominal scales, *Educ. Psychol. Measurement* 20(1):37-46., 20(1), 37 46.
- Congalton, G. R. (1991). A review of assessing the accuracy of classifications of remotely sensed Data. *Remote Sensing of Environment*, 37(1), 35 46.
- Congalton, R. G., & Green, K. (2008). Assessing the Accuracy of Remotely Sensed Data: Principles and Practice, 2nd ed, 183. Boca Raton, FL: CRC Press/Taylor & Francis Group.
- Curran, P. J., & Williamson, H. D. (1985). The accuracy of ground data used in remote sensing investigations. *International Journal of Remote Sensing*, 6, 1737 1651.
- Estes, J. E., & Loveland., T. R. (1999). Characteristics, sources, and management of remotely-sensed data. In *Geographical Information Systems: Principles, Techniques, Applications, and Management*, 2nd ed., ed. P. Longley, M. Goodchild, D. J. Maguire, and D. W. Rhind, 667. New York: John Wiley & Sons.
- Foody, M. G. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(185).
- Franklin, S. E. (2002). Evidential reasoning with Landsat TM, DEM and GIS data for land cover classification in support of grizzly bear habitat mapping. *International Journal of Remote Sensing*, 23, 4633.
- G.M.Foody. (1996). Foody, G. M. 1996. Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. *Int J Remote Sensing* 17:1317. *International Journal of Remote Sensing* 17(17), 1317.
- Gallego, F. J. (2004). Remote sensing and land cover area estimation. *International Journal of Remote Sensing*, 25, 3019.
- Ghimire, B., Rogan, J., & Miller, J. (2010). Contextual land cover classification: incorporating spatial dependence in land cover classification models using random forests and the Getis statistic. *Remote Sensing Lett* 1(45).
- Gong,P, & Howarth, P. J. (1992). Frequency-based contextual classification and gray-level vector reduction for land-use identification. *Photogramm Eng Remote Sens* 58(423).
- Kontoes, C., G.Wilkinson, A.Burril, S.Goffredo, & Megier, J. (1993). An experimental System from the Integration of GIS Data in Knowledge-base Image Analysis for Remote Sensing of Agriculture. *International Journal of Geographical Information Systems*, 7(3), 247 262.
- Lefsky, M. A., & Cohen, W. B. (2003). Selection of remotely sensed data. In *Remote Sensing of Forest Environments: Concepts and Case Studies*, ed. M. A. Wulder and S. E. Franklin, 13. Boston, MA: Kluwer Academic.
- Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28, 823.
- Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86(554).
- R.M.Lillisand, & R.M.Kiefer. (2000). Remote Sensing and Image Interpretation. 4th Edition, New York: John Wiley and Sons 724 p.
- San Miguel-Ayanz, J., & Biging, G. S. (1997). Comparison of single-stage and multi-stage classification approaches for cover type mapping with TM and SPOT data. *Remote Sensing of Environment* 59 92.
- Stuckens, J., Coppin, P. R., & Bauer, M. E. (2000). Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment*, 71, 282.