Chapter 13
Geospatial Technologies for Semiautomated Baseline Database Generation for Large-Scale Land Resource Inventory

S. Chattaraj, S. K. Singh, S. K. Ray, V. Ramamurthy, A. Daripa, and G. P. Obi Reddy

Abstract The goal of land resource inventory is to enable the lab-to-land transfer of agro-technology on a sustainable basis through identification of homogeneous soil management units. The identification of homogeneous landscape ecological unit (LEU) boundaries for soil mapping through conventional methods is time-consuming and laborious. Hence, it is necessary to develop a semiautomated geospatial framework for delivering reliable soil resource information to the users on time. In the present chapter, the approach for semiautomation in landform delineation using high-resolution IRS Cartosat-1 and LISS-IV data was discussed. Cartosat-1 stereopair data are processed to generate the digital terrain model (DTM) of 10 m spatial resolution. The digital terrain analysis was carried out to generate contour, drainage, slope, and hillshade for landform delineation in two distinct terrain conditions. *Object-based slope classification algorithm* is developed by following USDA-NRCS slope class thresholds to hasten the process of landform identification. The land use/land cover (LULC) map of the area is generated based on the *rabi* season data of Cartosat-1 merged LISS-IV (2.5 m) as well as high-resolution (0.5 m) public domain imagery at the backend so as to get the reliable land use boundary at cadastral level through feature optimization algorithm in eCognition software using near-infrared (NIR) and Normalized Difference Vegetation Index (NDVI) data. The integration of three secondary layers, i.e., landform, slope, and LULC, are achieved through the *hierarchical object-based segmentation algorithm* to develop landscape ecological unit (LEU) map. The logical automation algorithm developed at each stage assists in optimizing sampling intensity, which leads to a considerable saving of man power, labor, cost, and time.

S. Chattaraj (✉) · S. K. Singh · A. Daripa · G. P. O. Reddy
ICAR-National Bureau of Soil Survey & Land Use Planning, Nagpur, India

S. K. Ray
ICAR-National Bureau of Soil Survey & Land Use Planning, Regional Centre, Jorhat, India

V. Ramamurthy
ICAR-National Bureau of Soil Survey & Land Use Planning, Regional Centre, Bengaluru, India

G. P. O. Reddy, S. K. Singh (eds.), *Geospatial Technologies in Land Resources Mapping, Monitoring and Management*, Geotechnologies and the Environment 21,
https://doi.org/10.1007/978-3-319-78711-4_13

Keywords Baseline database · Digital terrain modeling · Object-based digital terrain classification

13.1 Introduction

The application of satellite remote sensing data products for small- and medium-scale soil mapping is widely accepted (Soil Survey Division Staff 1995), but its use in large-scale soil mapping is restricted till date not only for the limited availability of high-resolution data but also due to lack of understanding the too much details present in the high-resolution data. Large-scale soil mapping is mostly done following conventional methods that are time-consuming and expensive and have low repetitive value especially in difficult and inaccessible terrain. However, with the recent advances in satellite data processing and analysis, availability of high-resolution satellite data like IRS-R2 LISS-IV data (5.8 m) and Cartosat-1 can now be utilized well for large-scale soil mapping. Srivastava and Saxena (2004) discussed the technique of large-scale soil mapping (1:12,500 scale) in a basaltic terrain with a PLU approach and differentiated soil types using topographic information available in the Survey of India toposheet and LULC information from IRS-1C PAN merged data of two seasons (*kharif* and *rabi*). Similar exercise was also carried out by Nagaraju et al. 2014 using Cartosat-1 and IRS-R2 LISS-IV data. However, the traditional way of landform extraction by an interpreter through the topographic maps, aerial photograph, or satellite imagery followed by ground truthing is accepted and appropriate. In one hand, the traditional way is relatively time-consuming, and the results are subjected to interpreter's biasness, and also not reproducible. On the other hand, pixel-based digital landform and LULC mapping of high-resolution data results in noise at the larger scale owing to the presence of minute details in the data. That's why object-based image analysis (OBIA) comprising of image objects, i.e., groups of pixels that are similar to one another based on a measure of spectral properties (i.e., color), size, shape, and texture, as well as context from a neighborhood surrounding the pixels, has gained increasing attention in landform and LULC research from the last decade (Drăguţ and Blaschk 2006; Eisank et al. 2011; d'Oleire-Oltmanns et al. 2013; Chattaraj et al. 2017).

The goal of soil/land resource inventory is to identify and delineate homogeneous soil patterns formed within a complex, heterogeneous soil-forming environment to enable the lab-to-land transfer of agro-technology in a sustainable basis. Successful mapping of soil resources on large scale is highly dependent on precise information of landforms (the testimony of past climate as well as topographic factors), slope, and LULC (the indicators of present climate and management conditions). It is realized that the identification and delineation of homogeneous landscape ecological unit boundaries for soil mapping through conventional methods is time-consuming. Further, the experienced man power to carry out soil survey is also declining rapidly. Hence, it is necessary to develop a semiautomated geospatial framework so that reliable information on soil resource is delivered to the users in time. However, the key to successful knowledge-based modeling depends on how effectively the implicit knowledge understanding on the target objects is transformed into explicit decision rules (Cheng and Han 2016). The present chapter discusses the approach for semiautomation in slope, landforms, and LULC classification for generating the Hierarchical Landscape Ecological Unit (LEU) segmentation model using high-resolution Cartosat-1 and IRS-R2 LISS-IV data.

13.2 Methodology Framework

The overall methodology flow diagram is presented in Fig. 13.1.

The steps involved are:

First step is the generation of digital terrain model (DTM) particularly in the undulating terrain using Cartosat-1 data of 1 m resolution. The primary terrain attributes, namely, contour, drainage, hillshade, slope, and curvatures, are derived from DTM, which have been used as input layers for developing precise and quantified data on landforms (Fig. 13.1a).

Second step is the generation of LULC maps using IRS-R2 LISS-IV data of 5.8 m resolution. Derived LULC map superimposed on landform and slope map to develop Landscape Ecological Unit (LEU) map, the base map of soil/land resource inventory at larger scale. LEUs are defined by a set of symbol D2s, D4w1, U4w4, D2d, etc., consisting of letters and numerals. First letter in capital is the landform, second numeral is slope class, and third letter and numeral is LULC (Fig. 13.1b).

Third and final step is the **e**xtensive traversing and ground truth collection through mini-pits and profile investigations in well-defined strips representing assemblage of LEUs. Establishing phases of soil series and developing soil-landform relationship are the next part of third step. However, the third step is beyond the scope of the present paper. Hence, the development of object-based models for delineation of LEUs is the prime focus of the chapter (Fig. 13.1c).

13.2.1 Semiautomated Modeling

13.2.1.1 Digital Terrain Modeling

Cartosat-1 stereo pair data were processed to generate the digital terrain model (DTM) of 10 m spatial resolution using rigorous math model (Toutin's Model). In the model, OrthoEngine of Geomatica version 14.0 is used to generate DTM following the sequence of steps, namely, projection setup, sensor data reading, collection of GCPs and tie points, block adjustment, model computation (Satellite Math Model), epipolar image generation, and digital surface model (DSM)

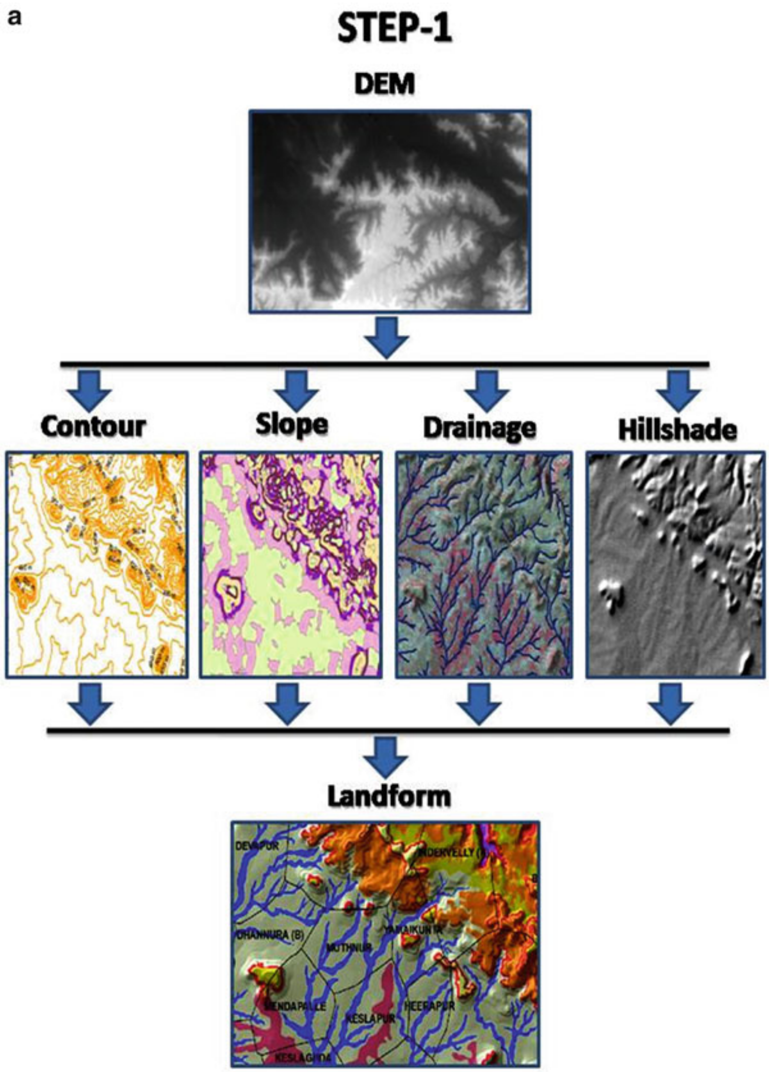

Fig. 13.1 (**a**)–(**c**) Steps in large-scale soil/land resource inventory

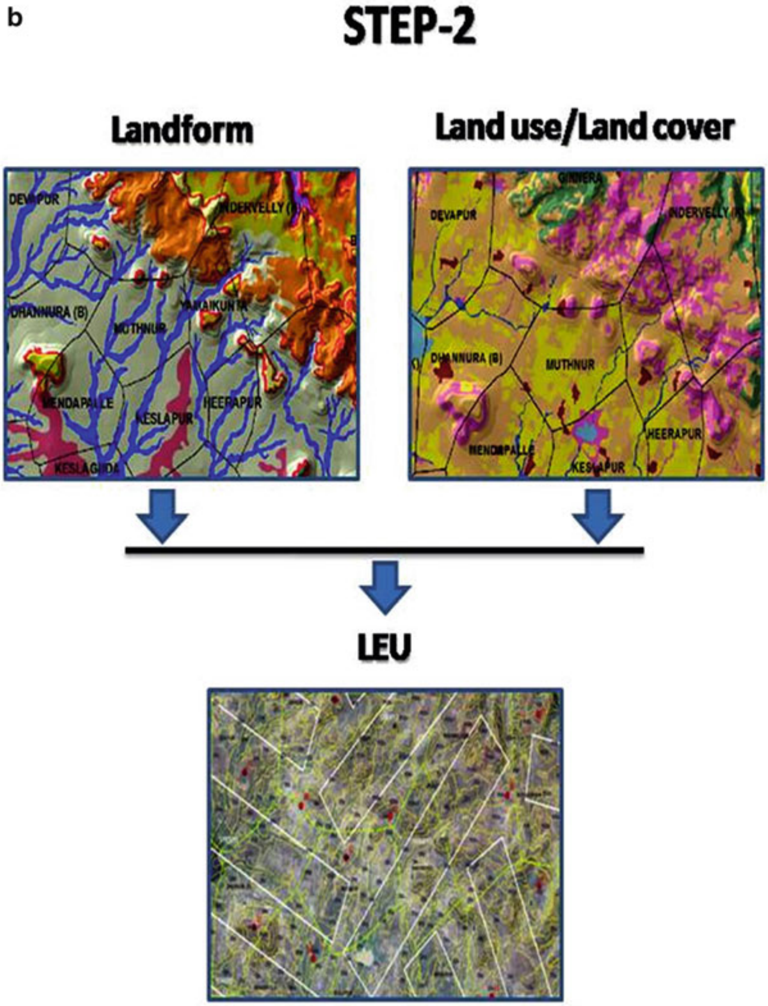

Fig. 13.1 (continued)

extraction. Balancing algorithm is applied to obtain the seamless mosaic DSM height. Filtering is done to convert bare earth model, DSM to DTM. Editing is done to smooth out the irregularities and create a quality output. RMSE statistics report is also generated to evaluate the accuracy of the DTM output (Fig. 13.2).

Further, DTM is subjected to a series of hydro-enforcement process including reconditioning, sinks and pit removal, flat and level water bodies, flat and level bank to bank, and gradient smoothening by DAT/EM and Arc Hydro tool, etc. This is

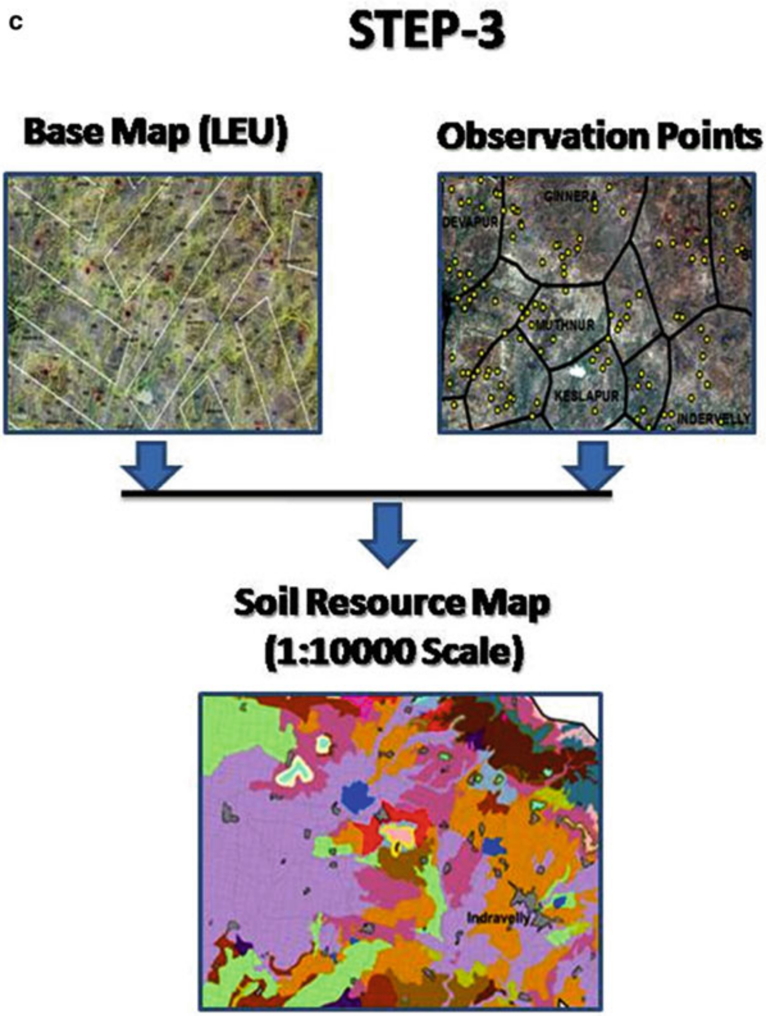

Fig. 13.1 (continued)

essentially needed to enrich the quality of the hydrological output such as slope, contour, and drainage (Romstad and Etzelmuller 2012). This altogether needed to improve the accuracy of landform mapping (Fig. 13.3).

13.2.1.2 Object-Based Digital Terrain Classification Models

The pixel-based classification procedure analyzes only the spectral properties, but the spatial or contextual information is lacking. Pixel-based methods applied to high-resolution images give a "salt and pepper" effect that contribute to the inaccuracy of

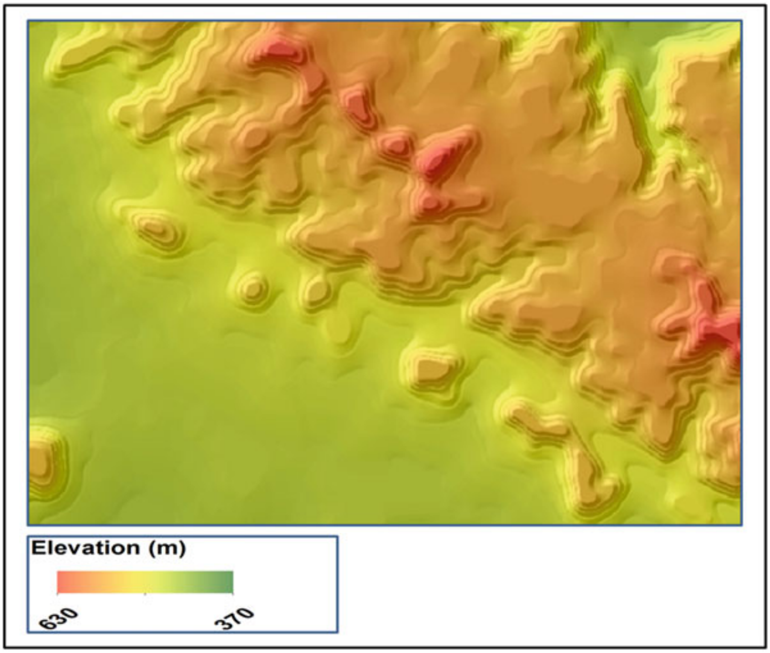

Fig. 13.2 DTM of 10 m resolution for a part of Indervelly block, Adilabad district, Telangana

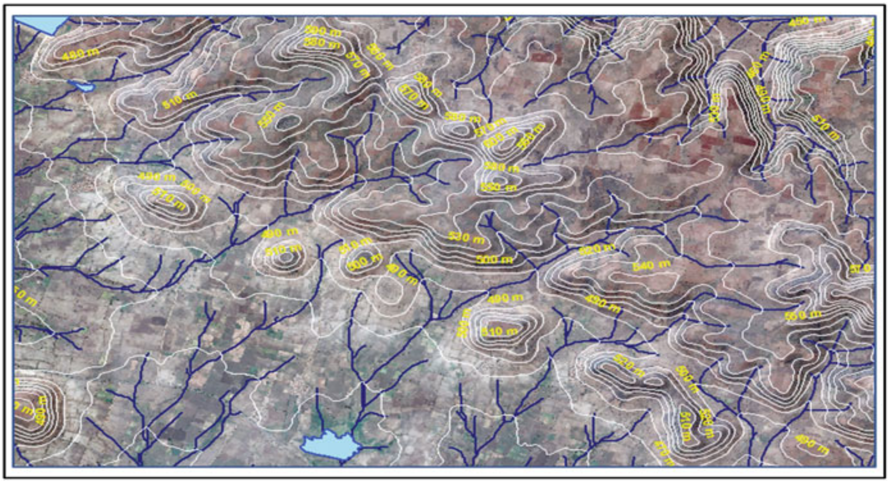

Fig. 13.3 Contour (10 m) and auto-drainage derived from DTM for a part of Indervelly block, Adilabad district, Telangana

the classification. For decades, geographic information system (GIS) specialists have theorized about the possibility of developing a fully or semiautomated classification procedure that would be an improvement over pixel-based procedures. The object-based modeling by taking into consideration the spectral and spatial/contextual properties of pixels and segmentation process with interactive learning algorithm promises to be more accurate than the pixel-based methods (Camargo et al. 2011). The following object-based semiautomated models are developed in the study.

13.2.1.3 Slope Classification Model

The raster slope layer output of DTM is taken as input in the object-based image analysis in the environment of eCognition® software. The slope layer was subjected to chessboard segmentation. Nine slope classes are created following the USDA-NRCS slope class threshold criteria. The criteria is fitted as fuzzy instead of hard rule using the less than and greater than "s-curve" membership function so as to get closer to the natural slope boundary. Morphology and contextual filters are applied to generate smooth slope class zones (Fig. 13.4).

13.2.1.4 Landform Classification Model

Case Study-I

The terrain attributes derived through digital terrain analysis of DTM layer, i.e., contour, drainage, slope, and curvature, are treated as input for landform delineation. The landform classification process is hastened taking into consideration the slope class zone, hillshade, contour, and auto-drainage pattern along with legacy physiography unit of 1:250k. The table (Table 13.1) below illustrates an example of logical rule set used for different landform units occurring in the Indervelly block of Telangana state (Fig. 13.5).

Case Study-II

The similar kind of exercised is carried out in the northeastern hilly region of Ri-Bhoi district, Meghalaya, where the objects resulting from segmentation are partitioned into subdomains based on thresholds given by the mean values of elevation and standard deviation of elevation, respectively, following the modeling approach given by Drăguț and Eisank (2012). The layer variable thresholds are modified as per the local condition. The rule set window (Fig. 13.6) and resultant landform (Fig. 13.7) are presented below.

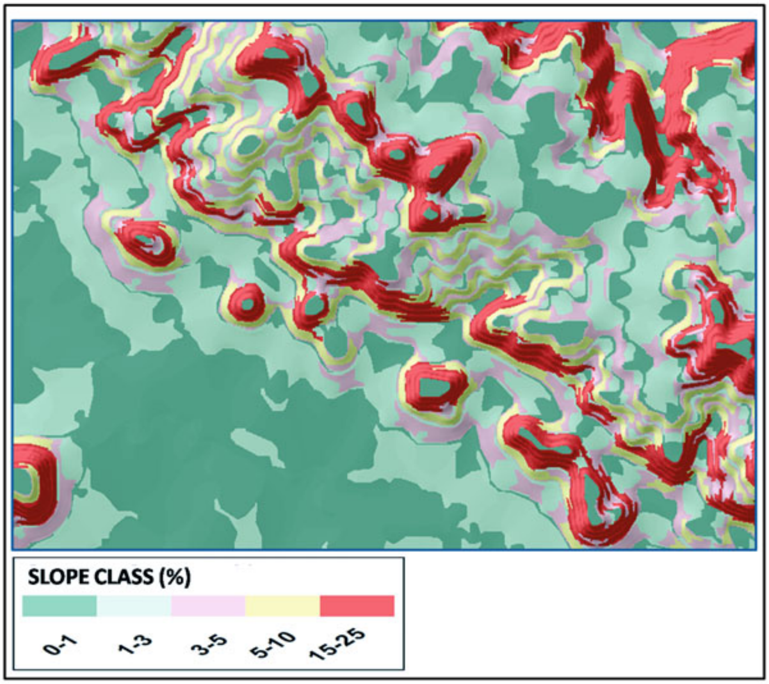

Fig. 13.4 Slope class zone derived from DTM for a part of Indervelly block, Adilabad district, Telangana

Table 13.1 Logical rule set used for different landforms

Landform	Logical ruleset condition
1. Undissected plateau	Slope range, 0–5%
	Relative boarder to escarpment, >80%
	Existence of drainage = false
	Relative topographic position = upper
2. Pediment	Side slope of plateau/upland
	Slope range, >1 to <15%
	Profile curvature = convex
	Presence of erosive features
3. Valley	Existence of drainage = true
	V-shaped contour with decreasing elevation gradient
	Profile curvature = concave
	Relative topographic position = lower

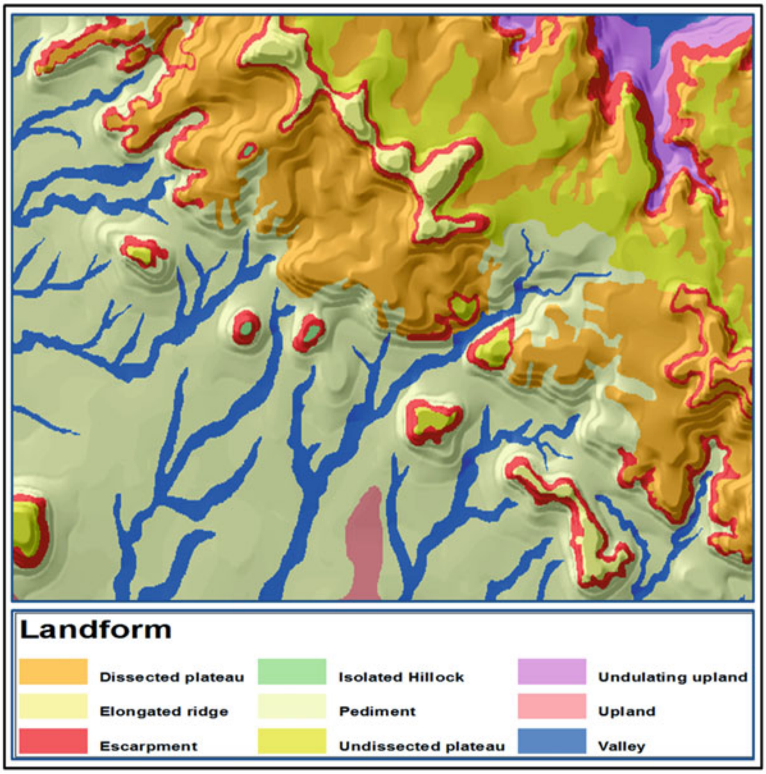

Fig. 13.5 Landform map on 1:10000 scale derived from DTM as a part of Indervelly block, Adilabad district, Telangana

Fig. 13.6 Rule set algorithm for delineating landform in Ri-Bhoi district, Meghalaya

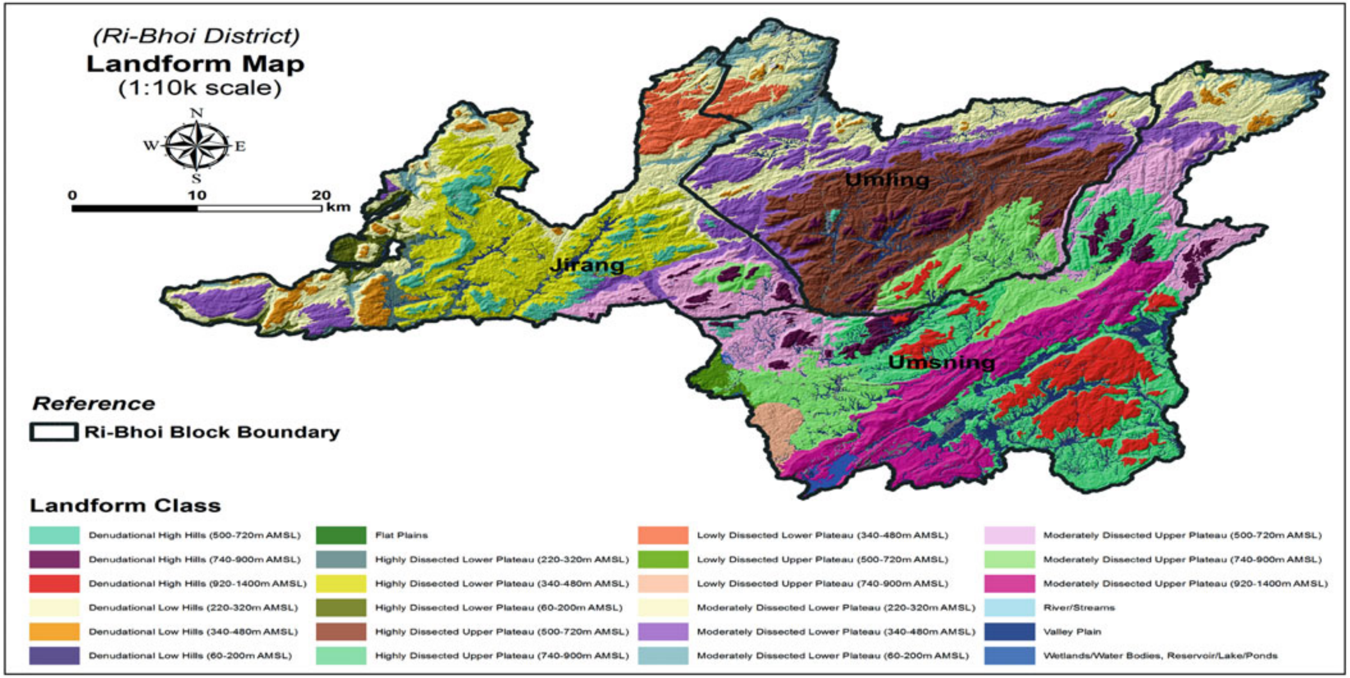

Fig. 13.7 Landform map derived from DTM for Ri-Bhoi District of Meghalaya state

13.3 Accuracy Assessment

To determine the accuracy of the OBIA-based digital landform output, independent reference landform is needed to be delineated based on visual interpretation in the study area (i.e., watershed). This will help in evaluating the accuracy of landform modeling on a larger scale, i.e., watershed level. The reference landform delineation is completed manually by using the background information of IRS-P6 LISS-IV imagery and DTM-based output of slope, contour and drainage pattern on a shaded relief layer. Finally, the accuracy is assessed based on the following three measures (d'Oleire-Oltmanns et al. 2013):

1. User's accuracy (UA), the percentage of correctly classified area from the total classified area
2. Producer's accuracy (PA), the percentage of correctly classified area from the total reference
3. Detection rate, the percentage of reference data that have been detected by the classification (also including partial detection)

13.3.1 An Example of Accuracy Assessment

Visual illustration of classified and reference landforms of Tandulwani watershed of Katol tehsil, Maharashtra (Chattaraj et al. 2017) are illustrated in Fig. 13.8. The top image of each landform section illustrates the classification results (solid color fill) as well as the reference polygon (black outlines) overlayed with shaded relief draped as a base layer. The three color-coded insets below in each section display examples of good matches (green box), as well as underestimations (yellow box) and overestimations (red box). Similar approach of illustration was also documented by d'Oleire-Oltmanns et al. (2013). Visual comparison of different landform segments at the two chosen scales illustrates the GEOBIA modeled landforms hold good even at larger scale. The values of classification accuracies and their graphical representation are given in Table 13.2 and Fig. 13.9, respectively.

For each individual landform, the UA, PA, and detection rate were calculated. A visual comparison of the GEOBIA modeled landform map to the visually interpreted reference landforms is shown in Fig. 13.10 revealing a highly satisfactory areal extent matching of the modeled output. Similarly, the quantitative assessment result of the modeling performance documents a high-level accuracy as indicated by the excellent kappa (0.91) and overall accuracy (92.8%) statistics. The UA and PA for all the landforms have achieved more than 90% accuracy except for lower alluvial plain and upper pediment (Table 13.2). It is noteworthy that the detection rate for all the landform units are around 100%. This indicates the sound performance of the

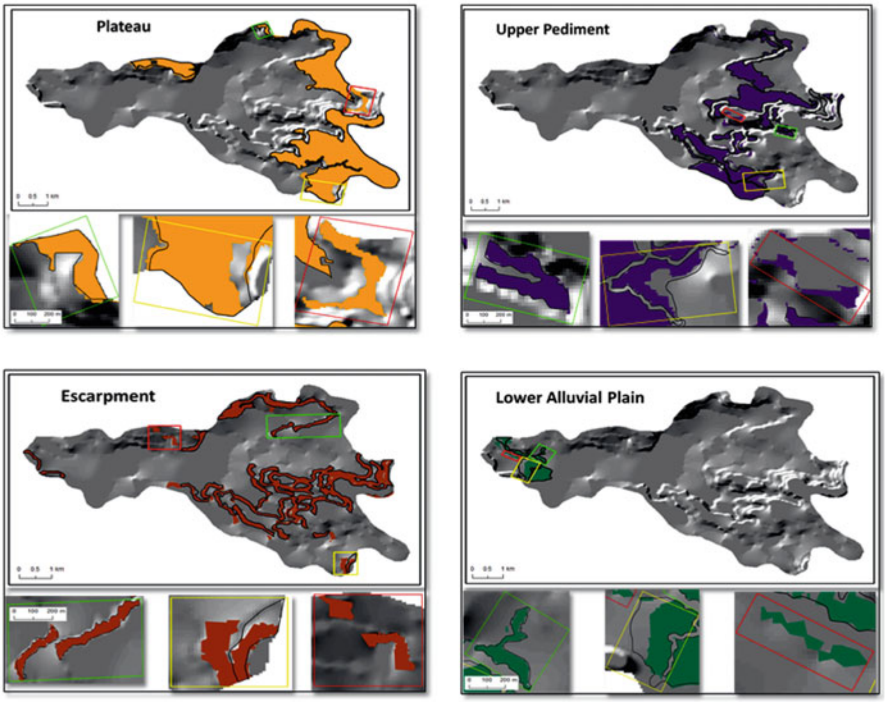

Fig. 13.8 Classification results for major landform units (solid color fill) and the corresponding reference landform units (black outlines) are illustrated. A shaded relief layer is displayed in the background. Three color-coded insets show examples of (green) good matches between classification and reference, (yellow) underestimations of reference, and (red) overestimations of reference

knowledge-based modeling to capture the existence of the landform units occurring in the watershed including partial detection.

13.4 Object-Based Land Use/Land Cover Classification Model

The LULC map was prepared based on the current *rabi* season data of Cartosat-1 merged LISS-IV (2.5 m) as well as high-resolution (0.5 m) public domain imagery at the backend so as to get the reliable land use boundary at cadastral level. The delineation of subclasses, viz., single- and double-cropped areas within the agriculture zone, was done using novel *LULC subclass classification algorithm* (Fig. 13.11). The merged data was segmented into spectrally homogeneous region using multiresolution segmentation algorithm. The optimum scale parameter for segmentation of the layer was achieved through estimation of scale parameter (ESP)

Table 13.2 Accuracy statistics of the GEOBIA modeled landforms in relation to reference landforms

Accuracy measures	Calculations	Isolated hillock	Plateau	Escarpment	Upper pediment	Lower pediment	Upper alluvial plain	Lower alluvial plain	Channel
User's accuracy	(Overlap area/classified area) *100	94.86	93.43	86.81	92.35	96.48	93.16	87.44	90.37
Producer's accuracy	(Overlap area/reference)*100	97.24	96.9	91.1	88.08	91.01	98.44	79.97	97.03
Detection rate	(Amount of classified refrence/ total reference)*100	100	100	100	96.47	98.2	99.4	97.3	100

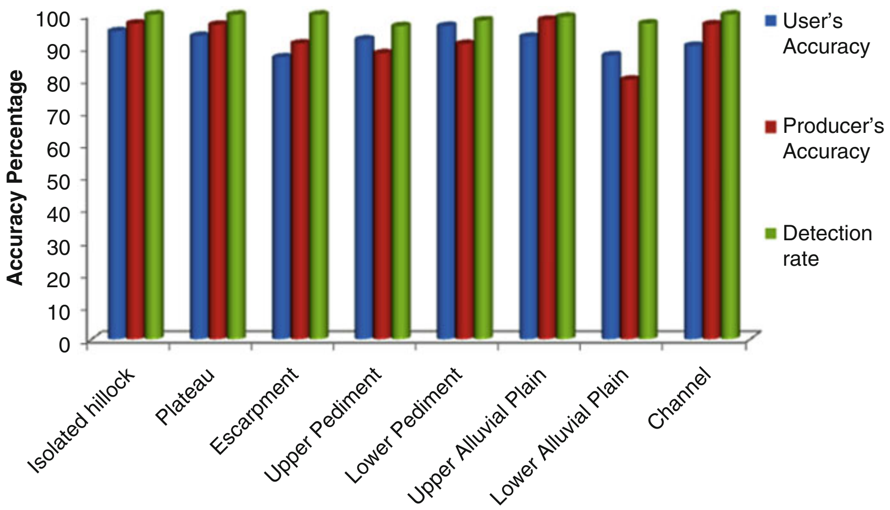

Fig. 13.9 Graphical representation of classification accuracy report across different landform units in the watershed

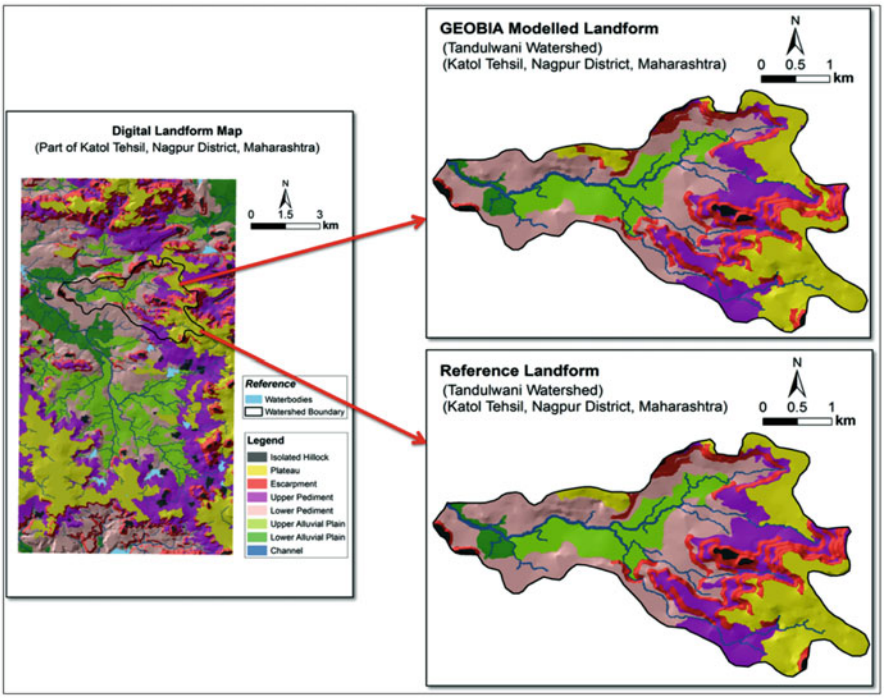

Fig. 13.10 Landform map of the study area as well as the modeled and reference landform map of the Tandulwani watershed as validation site

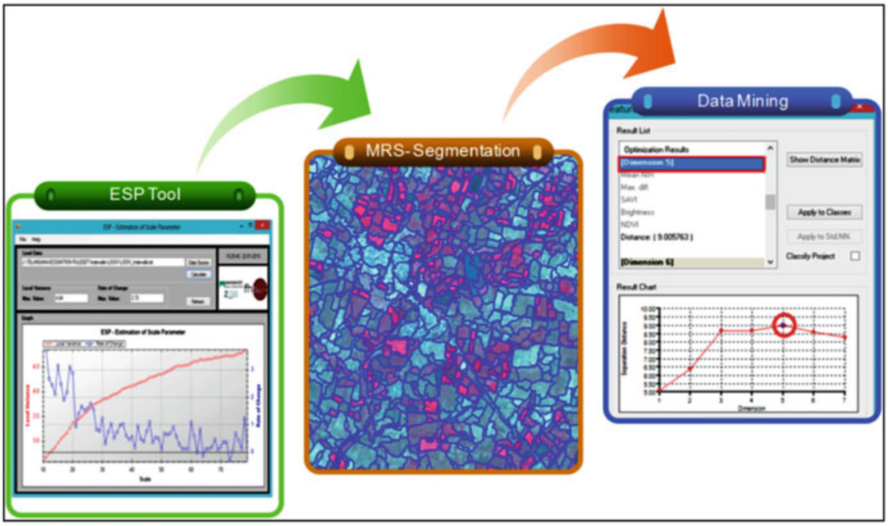

Fig. 13.11 LULC subclass classification model workflow

analysis tool. The point of interest lies where the local variance and rate of change are minimum in the graphical output. The data mining technique, i.e., feature space optimization, was applied to extract the double-cropped area based on certain number of layer variables and vegetation indices combination as obtained through the maximum separation distance. Following such scheme LULC map for a part of Indervelly block is given in Fig. 13.12.

13.5 Hierarchical Landscape Ecological Unit (LEU) Model

The integration of three secondary layers, i.e., landform, slope, and land use, was achieved through the *hierarchical object-based segmentation algorithm* taking into consideration the area, morphology of the landform units, and its relation with the neighbor objects to develop landscape ecological unit (LEU) map. The segmentation was accomplished in three levels:

1. Level-I: First level segmentation was done based on the landform layer.
2. Level-II: This segmentation was run within each of the first level segment based on fuzzy threshold-based slope class. Second level intermediate output gave rise to landform-slope unit.
3. Level-III: The landform-slope segments of second level were further subdivided into landform-slope-land use unit, i.e., LEU, by incorporating the land use factor. The logical condition used to incorporate the land use factor is that the minimum

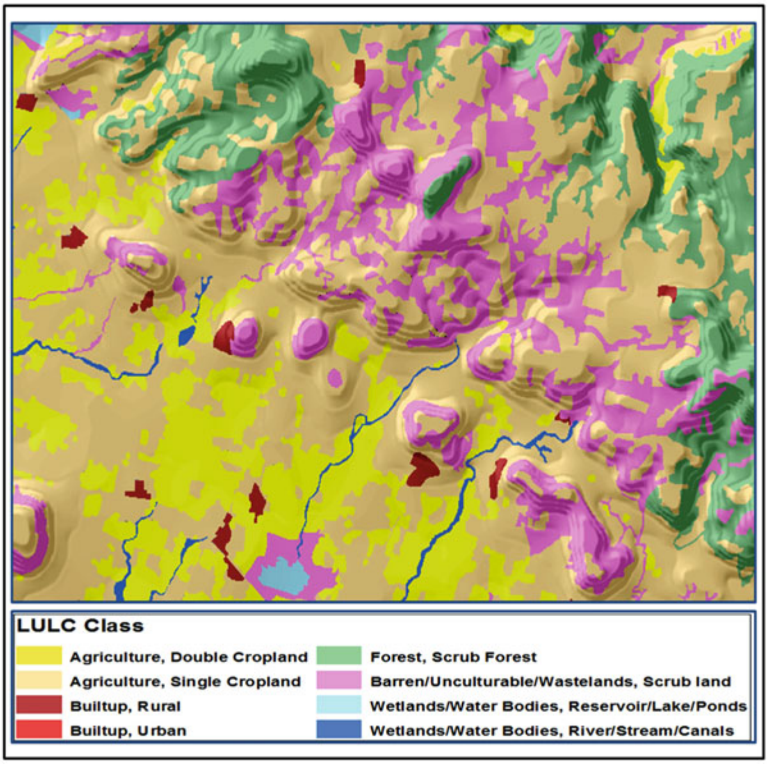

Fig. 13.12 Land use/land cover map on 1:10000 scale for a part of Indervelly block, Adilabad district, Telangana

overlap with the thematic polygon, i.e., level-II segment, will be more than or equal to 60%. The criteria ensure the continuity of LEU zone vis-à-vis soil boundary by ignoring negligible change in land use. Figure 13.13 explains the steps involved in the delineation of LEU.

13.6 Base Map in LRI Project

This LEU map has been used as base for developing soil-landform relationship for mapping soils on 1:10000 scales. Transacts were demarcated in GIS-based geo-database framework by assimilating the legacy data of 1:250 k scale and expert knowledge as shown in Fig. 13.14.

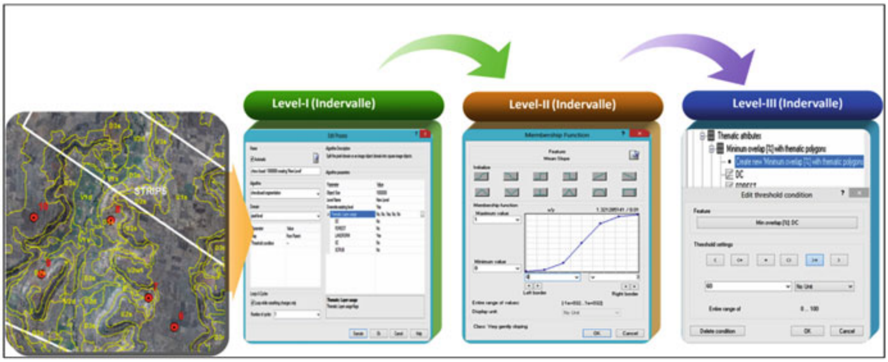

Fig. 13.13 Hierarchical object-based segmentation algorithm process for generating LEU maps

Fig. 13.14 Base map on 1:10000 scales for a part of Indervelly block, Adilabad district, Telangana

13.7 Conclusions

The logical automation algorithm developed at each stage results in considerable reduction in time for base map preparation. This will assist in optimizing sampling intensity, which leads to a considerable saving of man power, labor, cost, and most importantly the time. Finally, a hierarchical geo-database structure having unified schema is proposed for deploying in the *National Soil Geo-portal* to disseminate the information in a user-friendly way.

References

Camargo FC, Almeida T, Florenzano C, Heipke R, Feitosa, Costa G (2011) ASTER/Terra imagery and a multilevel semantic network for semi-automated classification of landforms in a subtropical area. Photogramm Eng Remote Sens 77:619–629

Chattaraj S, Srivastava R, Barthwal AK, Giri JD, Mohekar DS, Reddy GPO, Daripa A, Chatterji S, Singh SK (2017) Semi-automated object-based landform classification modelling in a part of the Deccan Plateau of central India. Int J Remote Sens 38(17):4855–4867

Cheng G, Han J (2016) A survey on object detection in optical remote sensing images. ISPRS J Photogramm Remote Sens 117:11–28

d'Oleire-Oltmanns S, Eisank C, Drăguţ L, Blaschke T (2013) An object-based workflow to extract landforms at multiple scales from two distinct data types. IEEE Geosci Remote Sens Lett 10(4):947–951

Drăguţ L, Blaschke T (2006) Automated classification of landform elements using object-based image analysis. Geomorphology 81(3–4):330–344

Drăguţ L, Eisank C (2012) Automated object-based classification of topography from SRTM data. Geomorphology 141-142:21–33

Eisank C, Drăguţ L, Blaschke T (2011) A generic procedure for semantics-oriented landform classification using object-based image analysis. In: Geomorphology 2011 conference paper. Elsevier, Redlands, pp 125–128

Nagaraju MSS, Kumar N, Srivastava R, Das SN (2014) Cadastral level soil mapping in basaltic terrain using Cartosat-1-derived products. Int J Remote Sens 35(10):3764–3781

Romstad B, Etzelmuller B (2012) Mean-curvature watersheds: a simple method for segmentation of a digital elevation model into terrain units. Geomorphology 139/140:293–302

Soil Survey Division Staff (1995) Soil survey manual, USDA handbook no. 18, revised edn. Scientific Publishers, Jodhpur

Srivastava R, Saxena RK (2004) Techniques of large scale soil mapping in basaltic terrain using satellite remote sensing data. Int J Remote Sens 25(4):679–688