Proceedings of the *FOSS/GRASS Users Conference* − Bangkok, Thailand, 12−14 September 2004

Classification Analysis of LANDSAT Images of Mixed Coniferous and Deciduous Riparian Forest in Nature Conservation Zone Using GRASS/PostGIS Link

Mikhail Gromyko*, Alexander Shevlakov**

*Sikhote−Alin State Biosphere Reserve, Terney, Primorskii Krai, Russia, 692150,t.+7(42374)31368
Email:gromyko_m@mail.ru
**Free Software and GIS Consultant, Moscow, Russia, t.+7(916)1739952,
Email:alex.shevlakov@motivation.ru

Abstract

Several patterns of riparian forests in the Sikhote−Alin Biosphere Reserve in Northern Primoriye were classified from two LANDSAT images taken in two different seasons, with various tree composition from dark coniferous forest through mixed to deciduous forests. The seasonal changes in the reflectance of the forested terrain were studied to use one of the two images for the forest classification. The GRASS unsupervised classification was used to create several maps with different number of initial classes. The automated spatial RDBMS link was used to search for the correlations of the different classified patterns with ground information and to select combinations of parameters which have effects on sun rays reflectance.

1. Introduction

It has been in the field of interest for decades to apply the remote sensing techniques in forest classification. The coarse resolution grid images obtained in early years can be used to separate vegetation from un−forested lands, or to assess large−scale conditions that influence on the forest, while the higher resolution grids make possible yet more precise analysis of the images, including the species composition analysis [2,3].

In areas with complex topography reflectance of the forest cover should be taken in account, while for the plain areas the task of classification is more simple. Various techniques, *e.g.* , preprocessing the satellite bands, make possible to yet increase the accuracy of classification. The robust computations using the forest inventory data as well as remotely obtained data help to produce models with prediction of forest types.

2. Vegetation in the Study Area

The vegetation cover of the riparian valleys in the middle course of streams of the eastern slopes of the Sikhote−Alin varies much in structure and composition [1].

The forests are of deciduous, mixed and coniferous types. The following different forest types occur in the area of study:

1. The deciduous forest stands on narrow sand and pebble−stone banks of the streams with dominance of chosenia (*Chosenia arbutifolia*) are often a result of catastrophic floods. These chosenia stands are characteristic in their fast growth and mono−dominance in early maturity. Other tree species are present in older stands, such as poplar (*Populus maximoviczii*), Japanese elm (*Ulmus japonica*), Manchurian ash (*Fraxinus mandshurica*), Yeddo spruce (*Picea ajanensis*), Manchurian alder (*Alnus hirsuta*), mono maple (*Acer mono*) and the Korean pine (*Pinus koraiensis*).

2. The poplar stands are found mainly in broad valleys of large rivers and are located on higher terraces than the chosenia stands. The alder stands are found at the damped parts of the river valleys, their composition may also include broadleaf white birch (*Betula plathyphylla*), larch (*Larix dahurica*); the undergrowth is with other deciduous species, *e.g.* , Amur corktree (*Phellodendron amurense*), false spiraea (*Sorbaria sorbifolia*), bird cherry (*Padus asiatica*), buckthorn (*Ramnus dahurica*).

3. The deciduous forest stands with elm, maple and ash in dominance are mostly found on the first riparian terraces and estuaries oriented to the south. The dominant species are Japanese elm and cut−leaved elm (*U. japonica* , *U. laciniata*), Manchurian ash−tree (*F. mandshurica*) and mono maple (*A. mono*); the accompanying species are poplar (*P. maximoviczii*) and Korean pine (*P. koraiensis*). The understorey is composed of Manchurian lilac (*Syringa amurensis*), bird cherry (*P. asiatica*) and Amur corktree (*Ph. amurense*), especially succeeding after forest fire. The undergrowth is typically composed of Manchurian honeysuckle (*Lonicera mandshurica*), mock orange (*Philadelphus tenuifolia*), Siberian ginseng (*Eleutherococcus senticosus*).

   The mixed valley pine and deciduous forests had been most common in the area of study before the industrial logging was in practice here from 1909 until 1935, the year the nature reserve was created. Besides logging, fire has been the other forest deteriorating factor. Currently, only limited number of areas are occupied by these forests on riparian terraces, on well−developed and moistened soils.

4. The "damp" (or "humid") pine stands occur in broad valleys on raises of modern riparian terraces; they are typically composed of 3 or 4 Korean pine trees and to 5 (of 10) deciduous species in the upper storey, with typical single poplar (*P. maximoviczii*), up to 160 cm in diameter. The understorey is composed of the deciduous species mixed with conifers. The dense undergrowth consists of Manchurian hazel (*Corylus mandshurica*), mock orange (*Ph. tenuifolia*). The grass storey is up to 60−70 cm high and covers 100 percent of the soil surface.

5. The "wet" pine stands occupy the lower terraces and sub−terrace parts, in summer they are often flooded. They are composed of Korean pine (*P. koraiensis,* to 3 trees), poplar, ash−tree and Manchurian basswood (*Tilia amurensis*), and single Mongolian oak (*Quercus mongolica*), silver birch (*Betula costata*). The understorey includes conifers (to 3 of 10 trees), Manchurian alder and lilac, and Pallas's apple−tree (*Malus pallasiana*). These pine stands are found rather seldom and occur as fragments inside the "damp" pine stands.

6. The "pine with dark conifers and tilia" stands can consist of to 7 pines and to 3 tilia (of 10); they mostly occur in the same conditions as the "damp" pine stands, however in more narrow valleys of streams. The understorey is mainly composed of Eastern Siberian fir (*Abies nephrolepis*) with admixture of pine and spruce.

7. The "pine with fir and silver birch" stands are found on the upper riparian terraces adjoining the northern slopes; the dominant species is Korean pine (to 8 trees of 10). The understorey mostly consists of conifers (pine, sruce and fir).

8. The "spruce and fir" forest stands are found at the upper courses of the streams, and less often at the middle courses. The accompanying species in the upper storey are pine and larch. In the understorey, the fir and spruce are equal, and only few alder, pine, Ukurundu maple (*Acer ukurunduense*) trees occur there.

9. The "white birch" stands are secondary forests that replace the primary ones (*i.e.* , the pine, spruce and fir) as result of repeating fires. They are found on terraces of various levels and differ between in composition of the understorey, depending on the moisture of the soil. Birch, oak, and also bird cherry (*Padus maximoviczii*), maple (*Acer* spp.), larch and tilia occur in the upper storey. The understorey after 60−80 years consists of pine, spruce and fir, especially in the stands that border coniferous stands.

10. The "larch" stands are seldom found in the study area, they can occur on various level terraces; as well as birch stands, these are the secondary formations growing on post−fire territory. The sparse upper storey also includes pine, spruce, fir and birch. In the understorey, the dominant species are birch, fir and alder.

3. Methodology

A rectangular scene 15 x 15 km, *i.e.* , 500 x 500 pixels, was selected as the study area. The slopes of all expositions were masked and removed from this scene, as well as areas without vegetation, leaving only forested plain areas of the riparian valleys (12.7% of the scene, or 31,823 pixels).

The GRASS unsupervised classification module *i.cluster* was used to create the signatures of the classification maps with the 2,3,5,7 and 9 classes, based on the combination of bands 1,2,3,4,5 and 7 of LANDSAT 7 ETM+ images. The maximum likelihood algorithm was used to produce classes (*i.maxlik*). The classified maps were exported to GRASS 5.3 sites (*r.to.sites*) and then to vectors (*s.to.vect*). The result was then exported to spatial database PostGIS with *v.out.shp* (GRASS) and *shp2pgsql* (PostGIS).

The PostGIS point entities representing pixels were processed with GEOS function "within" to find the number of pixels contained within each of the polygons on the forest stands map (scale 1:25,000) exported to PostGIS from GRASS 5.3 with *v.to.pg* module [4]. Given that areas of the polygons vary from 0.5 hectares to 136 hectares (mean = 18.36 hectares, st.dev. = 18.37 hectares), the RMS error during the coordinate transformation equal to 1 pixel, only polygons with area of 70 pixels and more (*i.e.* , those of over 6.3 hectares) were selected for analysis so that the area overlap of circle palette would be 85% had the palette center been shifted by 1 pixel.

There were therefore selected for analysis 105 polygons of 156, *i.e.* , 93.5% of the total area (mean = 25.49 hectares, st.dev. = 18.58 hectares). The forest stands were grouped into a number of various classes produced from the database in order to assign the group labels to polygons on the map, based on the tree species composition, such as percent of each species in the stand. As example, in the assessment of the deciduous to coniferous forests classification accuracy, two groups were formed on the database of forest stands by a specific query, and they were compared with the result of the N−classes unsupervised classification with assignment of the label to a polygon as that of the class having 50+ percent of pixels within the polygon. Then those labels were generalized: label ’Conifer’ assigned to polys with certain (m) classes and ’Deciduous’ to the other (N−m) classes. A polygon is considered ’correctly classified’ when its classification label assigned in this way coincides with the label assigned through the specific database query, grouping polygons by the dominant forest species column and numeric value of species composition column (*cf.* Table 1).

4. Results and Discussion

It is worth to note that we can not decide in advance on how the classification is related with the composition of any forest stand, but we rather are looking for stands that show good correlations between the database information and the classification, analyze these correlations, and on this basis assign specific labels to the classes. Moreover, the number of classes with which the correlations show maximum predictability of ’correct’ classification is not known. We also take in account the high level of variety of the stand species composition within any stand, where patches of various dominant species occur, and the average composition in the database is related with the whole stand.

It is clear that in our study area there is be, at least, a coarse but accurate classification of two different forest type groups, *i.e.* , the coniferous and deciduous group of types. We also hope to find the classification scheme that correlates strongly with some of ground information in the database, whereas the number of classes is submitted by the user.

The comparison of the results for the two images shows the image taken in autumn (October 29, 2001) to be more homogeneous than the one taken in summer (July 13, 1999) (*cf.* Tab.1), the number of polygons containing over 50 percent of the same class being about three times as many as for summer, for the 7−class unsupervised classification. This conclusion can be further supported by the class separability matrix values, calculated by GRASS 5.3 module *i.cluster* , *e.g.* , for 2−classes case those were 1.1 in autumn and 0.8 in summer. It is possible to conclude that the autumn image has better separability of two forest type groups, *i.e.* , deciduous versus coniferous (*cf.* Tab.1), the fact that can be taken in account that by October 29, trees in almost all deciduous species (except oak) of the macro−region have dropped dry leaves.

	Summer			**Autumn**		
N classes	**N_p (50+)**	**N_p**	**% of area**	**N_p (50+)**	**N_p**	**% of area**
2	105	82	83.0	105	91	89.5
3	90	74	86.7	93	77	83.9
5	45	42	95.7	69	62	91.5
7	17	17	100	60	54	89.2
9	14	14	100	38	36	95.8

Table 1. Comparison of the results of the 2−classes generalization for the 2 images taken in different seasons.

N_p (50+) − number of polygons with 50 percent or more of one class; N_p − number of ’correctly classified’ polygons. Here ’classified correctly’ means ’those whose post−classification label coincides with the label assigned through the database query’.

We looked for the correlations for various classes and assigned labels to classes where it was possible (Fig.3). The criterion for the assignment of the label was the number of ’correctly classified’ polygons divided by total number of polygons under the same class, to be over 50 percent. Otherwise, no label was assigned to the group (marked with ’?’ on the diagram).

Case 1: 2 classes
Labels "Conifer(1)" and "Deciduous(1)" were assigned with percentage of ’correctly classified’ polygons 85.7 (60 of 70) for the conifers and 88.6 (31 of 35) for the deciduous forest.

Case 2: 3 classes
Labels "Conifer A", "Conifer B" and "Deciduous(2)" were assigned with percentage 88.9 (24 of 27), 60.4 (26 of 43) and 100 (all of 22), respectively. The sub−group "Conifers" (classes 1 and 2) includes forests of types 3,4 and 5, *i.e.* , the most common coniferous forest in the area: deciduous elm, ash and poplar with K. pine, "damp" and "wet" K. pine forest. This sub−group is divided in "A" and "B" according to the percentage of deciduous − conifer mix, *i.e.* , the "Conifer B" includes both stands with dominant conifers (5 of 10 trees) and dominant

deciduous (elm, ash and poplar), while "Conifer A" mostly includes coniferous type stands. The links between classes of various levels signify the "inheritance", *e.g.* , the plots labeled "Conifer(1)" were labeled as both "Conifer A" and "Conifer B" in the 3−classes case, and the same is true about some deciduous stands; while none of the "Conifer (1)" was labeled "Deciduous(2)", and none of the "Deciduous(1)" as "Conifer A".

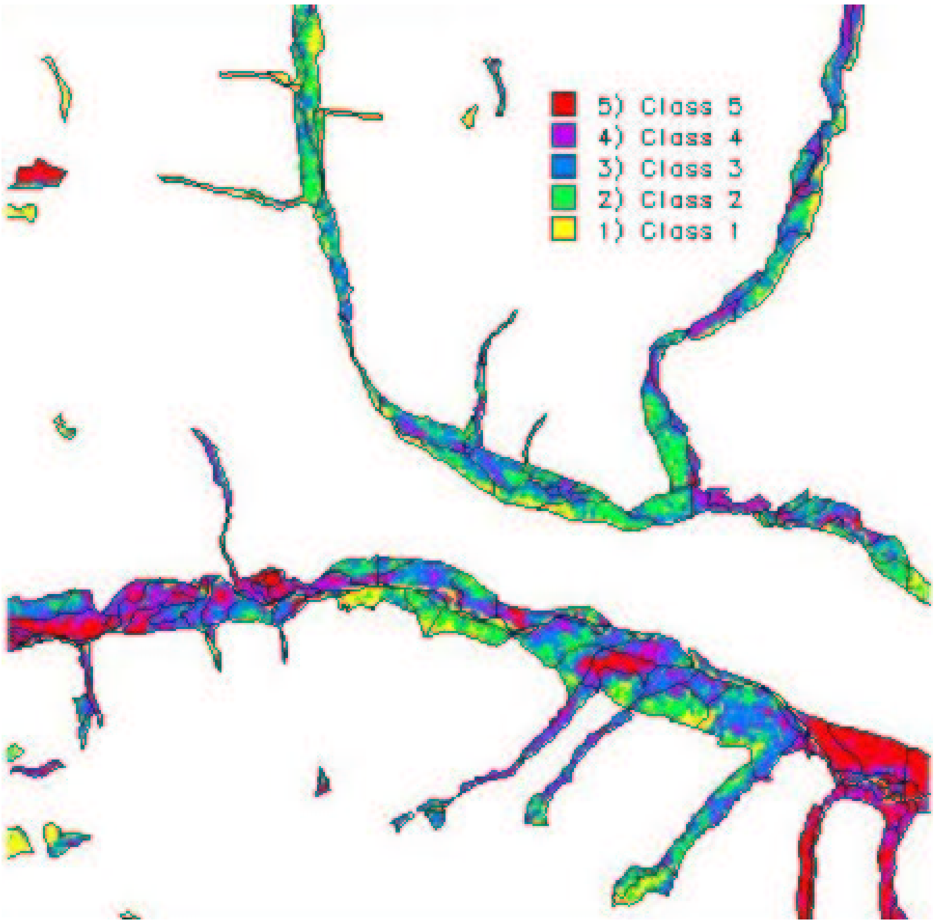

Figure 1. Unsupervised classification (5 classes) of the masked image (July 13, 1999).

Case 3: 5 classes (*cf.* Fig.1 and 2)
Labels "Dark Conifer", "Medium Conifer" and "Little Conifer" were assigned to 3 classes that followed from "A" and "B" of the 3−classes case. The "D. Conifer" label ("Class 1") was compared against selection from database with 6+ conifer (pine,spruce and fir) trees of 10 on a stand; the percentage of ’correctly classified’ in such way polygons was found 77.8% (7 of 9). The "M. Conifer" label ("Class 2") was compared against selection with 4 − 5 conifers on a stand, percentage 80.0% (20 of 25). The "L.Conifer" ("Class 3") was tested against selection with 3 conifers, percentage 54.5% (6 of 11). For all 3 groups, the conifer species were K.pine, spruce and fir, the deciduous species were poplar, elm, ash and birch. Other species (oak, larch, alder, maple, chosenia) present at least as 1 of 10 were found only in 15.5% of polys.

The deciduous group ("Classes 4 and 5") with labels "A" and "B" differ only in amount of conifer: the "A" group has 1 or 2 conifer of 10 trees on a stand (60%, 9 of 15 stands), while the "B" group has none in 3 plots of 5. The species composition of the stands in these two groups are similar.

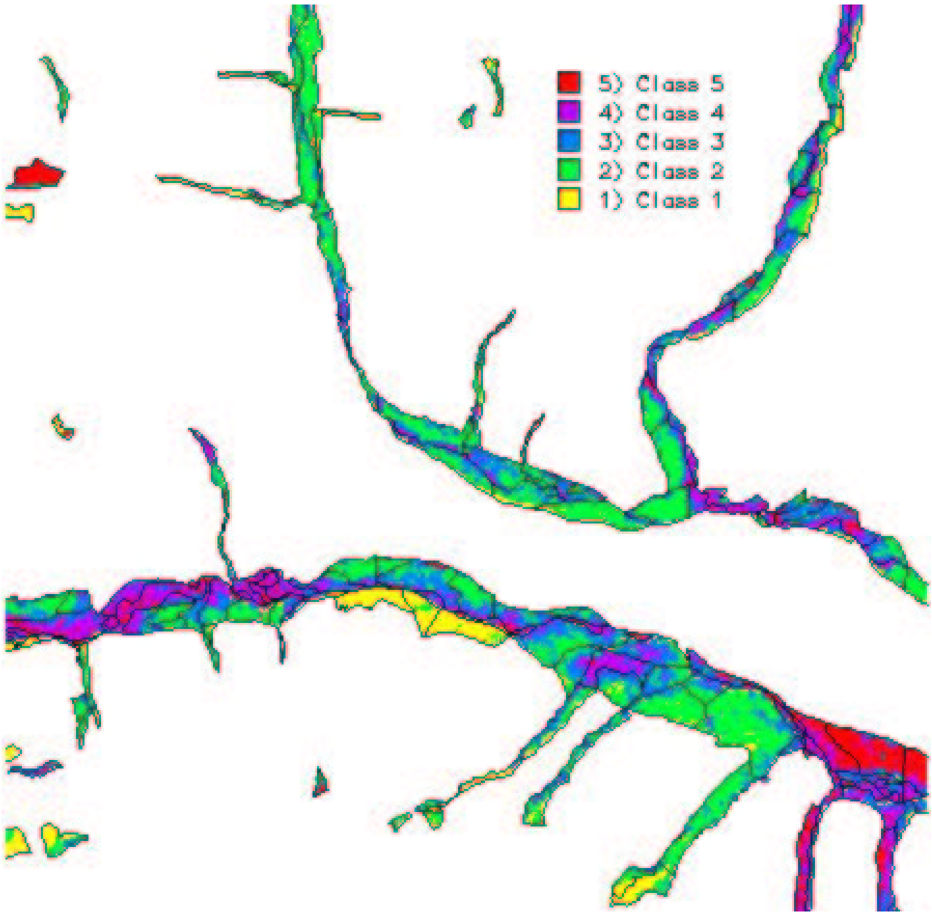

Figure 2. Unsupervised classification (5 classes) of the masked image (October 26, 2001).

Case 4: 7 classes
Two groups of conifers with 5+ trees of 10 were labeled "D.Con. I" and "D.Con. II" having the same composition structure, with no evident clue for splitting apart those classes by i.cluster after querying the database. The numbers of polys in the groups were 7 and 5, respectively, all of them having 5+ of 10 conifers. The 'M.Con' label was assigned to the group with 70.6% (12 of 17) of 5+ conifer stands. The last conifer group was too small to decide on its characteristic composition (marked with question mark, Fig. 3). The three other groups are mixed deciduous (labeled 'I' and 'II'), having similar composition of birch, elm, poplar, chosenia, tilia and maple. The group labeled 'Oak' consists of 1 polygon with 5 oak−trees of 10 on a stand.

Case 5: 9 classes
Of all 9 classes, only two conifers could be labeled (both having similar

composition with 5+ conifers) and one deciduous (namely, 'Oak'), while the rest had too few polygons to decide on the characteristic composition; nevertheless, species given after the '?' may have been the dominant for those stands. The statistics calculated from this classification had too few counted polygons for the analysis.

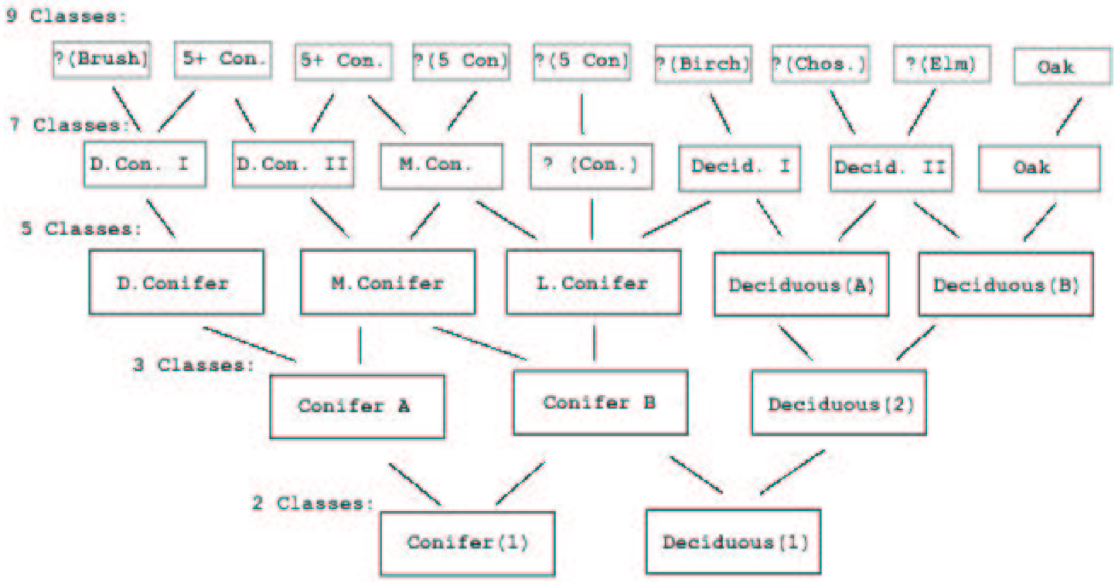

Figure 3. Post−classification species composition analysis.

5. Conclusions

In our study, we were able to find the correlation between the unsupervised classes (N classes = 5) and the percent of conifer−deciduous mix in species composition of stands. The deciduous species, such as dominant birch, chosenia and elm, could not be separated in their own classes. In contrary to this, forest with the dominant oak is likely to split from other deciduous forests. For conifers, the Korean pine, spruce and fir could not be found anywhere in the area other than in mixture with each other, therefore their classification was basically as one "conifer" ranging from 1 to 10 trees of 10 on a stand.

Of 10 different types of riparian forest characteristic for the macro−region of the eastern slopes of the Sikhote−Alin described above, types 1,2,3,4,5,6 and 9 are present with types 3 and 4 being most common. Another study area may be selected for conduct of a similar study and it has to include a wider range of different forest types, and a larger number of polygons altogether, to calculate better statistics.

References

[1] B.P.Kolesnikov. The Vegetation of the Eastern Slopes of the Middle Sikhote−Alin. −− In Transactions of the Sikhote−Alin State Reserve, V.I, Moscow, 1938, pp. 25−207.

[2] G.Xiang, Zh.Zhu, M.Hoppus, M.Fleming. Application Of Decision−Tree Techniques To Forest Group And Basal Area Mapping Using Satellite Imagery And Forest Inventory Data. −− Pecora 15 /Land Satellite Information IV /ISPRC Comission I /FIEOS 2002 Conference Proceedings.

[3] H.Saito, Kh. Aziz Hamzah, and H.Sawada. Classification Of Forest In Malaysia Using JERS−1 SAR And Landsat Tm Data. −− Proceedings of the Asian Conference on Remote Sensing, 1995.

[4] A. Shevlakov. GRASS Vector Data Management And Operation Enhancement By Using RDBMS PostgreSQL 2−D Vectors. −− Proceedings Of The Open Source GIS − GRASS Users Conference 2002. Trento, Italy, 11−13 September 2002.