Geoinformation for European-wide Integration, Benes (ed.)
 ISBN 90-77017-71-2

The preparation of map of land use in mazowieckie voivodeship based on satellite images from NOAA Satellite with AVHRR Scanner

Malgorzata Krowczynska
University of Warsaw, Poland

Keywords: map of land use, classification, satellite NOAA

ABSTRACT: People have always changed their environment in order to satisfy their needs. One of the means to evaluate its present state and identify any changes in land use are land cover maps. In my work a method of classification of low resolution AVHRR/NOAA images from May and September of 1999 was created. The goal was to make a 1:500 000 scale map of land use of mazowieckie voivodeship.

One of the main problems with a proper classification of images from NOAA series satellites with AVHRR scanner is their low spatial resolution. In the Polish reality of small parcels a few of 1100 m x 1100 m pixels are filed with just one form of land cover. Most of the pixels are the so called mixels composed of spectral values representing different forms of land use. In order to find fields filled with just one form of land cover low resolution AVHRR/NOAA images with a land use map from year 1980 at Institute of Geodesy and Cartography in Warsaw was mixed. After analysing both AVHRR/NOAA images and the map 8 land use classes on research polygon was created: water bodies, coniferous forests, leaf forests mixed forests, meadows and pastures, farming land, built-up areas and idle land. Those were classified on the AVHRR/NOAA image and are present on the 1:500 000 land use map of mazowieckie voivodeship.

Using a developed algorithm a computer automatically picked the most homogeneous fields. These fields were used as training fields for the land use classes mentioned above in supervised classification.

Evaluation of this classification, which is present on the land use map, was based on a digital land use map of the research polygon. It was conducted on two levels of essentiality general - informing about a number of pixels on the map and on the image after classification and on precise level that informs about the land use on pixels of classification image.

1 INTRODUCTION

Evaluation of classification, which is present on the land use map, was based on digital land use map of the research polygon. It was conducted on two levels of essentiality general - informing about a number of pixels on the map and on the image after classification and on precise level that informs about the land use on pixels of classification image.

From the beginning of the 90's, Poland has been going through a period of economical, social and judiciary changes. These include big changes in administration and property law. Poland is applying for membership of the European Union. This situation makes information about land use on the international, national and regional level very valuable.

Satellite imagery is widely used in preparation of land use databases. A good example is the CORINE program based on LANDSAT TM images. It has covered many European countries, including Poland.

Environmental satellites LANDSAT and SPOT provide images of high ground resolution with information in detail about earth. Such information feeds data bases that require the precise information at the price of very costly images and large amount of work needed to process them. Updating such a date base is time consuming and expensive.

The NOAA satellite provides images of rather low spatial resolution (1100m x 1100m). They contain, however, enough information to be used in preparation of a review map of land use at the scale of 1:500 000 for the mazowieckie voivodeship. The advantage of NOAA images is that they are acquired frequently, thus making it easy finding scenes without clouds.

A result of classification of those images is a numeric map of land use of mazowieckie voivodeship, which can directly feed the geographical information system.

2 METHODOLODY

2.1 *Selection of images and their preparation for analysis*

Images made with AVHRR scanner of NOAA satellite used to carry out the classification were from two dates: 20 May and 4 September. The selection of dates was not accidental, in May plants are at the early stage of phenological cycle and in September annual plants reach their maximum growth. Compilation from those two dates allows finding complete information about plant cover present on the considered area. While looking for the right picture on the research polygon special care was taken to find the ones without thin cloud layer. The clouds do not influence visual interpretation because they are hardly seen by an interpreter. For automated classification it may lead however to many mistakes, because the original values of pixels changed.

From the whole of the NOAA image a fragment of 300 columns and 300 rows, in which mazowieckie voivodeship is located, was cut off. Earlier all of the scene was geometrized. Based on 9 points the scene was transformed to equiarea, conical Albers's projection. Transformation was done with an accuracy of 0.9 the size of NOAA pixel.

2.2 *Preparation of numerical land map*

The map of land use (scale 1:500 000) prepared at IGIK from LANDSAT MSS imagery from 1977 and 1978 at different phenological seasons was converted to numeric format by means of digitising the area of mazowieckie voivodeship.

Considering the scale of the map and the method of preparation of satellite data the following geographical environment elements characterising land use in Poland were chosen:

1. Hydrography
2. Forest
   - Coniferous
   - Leaf
   - Mixed
3. Meadows and pastures
4. Farming land
   - With dominance of large area fields
   - With dominance of small area fields
5. Built-up areas
   - Compact houses
   - Detached houses
6. Industrial areas
7. Areas degraded by human activity
8. Idle land

The map was transformed to equiarea, conical, Albers's projection. *Geometrization* was conducted using 13 tic points. The transformation error is about 80 m. So prepared map was used to isolate training fields and in the last part of the work to verify classification done on the AVHRR/NOAA image.

2.3 *Preparation of spectral composition of selected channels*

The set of several channels is used for classification to enhance information carried by an image. Pictures taken in different channels may duplicate information, for example may increase contrast of pixels of 4th channel of AVHRR/NOAA and increase contrast of pixels of 5th channel of AVHRR/NOAA. In such case one image can be discarded, because it does not add any substantial information to object identification. [Ciołkosz, Kęsik, 1989]. In order to measure correlation level of NOAA satellite spectral channels a correlation factor was created, (Table I). Quite good correlation is noted between channel 1 (visible light) and channel 2 (infrared). It equals 0.86. Also channel 2 is highly correlated with NDVI. Low correlation level is observed between channel 2 and 4 (thermal).It equals -0.02. Same kind of correlation exists between NDVI and channels 1 and 4.

Table I

SPECTRAL CHANEL	SPECTRAL CHANEL OF AVHRR NOAA IMAGE		
	1	2	4
1			
2	0.86		
4	-0.4	-0.02	
NDVI	-0.2	0.65	0.3

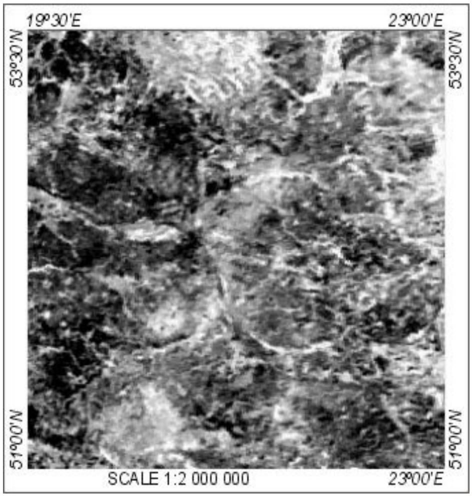

Fig. 1 AVHRR/NOAA image NDVI registered on 20 V 1999 over mazowieckie voivodship

When evaluating correlation level apart from standard channels the image of NDVI factor was included because it carries valuable information (Fig.1). Images both weight and differential stress out different details on both original pictures while extinguishing details common for both images [Sabins, 1987].

NDVI value was calculated using modified equation:

$$NDVI = ((Y2-Y1)/(Y2+Y1)) * A$$

where:

$$Y1 = 0.1065 * 4 * AVHRR/NOAA\_K1 - 4.01$$

$$Y2 = 0.1062 * 4 * AVHRR/NOAA\_K2 - 3.76$$

$$A = 1.21$$

2)

AVHRR/NOAA_K1 - channel 1 z AVHRR/NOAA (0.58-0.68m)

AVHRR/NOAA_K2 - channel 2 z AVHRR/NOAA (0.7-1.1m)

Acquired values based on equation (2) belong to the interval <-0.208,0.847>. They were converted to integer values from 0 to 65535.

A combination of channels (May NDVI, September NDVI, channel 4) that had the smallest correlation, i.e. carried the most information (Fig. 2), was used for classi-fication.

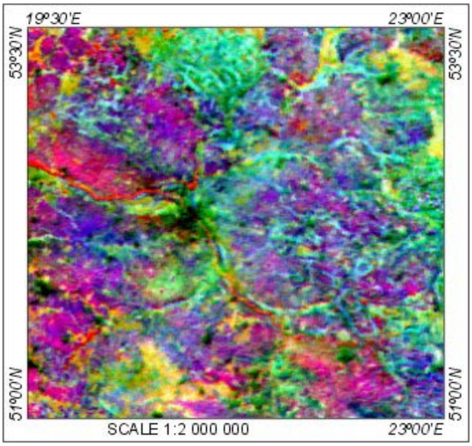

Fig. 2 AVHRR/NOAA multitemporal composite NDVI-20 V 99, NDVI -4 IX 99, channel 4 – 20 V 99

2.4 *Compilation of satellite data with numeric map of land use*

In order to separate training fields on the satellite image satellite data with land use map in system environment was joined. On the land use map ploughed grounds were described with attribute 1, pastures and meadows with attribute 2, built up area - 3, coniferous forests - 4, mixed forest - 5, idle land - 6, water bodies - 7 and unclassified areas - 8. When dividing the mentioned land cover classes to separate files, the values of those attributes was changed. To separate ploughed grounds 1 was assigned as their attribute while 0 to all the rest attribute 0. This gave a file with ploughed grounds pixels having values 1 and the background value 0. This procedure was followed for remaining classes.

Eight files connected to land use map were thus created. They were merged with three information layers (channels) of the NOAA satellite image. Merging was based on co-ordinates of ground objects. Both layers were in the same set of co-ordinates of Albers's projection.

The result of these transformations I separated files possessing information on occurrence of given class and its value recorded on AVHRR scanner image. The training polygons were ready.

2.5 *Selection of training polygons*

The purpose of classification of satellite images is to divide as accurately as possible the information into pixel groups called classes (creating clusters in multispectral space) and to define corresponding types of land use. [Domański, 1980]. The supervised classification method used is based on statistical parameters calculated from training polygons. Accuracy of defining training polygons is one of the main factors of accuracy of the whole classification procedure. When dealing with a training polygon one should therefore pick only pixels representing only one class.

Land use map described in paragraph "Preparation of numerical land map" consists of 11 components of geographical environment that describe land use in Poland. Those classes were separated using visual interpretation of colour images of high resolution LANDSAT MSS satellite image. Considering the size of pixels of 1100mx11000m for AVHRR/NOAA and 80mx80m for LANDSAT MSS it is not possible to automatically convert all of land use classes present on the map. After merging the land use map with satellite NOAA image and after adding a two pixel buffer (to eliminate transformation and raster format errors) some classes have very small areas. They are represented by only several pixels of the NOAA image. Classes of similar types and small areas were joined, because defining training polygons for such classes on AVHRR/NOAA image is difficult or even impossible. Table II presents 8 newly formed land use classes.

Table II

CODE	DESCRIPTION OF CLASSES	CODE	DESCRITPION OF COMBINED CLASSES
1	Water bodies	1	Water bodies
2	Coniferous forests	2	Coniferous forests
3	Leaf forests	3	Leaf forests
4	Mixed forests	4	Mixed forests
5	Meadows and pastures	5	Meadows and pastures
6	Farming land – large areas	6	Farming land
7	Farming land – small areas		
8	Built-up areas – compact settlement	7	Built-up areas
9	Built-up areas – dispersed development		
10	Industrial areas		
11	Idle land	8	Idle land

Table III

Code	Description of classes	Polygon area [km²]	Polygon area after classification [km²]	Classification/polygon [%]
1	Water bodies	454	454	-
2	Coniferous forests	8180	7914	96,7
3	Leaf forests	1005	504	50,1
4	Mixed forests	2501	2472	98,8
5	Meadows and pastures	6958	5172	75,0
6	Farming land	38784	41087	106,0
7	Built-up areas	1215	1494	123,0
8	Idle land	148	148	-

All changes shown in Table III were added to the numerical land use map resulting in an increase of the size of polygons representing each class.

2.6 *Classification of NOAA image*

The classification of the AVHRR/NOAA image was conducted on a colour composition made of channels [NDVI-May, NDVI-September, 4th channel - May], using algorithm presented in paragraph " Selection of training polygons". In this algorithm training polygons were picked automatically by the system as a result of merging numerical land use map for each type of land use with AVHRR/NOAA satellite image. Supervised classification function - highest probability of the ERDAS system was used in order to transform the image.

Quality and quantity analysis of classification of AVHRR/NOAA image

Quality and quantity analysis of classification of AVHRR/NOAA image using classification [NDVI-May, NDVI-September, 4th channel - May] was done. Quantity verification of classification was based on comparing number of pixels of given class present on the training polygon with number of pixels classified as this class.

Good results of quantity verification of classification of AVHRR/NOAA image were achieved for classes: "mixed forests" - 98,8%, "coniferous forest" - 96,7%, "farming land"106%. Error for these classes did not exceed 6%. Classes: "built-up areas" and "meadows and pastures" were classified with quite good accuracy "built-up areas" - 123%, "meadows and pastures" - 75%. "Leaf forest" class was affected with largest error estimated, its accuracy was only 50,1%.

Two classes "water bodies" and "idle land" require a separate explanation. These classes correspond to a very small number of pixels matching certain homogenous level. Therefore they not included into classification process. They will be later added as separate layers to the classification results. Before quality analysis pixels representing "water bodies" and "idle land" were subtracted from classes "farming land" and "coniferous forest" because the pixels of those classes were present within them. As the result of this transformation equal number of pixels from classification and from the land use map was reached.

The quantity analysis presented above allows only assessing the number of classes present on an AVHRR/NOAA image. It does not give any information about pixel structure of this image. To get this kind of information quality analysis was conducted. It was done by comparing samples (300 pixels) randomly chosen from the land use map. Accuracy of the whole classification, as well as of separate classes, was calculated by comparing categories defined by the sample with a class form the classification. Table IV shows the number.

Table IV

Image / Map	Coniferous forests	Mixed forests	Leaf forests	Meadows and pastures	Farming land	Built-up areas	Total	Pixels properly classified [%]
Coniferous forests	47	0	0	0	9	2	58	81,0
Mixed forests	0	5	3	0	0	0	8	62,5
Leaf forests	0	2	5	0	3	0	10	50,0
Meadows and pastures	1	0	0	28	10	0	39	71,7
Farming land	4	0	1	8	157	2	172	91,2
Built-up areas	2	0	0	0	1	10	13	76,9
Total	54	7	9	36	180	14	300	84,0

of pixels from the sample categorised to the corresponding class. This information can be read in rows. for example: out of 58 of the sample pixels belonging to "coniferous forests" 47 were classified correctly, 9 as farming land. The columns of this table show the number of pixels classified to certain class, hat correspond to appropriate sample fields. For example: in the column "farming land" one can read that out of 180 pixels in the sample field "coniferous forests" 9 pixels are present, "leaf forests" - 3, "meadows and pastures" - 10, "built up areas" - 1 and "farming land" 157 pixels. Considering the above examples it can be assumed that out of 58 pixels of sample field "coniferous forests" properly classified are 47/58*100%=81% of pixels. Table IV presents also (right hand, bottom corner) a summary of the accuracy of the classification. It means the share of properly classified pixels, those present on the main diagonal of the table.

To recapitulate: quality analysis shows that in classification [NDVI-May, NDVI-September, 4-channel- May] most classes are properly classified. Class "farming land" properly classified is 91,2%, only slightly result, equal to 81%, gave "coniferous forest". Classes "built up areas" and "meadows and pastures" gave results above 70%. "Mixed forest" and "leaf forest" can have quite low accuracy of 62,5% and 50%. Results achieved by quantity analysis are strange. Number of pixels of this class in classification [NDVI-May, NDVI-September, 4-channel- May] is different from "reality" only by 2%. This would suggest very good classification, but under close inspection it was discovered that only 62,5% is actually matched with "mixed forest".

Taking into consideration a disputed piece of the AVHRR/NOAA image, total accuracy of classification is of 84%. Which is good given very low ground resolution of images taken with AVHRR scanner. The reason behind such good results is that images were from two phenological seasons. This helped to separate the spectral characteristics of some land use forms.

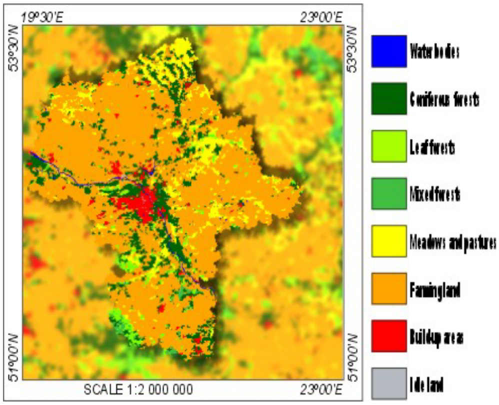

Fig 3. Land use map of mazowieckie voivodship

3 SUMMARY

On NOAA satellite images it is quite easy to recognise "farming land" and "coniferous forest". Classes corresponding to those land use forms were detected with the highest accuracy of over 85%. It needs to be stressed that the main factor was that polygons were large and filled with only one form of land use.

Classification of meadows was difficult because their polygons are very narrow and their pixels are often "mixels". The result was that some small meadows were not detected during classification of AVHRR/NOAA image. Those that were present had too large areas.

A similar case concerns built up areas. Generally speaking it is difficult to identify them on images taken with the AVHRR scanner of the NOAA satellite. Cities of over 20 000 people were in most cases classified properly. Very large agglomerations, for example Warsaw, have too large areas "taking" space that in reality is coniferous forest.

In many cases “leaf forest”' and “mixed forest” were classified mistakenly. Classification met some problems. Both classes are mixed together meaning that in places where mixed forest ought to be present classification showed leaf forest and vice versa. The reasons behind this are probably the difficulties in defining areas of mixed forests on the one hand and the lack of homogeneous fields representing this kind of forest in mazowieckie voivodeship on the other.

Preliminary comparison of the land use map prepared at the Institute of Geodesy and Cartography in Warsaw in 1980 with the map based on low resolution AVHRR/NOAA satellite images shows some changes in built up areas. Large cities are increasing their area, meadows and pastures also are increasing their area. This may be linked to the ending of the process of artificial drying of meadows in order to use them as farming land.

REFERENCES

Ciołkosz, A., Kęsik, A., (1989): Teledetekcja satelitarna. PWN, Warszawa.

Davis, F.W., Simonett, D.S., (1991): GIS and Remote Sensing. Longman, Vol. 1, London.

Domański, J., (1980): Metody analizy cyfrowej wielospektralnych danych teledetekcyjnych. Przegląd Geodezyjny. Nr 3.

Egenhofer, M.J., Herring, J.R., (1991): High-level Spatial Data Structures for GIS. Longman, Vol. 1, London.

Lewiński, S., (1993): Ocena szczegółowości kartowania użytkowania ziemi na podstawie zdjęć wykonanych skanerem AVHRR z satelity NOAA. Rozprawa doktorska, Warszawa.

Rożemski, K., (1990): Zastosowanie kontrolnych punktów terenowych do geograficznej lokalizacji obrazów satelitarnych AVHRR/NOAA. Przegląd Geodezyjny, Nr 11.

Sabins, F.F., (1987): Principles and interpretation, Remote sensing, San Francisco.

Struzik, P., (1999): Zastosowanie informacji satelitarnej AVHRR/NOAA do klasyfikacji rodzaju pokrycia terenu. Instytut Meteorologii Gospodarki Wodnej, Warszawa.