ISO5725-2 STANDARD APPLICATION TO VERIFICATION OF ORTHOPHOTO-BASED IMPERVIOUS SURFACE AREA AND IMPERVIOUSNESS FACTOR DETERMINATION

B. Hejmanowska, W. Drzewiecki, A. Wróbel

Dept. of Geoinformation, Photogrammetry and Remote Sensing of Environment, AGH University of Science and Technology, Krakow, Poland – (galia, drzewiec, awrobel)@agh.edu.pl

KEY WORDS: Imperviousness Factor, Photointerpretation, Accuracy, IKONOS, Airborne Ortophotomap

ABSTRACT:
The main aim of the research described in the paper was analyse the accuracy of photointerpretation of impervious surface using IKONOS images and its influence on the imperviousness factor determination. Two kinds of IKONOS image were chosen to tests: panchromatic and colour pansharp. Airborne ortopho (pixel size of 0.2m) was applied as a reference. Six operators, digitised three times, two kinds of IKONOS image, on the six test areas (300x300m). Accuracy analysis was performed applying different parameters, among others: RMS and reproducibility (ISO 5725-2). Then, each test area was grided with 30m pixel size (simulation Landsat image) and imperviousness factor was in each pixel determined. Mean error for PAN image was ca. 20% and for RGB image ca. 10%.

1. INTRODUCTION

In many cases traditional land-use / land cover map created through classification of satellite images does not provide us with information necessary for evaluation of changes occurring in the landscape. The process of landscape urbanization can be given as an example. In this case the changes may be twofold. What can change is not only the type of land-use (eg. from agriculture to residential area), but also the level of urbanisation within the same land-use type. For former the traditional land-use map is enough, for latter may not. Because what changes here it is not the land-use type, but the proportions of different kinds of land cover inside the same land-use class. Very detailed land-use / land cover map made from high resolution satellite images or air photos could be proposed in such a case, but such a map is very laborious and expensive when large area is taken into consideration. Moreover, this kind of images may be not available for past years. Continuum-based classification of medium-resolution satellite images may be seen as a viable alternative (see e.g. Clapham 2003, Xian and Crane 2005, Xian 2006). As a result of such classification a map of imperviousness factor is obtained. The imperviousness factor can be defined as a percentage of the area (e.g. percentage of the image pixel) covered by impervious surfaces (such as roofs, asphalt roads, parking lots, etc.).

Medium resolution satellite images have been used for the assessment of the ground surface imperviousness from 1970s (see Jackson 1975). Initially the methodology was based on supervised or unsupervised image classification techniques, but because of the resolution of these images the results were often not satisfactory. Then many new approaches have been developed, including among others artificial neural network, spectral mixture analysis or regression tree approach. A review of up-to-date techniques can be found e.g. in Weng (2008). The accuracy of the imperviousness factor estimation reported in different studies is usually better then 20 per cent.

Regardless the approach applied, the information about the impervious surfaces acquired in the field or from higher resolution data is needed as a training (or calibration) data and also for accuracy assessment. The field data are rarely available

and in the most cases such information is acquired from digital aerial or satellite orthophotos. High resolution satellite images are commonly used for this purpose. Here we comes to the question about the accuracy of these training and more importantly validation (or control) datasets. In many cases high resolution satellite orthoimages are used. Despite our efforts we weren't able to find in literature any assessments of accuracy for imperviousness factor estimations based on photo interpretation of high resolution satellite imagery. Actually the only information about the accuracy of photo interpretation based imperviousness data was find in Deguchi and Sugio (1994). They use the aerial photographs in different scales (from 1:10000 to 1:23000) to obtain the reference dataset. They report the accuracy of the estimation of imperviousness factor by visual interpretation of these photographs to be about 10 per cent. We could expect similar or even worse accuracy of imperviousness factor derived by visual interpretation of high resolution satellite imagery. The verification of this assumption was set as a goal in research presented in our paper.

2. ISO 5725-2 STANDARD AND ITS APPLICATION TO PHOTO INTERPRETATION

Acquisition of spatial data should be accompanied by acquisition of information about their quality. In our opinion information about GIS data accuracy should be seen as one of the most important metadata, especially if the data are to be used in financial context as penalty (e.g. Integrated Administration Control System – IACS, in agriculture financial subsidies in EU) or taxation (e.g. cadastre, sewer waters). A necessity for such information is also stressed in official regulations. In the Directive of European Council from 14 March 2007 establishing the Infrastructure for Spatial Information in the European Community (INSPIRE) we can find the following statement: "metadata in spatial database shall include information on the quality and validity of spatial data sets" (Chapter II, Metadata, Article 5, p.2 c). GIS data metadata as defined in ISO 19113 standard contain among others: quality, spatial accuracy, temporal accuracy and thematic accuracy.

Usually in cartographic or surveying approaches (e.g. parcel area determination) data quality (e.g. accuracy of calculated area) is estimated by RMS. National regulations determine acceptable discrepancies according parcel area, elongation or calculation method. In case of visual interpretation of remotely sensed imagery not only factors mentioned above are important. The results are influenced by the process of photo interpretation as well. In this in mind we decided to test another approach – ISO 5725-2 standard usually applied to chemical measurements. ISO 5725-2 gives "the basic method for the determination of repeatability and reproducibility of a standard measurement method". It is typically designed for interlaboratory experiments in order to estimate repeatability and reproducibility of measurement method (of chemical contents for example), but it can also be used for other purposes. Hejmanowska et al. (2005) presented its application to repeatability and reproducibility assessment of orthophoto-based measurements done during the experiment for validation of land parcel areas measurement methods. Results of the research were presented in the final report of the project supported by Joint Research Centre (JRC) in Ispra in Italy where. In the report the huge experiment of agriculture parcel measurements with accuracy assessment was described. Statistical analyses involved in the project, based on ISO 5725, were inspired by S.Kay and J.Delincé from JRC. In a typical application of ISO 5725-2 standard for basic interlaboratory experiment, samples from q batches of materials representing q different levels of the content to be measured are sent to p laboratories. Each laboratory obtains n replicated results under repeatability conditions for each of the q levels. In the case of photo interpretation several test areas should be measured on different days by different operators.

In a typical interlaboratory experiment critical examination of the data is based on a "pooling factor": the observations are grouped according to this factor and mean and standard deviation within a given group is compared to the means and standard deviations within other groups. Typically the pooling factor is the factor "laboratory", because in each laboratory replicated results are obtained under repeatability conditions. For remote sensing applications measurements made by the same operator on several days can be considered as made under repeatability conditions. As a consequence, the data should be pooled by operators.

3. TEST AREA AND DATA

Measurement experiment was performed applying the remote sensing images covering region near Cracow, on the south of Poland. Two kinds of data were applied: IKONOS PAN (called PAN in the paper; pixel size of 1m) and colour IKONOS PAN-SHAR (called RGB in the paper; pixel size of 1m) registered 7.05.2003 and delivered as 16 bits GEOTIF. Besides, panchromatic airborne orthophoto (0.2m – pixel size) generated from photographs in scale of 1:13 000 was applied as a reference.

Test area is composed of 6 rectangles of 300x300 m (10x10 Landsat pixels): 2 dense urban (3rd and 4th test area, ca. 60% of impervious surface), 2 industrial (2nd and 5th, ca. 80% of impervious surface and 2 suburban (1st and 6th, ca. 30%), (Figure 1).   Test area no. 1 with 30 m grid overlaid on the IKONOS (RGB) image is shown as an example on the figure (Figure 2).
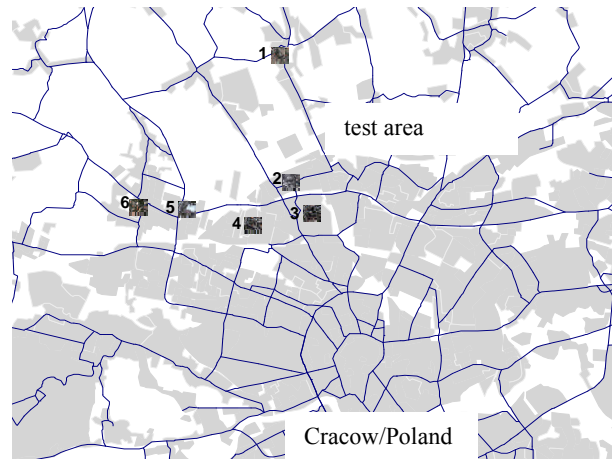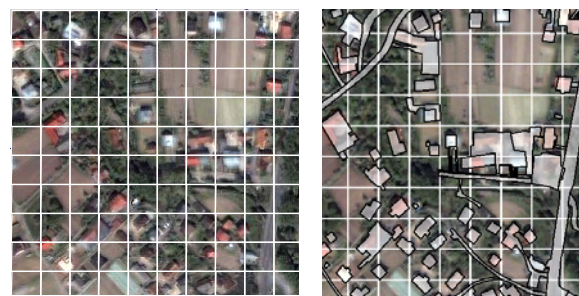

Figure 1. Test area (6 rectangles: 300x300m)

Figure 2 Test area no. 1 with 30m grid overlaid on the image, right also with reference area of impervious surface digitised on air orthophoto (0.2m - pixel size).

4. METHODOLOGY

Impervious surface area was digitised on the test area 1, 3, 4 and 6, and pervious surface in the case, when the surface is much smaller then impervious one (test area 2 and 5). Two groups of operators were chosen: 3 specialists and 3 beginners. Operator digitised firstly IKONOS PAN (3 times) then IKONOS RGB (3 times). Finally, specialist digitised ortophotomap with pixel size of 0.2 m and reference impervious surface was collected for each of six test area (Figure 2 right). Data analysis is composed of 2 parts: accuracy analysis of photointerpretation of IKONOS images (PAN, RGB) and research of its influence on its averaging in 30 m Landsat pixel size.

4.1 Part 1 – accuracy analysis of IKONOS photointerpretation

Results of the remote sensing images photointerpretation could be compared in different aspects: object recognition (object recognised or no), variation of the border shape or variation of area of the recognised object. Photointerpretation is both, time and cost consuming, so there is difficult to find in the literature some information about accuracy of the process, based on fotointerpretaion made by many operators. Usually fotointerpreatation is made once, and it is treated as a reference. One can assume, that the photointerpretation process is biased by operator, and accuracy depends on the experience of the person performing interpretation. In approach basing on the remote sensing, accuracy is usually determined by RMS, calculated from differences between mean value of

measurements and singular measurement. Instead of mean, the true value should be used, if it is accessible. Alternative method, to RMS, could be also involved to the accuracy assessment, especially if the experiment is properly prepared and measurements are repeated by a few operators. In this case, the experiment is similar to laboratory measurements usually made in chemical research, precisely determined by ISO standard. According ISO 5725 "Accuracy (trueness and precision) of measurement methods and results", among others, the following notation is used: sample, operator, laboratory, equipment, repeatability and reproducibility. As a sample, in our research we understand digitised impervious surface on IKONOS, on each of test area. In the experiment, 6 operators have been taking part using 2 equipments (IKONOS PAN and RGB). Each operator is understood as a laboratory, so we have 6 laboratories. Each test area has, as a reference, the impervious surface area, digitised on the aerial ortophoto. Six operators digitised PAN/RGB 3 times, so we have 18 measurements for each test area. Finally, 108 measurements were analysed because of 6 test area and 18 measurements of one test area. At the beginning, outliers were found for PAN and RGB images. According ISO 5725 outliers are found using graphical or numeric method, Corchan/Grubbs tests. In our research group, there was no statistician so we identified outliers in traditional way, as used usually in surveying. Relative differences between reference area and each measured area were calculated and histogram of the error was prepared for PAN and RGB (Figure 4). Outliers were defined as 5% of the external measurements on the histogram. If at least one of measurement is marked as an outlier, all group of measurements (3 repetitions) are discarded. After outliers removing, accuracy analysis was performed.

Assuming "y", as a result of measurement of the impervious surface area for each test area for each measurement, the following can be written:

$y = m + B + e$

where:

m – expected value (reference impervious area),

B – bias in repeatability condition (difference between measured area and reference area),

e - random error in repeatability condition.

Variance of B describes between laboratories variance:

$\mathrm{var}\,(B) = \sigma^2_{group}$

where:

σ_{group} – standard deviation (between laboratories).

Variance in repeatability condition for one laboratory is defined as following:

$\mathrm{var}\,(e) = \sigma^2_{l}$

σ_l – standard deviation within group (laboratory), calculated for the test area digitized 3 times by one operator.

Each operator (laboratory) is described by standard deviation σ_l. For all operators average variance is calculated, called variance of repeatability:

$\underline{\sigma^2_{repet}} = \underline{\mathrm{var}(e)} = \underline{\sigma^2_l}$.

Finally, accuracy is defined as standard deviation of reproducibility and it is the sum of the between groups variance and the within groups variance:

$\sigma^2_{reprod} = \sigma^2_{group} + \underline{\sigma^2_{repet}}$

where:

σ_{group}- standard deviation in group (one for one test area),

$\underline{\sigma_{repet}}$- average standard deviation of repetibility (average of six standard deviations for each operator).

Results of photointerpretation were validated using ISO standard and in traditional way basing on RMS. Besides

absolute RMS [sq m], relative area error (RRMS – Relative Root Mean Square = RMS/reference area with value: 0-100% or 0-1) was analysed.

4.2 Part 2 - influence of the photointerpretation accuracy on the imperviousness factor calculated in simulated Landsat pixel

Landsat classification (e.g. unmixing) requires training area, obtained from field surveying or from photointerpretation of VHR images (e.g. IKONOS). In our research, we ask the question: how does the photointerpretation accuracy influence on the percent of impervious surface in Landsat pixel?

For each test area, grid of 30m cell size was simulated. Then the impervious surface area was calculated in the grid. The procedure was performed for all measurements (18 observations for each test area) and for the reference. In each simulated Landsat pixel, area of impervious surface and percent of impervious surface area in the pixel was calculated for each measurement and for the reference. The percentage of impervious surface area in the pixel is called: imperviousness factor. Reference area and measured area were analysed in 100 pixels for the test area. Analysis was made in two aspects: comparison to the relative area error (RRMS) calculated generally for all test areas (data processing described in 4.1 Part 1) and evaluation of the absolute and relative error of imperviousness factor. Firstly, RRMS of the impervious area was calculated for each pixel. Then the differences between percent of impervious cover of the pixel basing on observations and on reference were calculated. Then, absolute RMS of imperviousness factor was calculated for each pixel. Finally, RMMS of impervious surface area was calculated for all pixels. Accuracy analysis was performed on the Landsat pixel level in 10 groups of imperviousness factor varies from: from 0% to 100% by each 10%.
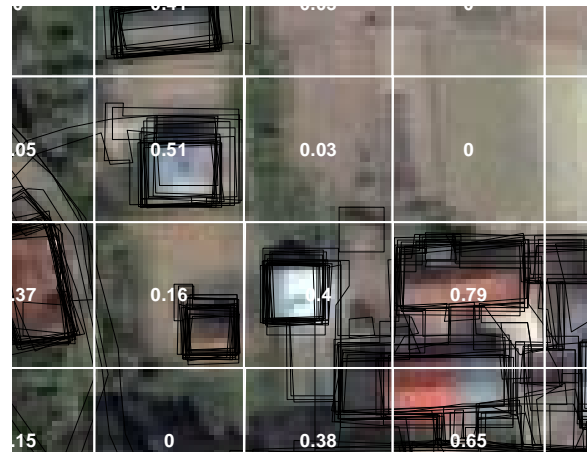

Figure 3 Part of test area no. 1 with overlaid all digitised on Pansharp colour IKONOS area; each pixel is labelled by percent of impervious surface [0,1].

5. RESULTS

Initially, all observations for PAN and RGB were statistical analysed. Histogram of relative area error for PAN and RGB is on the figure (Figure 4) presented. Mean of the difference between observed area and reference area (bias of the method)

calculated for all observations receive for PAN: -0.11 and for RGB: 0.02. Standard deviation is respectively: 0.23 and 0.17.
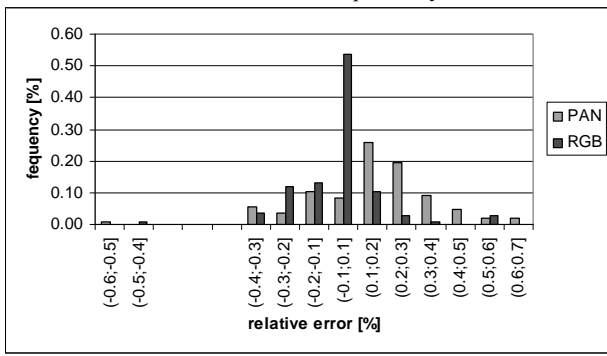

Figure 4 Relation between relative area error [%] and its frequency (histogram of errors for all data).

Reference area [m²]	σ_{group} [m²]	$\underline{\sigma}_{repet}$ [m²]	σ_{reprod} [m²]	σ_{reprod} ---------- ref. area [%]	RRMS (ref. area) [%]
23103	3022	1952	3598	0.156	0.236
14805	1922	2019	2787	0.188	0.248
52925	5669	3675	6756	0.128	0.208
51194	5284	2718	5942	0.116	0.246
19350	1376	1096	1759	0.091	0.178
25664	2268	2134	3114	0.121	0.180
mean				0.133	0.216

Table 1. Accuracy analysis PAN – repeatability, reproducibility and RRMS

Reference area [m²]	σ_{group} [m²]	$\underline{\sigma}_{repet}$ [m²]	σ_{reprod} [m²]	σ_{reprod} ---------- ref. area [%]	RRMS (ref. area) [%]
23103	2105	1212	2429	0.105	0.180
14805	1235	1142	1683	0.114	0.168
52925	2091	2395	3179	0.060	0.070
51194	1840	2194	2863	0.056	0.054
19350	1528	1097	1881	0.097	0.113
25664	2037	1754	2688	0.105	0.180
mean				0.089	0.127

Table 2. Accuracy analysis RGB – repeatability, reproducibility and RRMS

Test area no	Reference area [m²]	Δ (mean obs.- ref. area) [%]	δ (mean.) [%]	σ_{reprod} ---------- ref. area [%]	RRMS (ref. area) [%]
1	23103	-0.163	0.212	0.156	0.236
2	14805	0.131	0.194	0.188	0.248
3	52925	-0.167	0.159	0.128	0.208
4	51194	-0.234	0.139	0.116	0.246
5	19350	0.115	0.128	0.091	0.178
6	25664	-0.152	0.117	0.121	0.180
mean		0.158	0.133		0.216

Table 3. Accuracy analysis PAN – bias (Δ) , standard deviation (δ), reproducibility (σ_{reprod}) and RRMS

Test area no	Reference area [m²]	Δ (mean obs.- ref. area) [%]	δ (mean.) [%]	σ_{reprod} ---------- ref. area [%]	RRMS (ref. area) [%]
1	23103	0.126	0.118	0.105	0.180
2	14805	-0.026	0.184	0.114	0.168
3	52925	-0.025	0.069	0.060	0.070
4	51194	-0.001	0.058	0.056	0.054
5	19350	-0.051	0.117	0.097	0.113
6	25664	0.151	0.102	0.105	0.180
mean		0.029	0.108	0.089	0.127

Table 4. Accuracy analysis RGB – bias (Δ) , standard deviation (δ), reproducibility (σ_{reprod}) and RRMS
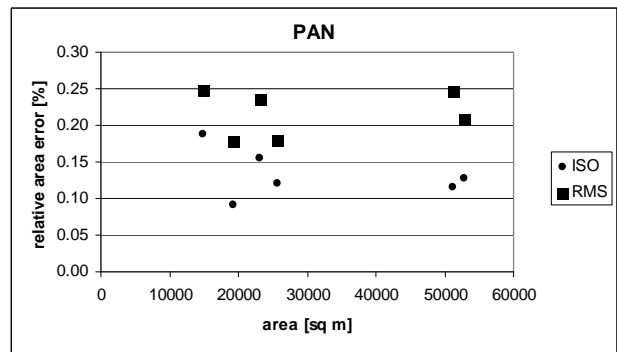

Figure 5 Relation between relative area error, RRMS [%] and test area [sq m], PAN
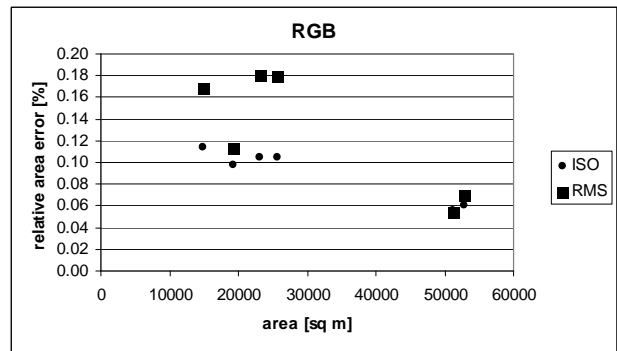

Figure 6 Relation between relative area error, RRMS [%] and test area [sq m], RGB
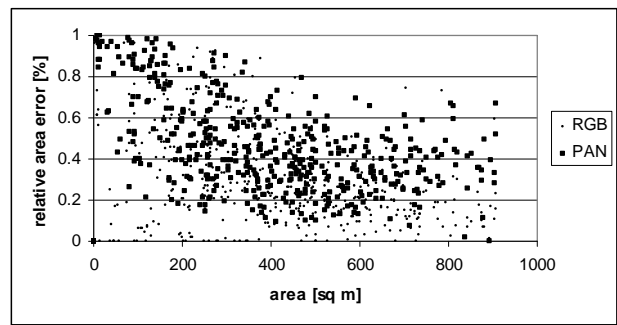

Figure 7 Relation between relative area error RRMS [%] and area
[sq m], PAN, RGB (in simulated Landsat pixels, 30m)

Observations were analysed to identify and discard outliers:
- PAN: 8 outliers (4 groups – 12 measurements),
- RGB: 5 outliers (3 groups – 9 measurements)

After outliers discarding standard deviations received 0.20 (PAN) and 0.14 (RGB). Repeatability, reproducibility and RMS for the test areas are for PAN in table (Table 1) presented and in table (Table 2) for RGB. Comparison between: bias (average observed area minus reference area), standard deviation (from the average observed area), reproducibility and RMS is possible in tables (Table 3, Table 4). Relationships between reproducibility, RRMS and the test area are shown on the diagrams (Figure 5 and Figure 6). Mean value of reproducibility for PAN is equal: 13.3% and respectively RRMS: 21.6%. Better results were obtained for RGB, mean reproducibility is equal: 8.9% and respectively RRMS: 12.7%. Photointerpretation of PAN was biased more than RGB (see Table 3 and Table 4). Mean bias on PAN for test areas 1,3,4,6 (digitized impervious surface) was -18% and for test areas 2, 5 (digitized pervious surface): +12.3% (average error of Δ in Table 3 is 16%). It means that in photointerpretation on PAN impervious surfaces were underestimated in compare to the reference. This relationships is not observed on the RGB, bias is smaller than on PAN, varies from minus to plus values and mean of RRMS is 2.9% (Table 4). Relationships between the RRM and test area are presented for PAN and RGB on the diagrams, respectively on Figure 5 and Figure 6. The bias, appearing in the PAN observations, is possible to be seen on the figure Figure 5 as a shift of RRMS up to the reproducibility (ISO). This phenomenon is not, in this scale, observed in RGB observations (compare Figure 5 and Figure 6). For example, RRMS for the test area 3 and 4 is almost equal to the reproducibility, and the bias is small.

The figures: Figure 7 - Figure 11 show the accuracy analysis of data, in simulated Landsat pixels (30m). Relation between the RRMS, relative area error of impervious surface in the pixel, and reference area of impervious surface is presented on diagram (Figure 7). Usually, the absolute area error increases with increasing of the measured area, but simultaneously RRMS is decreasing. The relationship is valid for the area from digitalisation of remote sensing images or surveying. It is easy to be observed in the case of cadastre parcels. Our object of interest was however other. There are in fact many polygons, analysed totally in test areas (Figure 5 and Figure 6). Therefore, the tendency of decreasing of RRMS with the area is slightly to be noticed on the diagram (Figure 5) and better in the RGB observations (Figure 6). The total measured impervious areas in test areas are varying between 10 000 – 60 000 sq m, and the RRMS varies from 6%, in RGB, to even 25 % in PAN observations.

Analysing the same relationship in grid of 30m, we could obtain even huge error for small areas heading to zero, because in this case RRMS heads to infinity. Maximum area in grid 30m is 9000 sq m, so in Landsat pixel impervious area varies from 0 to 900m, and RRMS varies in range of zero to infinity. But the decreasing tendency with area is noticeable on Figure 7 , even in the "cloud of points". On the other hand we use results of photointerpretation as a test area in Landsat image classification. Therefore, we compared imperviousness factor calculated for pixel on the basis on all observations and on the reference. Usually we assume result of one photointerpretation as a reference. In our research we have 18 measurements because 6 operators digitised the area 3 times.

Before statistical analysis, values of RMS for all test area and all pixels, were combined into 10 groups depending of the mean imperviousness factor from 0 to 100% by each 10%. Bias, and RMS of imperviousness factor are presented in 10 groups for PAN (Figure 8) and RGB (Figure 9) observations. Maximum bias in PAN measurements was slightly more then -30%, it means that operators underestimated impervious areas. Absolute RMS increases with increasing of imperviousness factor even more then 35% on IKONOS PAN. IKONOS PANSHARP (RGB) allowed obtaining much better results (Figure 9), small, neglected bias and RMS in some cases only slightly more then 10%.

Finally, the RRMS of impervious surface area was calculated in each pixels for comparison to the RRMS of impervious surface area calculated for all test area (compare Figure 5, 6 and Figure 10,11). In the analysis the value of the impervious area should be take into consideration, Fig, 5, 6: 10 000 – 60 000 sq m, and Figure 10, 11: 0 – 900 sq m.
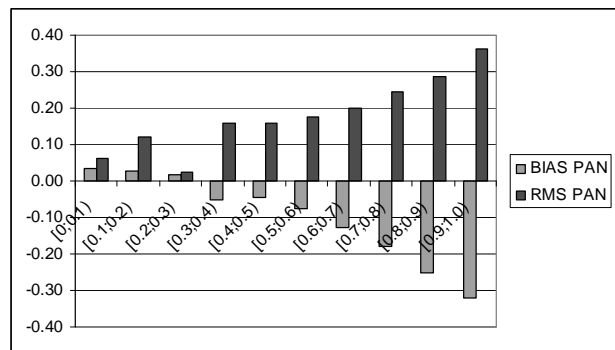

Figure 8 Relation between the absolute errors of imperviousness factor: RMS, bias and the imperviousness factor combined into 10 groups, PAN
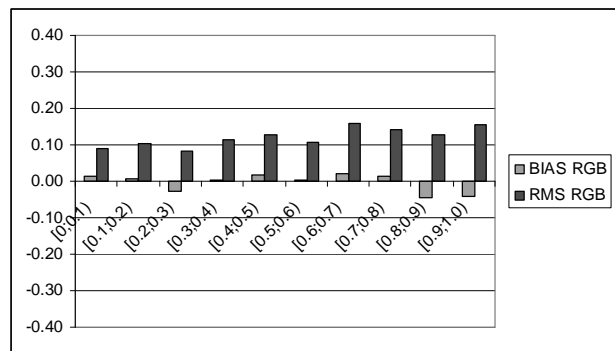

Figure 9 Relation between the absolute errors of imperviousness factor: RMS, bias and the imperviousness factor combined into 10 groups, RGB
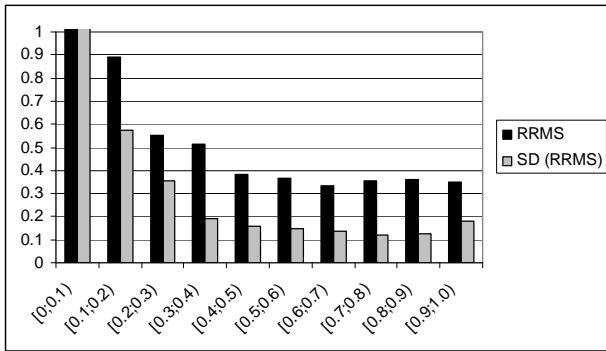
Figure 10 Relation between RRMS and standard deviation of RRMS of impervious surface area and the imperviousness factor combined into 10 groups, PAN
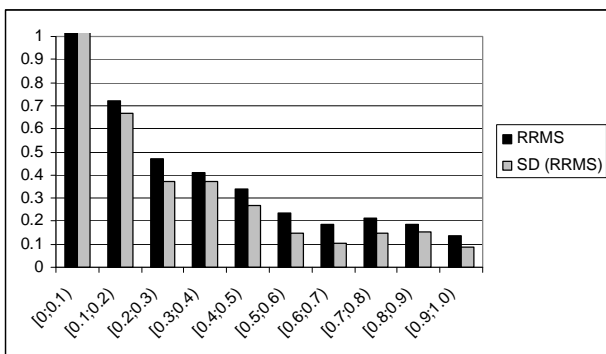

Figure 11 Relation between RRMS and standard deviation of RRMS of impervious surface area and the imperviousness factor combined into 10 groups, RGB

6. CONCLUSIONS

According measurements experiment and statistical analysis the following conclusions could be drawn out:
1. Accuracy of manual photointerpretation of panchromatic IKONOS was describe by:
    -   mean relative area error (RRMS): 22%,
    -   average bias: 16%,
    -   mean reproducibility: 13%.
2. Accuracy of manual photointerpretation of IKONOS PANSHARP was describe by:
    -   mean relative area error (RRMS): 13%,
    -   mean bias: 3%,
    -   mean reproducibility: 9%.

In our experiment, panchromatic IKONOS allowed for the impervious surface interpretation with less accuracy in compare to the colour IKONOS PAN-SHARP.

Accuracy of photointerpretation influences on the value of imperviousness factor, calculated in 30m grid and later on

applying it, as a reference in Landsat classification. In our experiment, we obtained also less accuracy for panchromatic image than for colour one:

-   PAN - mean RMS of imperviousness factor: 18%, mean bias: -10%,
-   RGB - mean RMS of imperviousness factor: 12%, mean bias: -4%.

Accuracy obtained from IKONOS RGB is comparable to the results published by Deguchi and Sugio (1994). Results from panchromatic images were significant worse then obtained from the colour images. Some explanation might be the spectral range of panchromatic channel of IKONOS covering also infra red, vegetations on the PAN image is bright and might be misrecognised as a bright concrete cover (or against). Generally, colour image contains more information useful for impervious surface recognition.

REFERENCES

Clapham, W. B., Jr., 2003. Continuum-based classification of remotely sensed imagery to describe urban sprawl on a watershed scale. *Remote Sensing of Environment*, 86, pp. 322-

Deguchi C., Sugio S., 1994. Estimations for percentage of impervious area by the use of satellite remote sensing imagery. *Water Science Technology*, 29, pp. 135-144.

Hejmanowska, B., Palm, R., Oszczak S., Ciećko, A., 2005. Validation of methods for measurement of land parcel areas. Final Report. AGH University of Science and Technology, Faculty of Mining Surveying and Environmental Engineering, Department of Photogrammetry and Remote Sensing Informatics.
http://home.agh.edu.pl/~galia/research/Area_Validation/Validation%20of%20method%20final%20report.pdf

Weng, Q., (ed.), 2008. *Remote Sensing of Impervious Surfaces.* CRC Press, Taylor&Francis Group, Boca Raton, London, New York

Xian, G., Crane, M., 2005. Assessments of urban growth in the Tampa Bay watershed using remote sensing data. *Remote Sensing of Environment*, 97, pp. 203-215.

Xian G., 2006. Assessing Urban Growth with Subpixel Impervious Surface Coverage. In: Weng Q., Quattrochi D. A. (eds.), *Urban Remote Sensing*. CRC Press, Taylor&Francis Group, Boca Raton, London, New York, pp. 179-199.