Assessing the Urban Encroachment Phenomenon in Egypt Using Satellite Imagery

Rami Y. Khamis *, Amr H. Ali **, and Michael Hahn *

Abstract: The agricultural land in Egypt is consistently threatened to diminish by the urban encroachment phenomenon, which extensively occurred after the Egyptian Revolution in 2011. The risk is considered particularly high since the fertile and productive land of Egypt is remarkably scarce and is profoundly shrinking due to the urban infringements. The unstable conditions of the country exceptionally allowed planned and unplanned urban expansion forms to appear more and more. In this research urban expansion is studied by utilizing satellite images with distinctly different resolutions and employing diverse remote sensing classification procedures. The analysis is based on a multi-temporal change detection procedure which identifies the urban expansion in terms of location and areal extent that occurred, comparing the situation before and after the revolution.

For a regional overview regarding the areal extent of the urban expansion the results of unsupervised classification of Landsat images gave most helpful hints. Furthermore, the results of supervised classification of GeoEye satellite images achieved the most satisfactory results in determining the locations of newly-constructed buildings.

Index Terms: Urban encroachment phenomenon, urban expansion analysis, multi-resolution satellite imagery, image classification, multi-temporal change detection, Landsat, RapidEye, and GeoEye.

1. INTRODUCTION

Urban encroachments have a surprising scale and a complexity level that surpasses our present tools and techniques which we use in order to understand how the phenomena work. Referred to Sejourne (2009) the urban encroachment on the agriculture land is occurred in spite of the good productivity of agricultural land, their sale for building purposes was more paying than the revenues from farming. As well, Sims et al. (2003) highlights the fact that practically all of Cairo's urban expansion occurred on rich and fertile agricultural land. Likewise, Mekawy et al. (2012) explains that Greater Cairo is an inflating chaotic megacity that has gone through a wave of planned low-density suburban expansion, in addition to the stretch of unplanned informal areas at its peripheries. Moreover, Sims, D. (2012) emphasized that the rate of non-authorized, informal buildings after the revolution has increased by 2.5 times over the rate before the revolution. Here it is important to clarify that such new constructions begin with a random local labor employment. The highest infringements took place in the region of Greater Cairo and Delta compared with the Nile River Valley cities till Aswan in the South.

This research aims to assess the urban expansion that took place in the region of Greater Cairo (Egypt) after January 2011 revolution with additional focus put on the agricultural land invasion, that is, urban encroachment, in terms of location and extent. For that purpose a framework of remote sensing procedures including image classification and chance detection is developed. Beyond the study region the processes can be applied onto urban spots all over the country and assess the emerging results.

Urban infringement on agricultural land is demonstrated in Fig. 1 with an area called Jazirat Muhammad. The images show the situation before the revolution in August 2010, after the revolution in June 2011 and finally in August 2013. The dramatic increase of infringements is clearly visible.

The relevance of the research is emphasized by knowing that only 3.5% of the total land area of Egypt (1,002,450km2) is cultivated according to the State Information Services (2013).

The aims of this research are to find out "where" and "how much" agricultural land was infringed by urban structures, and from the remote sensing point of view, to find out the best classification technique for urban encroachment and finally to evaluate the quality of the change detection results.

Fig. 1: Extensive urban encroachment in Egypt. (a) five month before the revolution, (b) five months after the revolution and (c) situation in August 2013 (Google Earth 2013).

* Rami Y. Khamis, Stuttgart University of Applied Sciences
** Amr H. Ali, Benha University, Faculty of Engineering-Shoubra
* Michael Hahn, Stuttgart University of Applied Sciences

2. Study Area

Egypt General Organization for Physical Planning in 1982 defined the Greater Cairo Region (GCR) to be composed of Cairo Governorate and most of Giza and Qaliubia Governorates, as seen in Fig. 2, Giza left, Cairo right and Qaliubia above them. Greater Cairo reflects the situation of an inflating chaotic megacity that has gone through a wave of planned low-density suburban expansion, in addition to the stretch of unplanned informal areas at its peripheries (Mekawy et al. 2012). According to the 1996 Census Greater Cairo had 11,395,335 citizens in 245 square kilometers (Sims 2000).

Freely accessible satellite images of Landsat cover all Greater Cairo, even the complete Nile Delta. Furthermore, satellite images of GeoEye and RapidEye are available that cover parts of Greater Cairo.

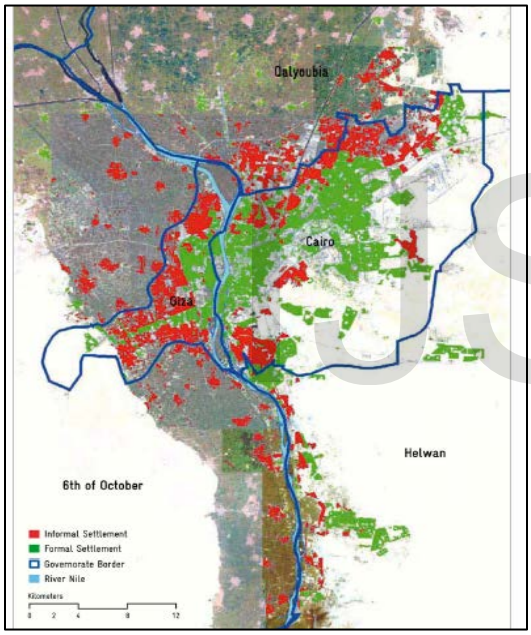


Fig. 2: Governorates of Greater Cairo including formal and informal settlements. Taken from: Kipper et al, (2009).

The Satellite images used for temporal change detection are Landsat images from years 2002 and 2013 (GSD 30m), RapidEye imagery from the years 2011 and 2012 (GSD 5m,) and GeoEye imagery from years 2010 and 2011 (GSD 0.5m pan-sharpened). For image classification and change detection for urban encroachment analysis Erdas Imagine 2010 is used which also supports quality related investigations of the results.

3. Methodology

With a descriptive analytical method first the urban encroachment phenomenon is defined and the current status of the situation in the study region from 2010 to 2013 is exposed. Furthermore, it allows supporting an analysis of urban encroachment based on a sequence of Remote Sensing procedures which are summarized in Fig. 3. Pre-processing procedures are vital to make the classification and change detection steps possible, faster and much smoother. Finally, the classification quality is assessed in a post-processing step.

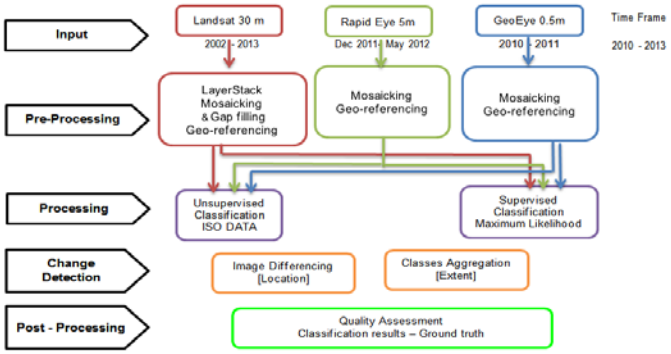


Fig. 3: A Remote sensing process for urban encroachment analysis.

The pre-processing stages handle the issues related to radiometric, atmospheric, topographic corrections, geometrical rectification, image registration and noise removal. It is important to consider using data from the same sensor, radiometric, spatial resolutions and near-anniversary acquisition dates to eliminate the effects of sun angle, seasons, and phonological difference (Schneider et al., 2003).Corrections are required to minimize the impact caused by these factors.

Image registration and multi-temporal radiometric corrections are perhaps the most important and indispensable steps in change detection methods. Precise geometric registration between multi-temporal images is essential to avoid largely spurious results, as image displacement will cause false change areas in the scene, as well as radiometric corrections, rectification errors caused by the variation in sensor characteristics, atmospheric condition, solar angle, and sensor view angle (Lillesand et al., 2004). To maintain the radiometric consistency, different radiometric correction methods are applied. The absolute radiometric correction (ARC) extracts the absolute reflectance of scene targets at the surface of the earth. The relative radiometric correction/normalization (RRC) reduces atmospheric and other unexpected variation among multiple images by adjusting the radiometric properties of target images to match a base image (Masroor Hussaina et al., 2013).

Thereafter, every set of satellite images from each of the three satellites will be handled separately into an image classification procedure. Image classification is the categorization of all the pixels of an image based either on

their spectral or spatial pattern into diverse land cover classes, which produces a thematic map and a quantitative summary statistic for these classes. Generally, there is no single correct way to go through image classification analysis. A chosen image classification approach depends on the data at hand, the hardware available and the targeted application (Lillesand et al., 2004)

According to Morans (2010) the relatively coarse resolution of Landsat (30m) imagery doesn't achieve satisfactory results for urban land use/land cover (LULC) classification in particular if the area is mixed with urban and rural spaces. He recommended to use high resolution satellite imagery (HRSI) to avoid this mixed-pixel problem. Furthermore, he found out that combining texture and spectral information in the classification procedure leads to an improvement in the overall accuracy of the results mainly because the heterogeneity problem is reduced. Also Paul M. Dare (2005) states that the Extraction and Classification of Homogeneous Objects (ECHO), as a multistage spectral-spatial classifier, brings both spectral and spatial/textural features together.

Based on that and in order to achieve better results, a hybrid approach for image classification is applied, combining the mostly automatic unsupervised classification with the supervised classification where the image analyst trains the computer algorithm how to categorize the pixels into the defined classes. (Jensen 1996).

Fig. 4 is the flowchart of the unsupervised classification based on Iterative Self-Organizing Data Analysis Technique (ISODATA).

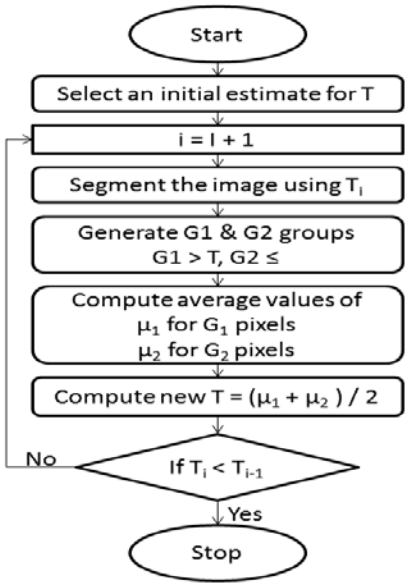


Fig. 4: flowchart of the ISODATA algorithm.

The process followed by a supervised classification based on Maximum Likelihood (ML) Classifier, where M. Oruc et al. (2001) state that maximum-likelihood classification leads to the most accurate classification results.

Maximum likelihood classification is used to categorize the three main land cover classes with their subclasses, agricultural land, urban areas and water.

This is based on the assumption that the probability distribution for each spectral class is of the form of a multivariate normal model with dimensions which equal the number of spectral bands.

Maximum likelihood derived from the Bayes theorem, which states that the a posteriori distribution P (i | ω), i.e., the probability that a pixel with feature vector ω belongs to class i, is given by:

$$P(i|\omega) = \frac{P(\omega|i)P(i)}{P(\omega)} \quad (1)$$

Where P(i | ω) is the likelihood function, P(i) is the a priori information, i.e., the probability that class i occurs in the study area and P(ω) is the probability that ω is observed, which can be written as:

$$P(\omega) = \sum_{i=1}^{M} P(\omega|i)P(i) \quad (2)$$

Where M is the number of classes. P (ω) is often treated as a normalization constant to ensure $\sum_{i=1}^{M} P(i|\omega)$ sums to 1. Pixel x is assigned to class i by the rule: $x \epsilon i \quad if\ P(i|\omega) \qquad for\ all \quad j \neq i.$

For more details about ML equation, see Ahmad, A., Quegan, S. (2012).

Focus of this research is put on the temporal pattern recognition based on the spectral characteristics of the pixels, using imagery of the same region yet with different capturing dates to detect newly-built urban features.

To define the location, where exactly the change took place, a simple image differencing operation was executed. For this operation the unsupervised classification outcome or the supervised classification outcome was used. For instance, the unsupervised classification result of GeoEye 2010 was subtracted from the unsupervised classification result of GeoEye 2011. To define the extent of change, for instance how much agricultural land was lost between the first and second date, aggregation of class agriculture with its subclasses was done, and the total area of agriculture was compared between the two dates. For this procedure, the unsupervised classification outcome or the outcome of the supervised classification was used (Congalton, R.; Green, K. 2009).

By aggregation of classes that share similar characteristics and appearance the process of change detection between two dates can be conducted. Through

this aggregation, the total change in area of each class from the first date to the second date can be calculated. To define the location of change, the results of classification of the first date is subtracted from that of the second date (image differencing).

The quality of the results of classification will then be assessed with respect to ground truth data which will be collected via visual inspection of the satellite images. Through using an error matrix and the kappa statistics $\hat{K}$ an adequate quality control can be achieved.

Equations (3 and 4) summarize how the kappa statistic is calculated (Moran 2010).

$$\hat{K} = \frac{observed\ accuracy - chance\ agreement}{1 - chance\ agreement} \quad (3)$$

$$\hat{K} = \frac{N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} X_{i+}.X_{+i}}{N^2 - \sum_{i=1}^{r} X_{i+}.X_{+i}} \quad (4)$$

Where

r = number of rows in the error matrix.

X_{ii} = the number of observation in row i and column i (on the major diagonal)

X_{i+} = total of observations in row i

X_{+i} = total of observations in column i

N = total number of observations included in matrix

4. Results, Comparisons and Discussion

All the mentioned procedures are executed on the three available data sets of Landsat, RapidEye and GeoEye in a separate manner.

4.1. Unsupervised Classification Results

Landsat unsupervised classification

The available LS-7 images are of year 2002, 2005 and 2010 while this of LS-8 is of year 2013. Line gaps that existed in LS-7 2005 and 2010 images could not be sufficiently compensated by a gap filling operation. Consequently all the images of these years were not used in this research because the line gaps lead to clear misclassification.

The results of unsupervised classification (USC) are shown in Fig. 5 for Landsat-7 2002 (14 classes that aggregated in 4 classes) and Fig. 6 for Landsat-8 2013. Table 1 contains the land use and land cover (LULC) areas of each class respectively along with the areal extent of change between years 2002 and 2013 for each class.

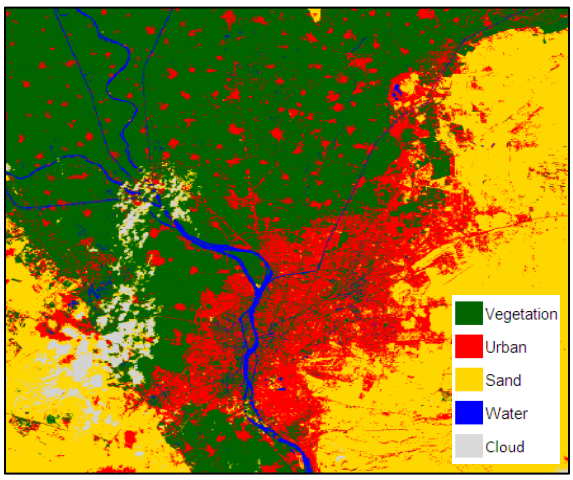


Fig. 5: Aggregated 14 classes for Landsat-7 2002 unsupervised classification.

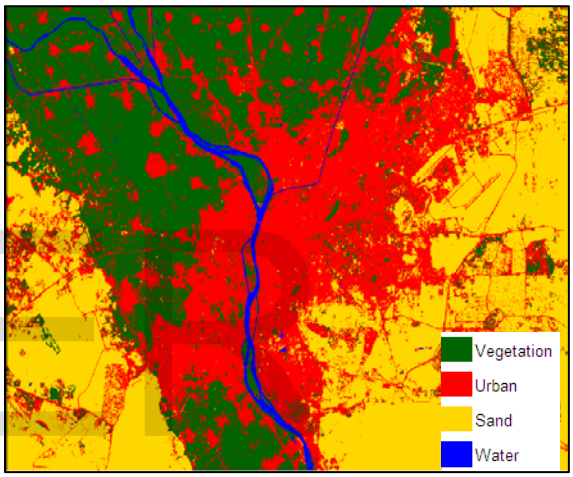
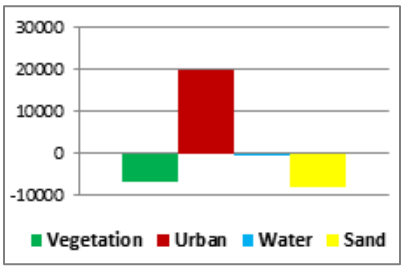


Fig. 6: Aggregated 14 classes for Landsat-8 2013 unsupervised classification.

TABLE 1 DIFFERENCE OF TOTAL AREAS BETWEEN AGGREGATED USC LS-7 2002 AND AGGREGATED LS-8 2013.

Landuse Type	Area (hectares) USC LS-8 2013	Area (hectares) USC LS-7 2002	Difference
Vegetation	99,607.50	106,325.00	-6,717.50
Urban	68,731.60	48,712.30	20,019.30
Water	3,886.20	4,158.00	-271.80
Sand	74,585.00	82,518.70	-7,933.70
Cloud	0	5096.43	-5096.43



Even though the classification procedure was unsupervised, we found, that this is a realistic result. Between 2002 and 2013 the vegetation, mainly the agricultural fields, decreased by 6717.5 hectares and the urban areas increased 20019.3 hectares.

In order to define the location of the changes, a simple subtraction of the USC LS-7 2002 outcome from the USC LS-8 2013 was done. A subset of the final result is shown in Fig. 7 where the urban encroachment is displayed in orange and urban expansion in green between years 2002 and 2013.

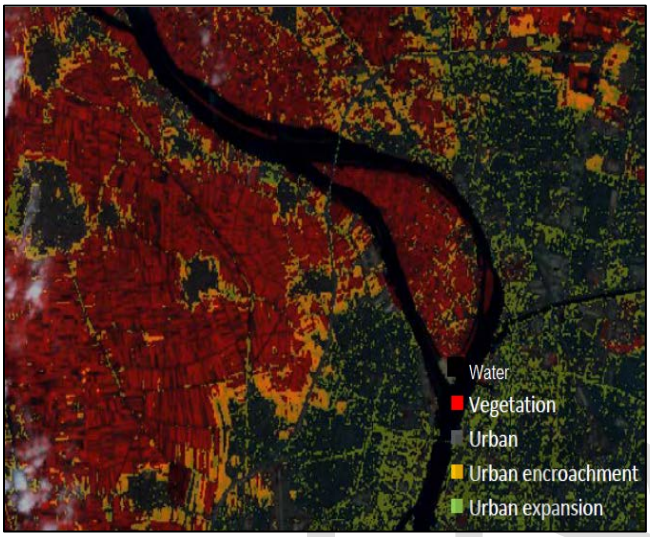


Fig. 7: Subset of Landsat USC image differencing (location of change).

RapidEye unsupervised classification

It is important to explain that the original image of RapidEye (RE) before going into the classification procedures. Band one is chosen to be represent the blue, while band two is for green and band three is for near-infrared. It is quite clear how the basic land cover classes on RapidEye satellite images can be categorized solely by looking at the image, vegetation in red, urban in gray and water in black.

After carefully examining the values of the areas of both the aggregated USC results of 2011 and 2012, it was found out that the total area of the vegetation increased, while the total urban area and water areas decreased. In reality this makes no sense at all. The reason for the increased vegetation total area in the RE 2012 satellite image with respect to RE 2011, is the water current (wind effect) on the Nile River, this water turbulence has developed a misclassification of the water to the vegetation class in the aggregated RE 2012 result, see Fig. 8 below, and accordingly a decrease of the water class was given.

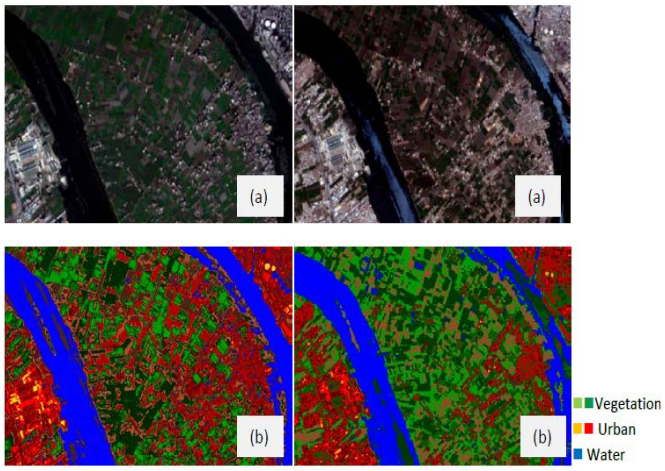


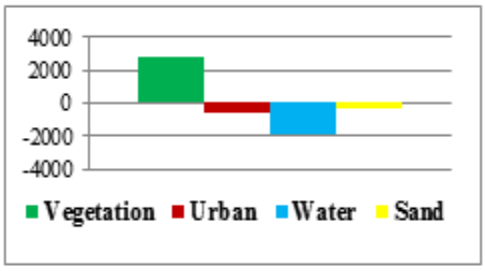
Fig. 8: Zooming on RE 2011 left and RE 2012 right: (a) original; (b) USC final result.

The main reason for the falsely-indicated urban class decrease from 2011 to 2012 is misclassification of the vegetation areas as urban areas in RE 2011 image occurred, which can be explained by the hard-to-classify soil moisture situation of the agricultural fields.

The creation of a mask for the water which was applied on the aggregated USC RE 2012 result was not satisfactory to overcome the water turbulence misclassification obstacle. Also digital number thresholding approach on the satellite image RE 2011, to bring the spectral signatures in the same range as RE 2012, was not helpful as required to overcome the soil moisture condition of the vegetation areas. The results were still not realistic, yet are compared in Table 2 which shows the differences of total areas.

TABLE 2 DIFFERENCE OF TOTAL AREAS BETWEEN AGGREGATED USC RE 2011 AND AGGREGATED USC RE 2012

Landuse Type	Area (hectares) USC RE 2012	Area (hectares) USC RE 2011	Difference
Vegetation	37937	35149.7	2787.3
Urban	25601.4	26211.2	-609.8
Water	1928.13	3810.57	-1882.44
Sand	2213.18	2508.2	-295.02



The approach of image differencing for change detection (for location of change) using the USC outcome of RE 2011 and RE 2012 respectively was attempted, yet as expected,

did not achieve satisfactory results (because it is based on the USC outcome) and thus will not be presented anymore.

GeoEye unsupervised classification

The satellite images of GeoEye-1 are High Resolution Satellite Imagery (HRSI) has a resolution of 0.5m. Thus, a bigger number of classes (16 classes) for the USC was essential. It is important to mention that the USC operation in Erdas Imagine 2010 took around 42 minutes with 16 classes defined, for an area slightly smaller than 16 km x 8 km. Here, two new classes emerged, one for ground covering 25% and another for shadow covering 20% of the USC result.

The USC aggregated classes result is illustrated with a close snap shot in Fig. 9, in comparison to the original GE 2010 satellite image and the USC 16 classes result. It is clear that the blue color still exists around the buildings and could not be separated from the blue water spectral class. Increasing the number of classes to 20 did not help to overcome this error.

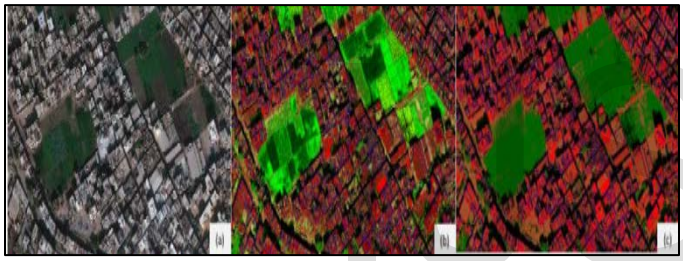


Fig. 9: GeoEye 2010: (a) Original, (b) USC with 16 classes, (c) USC aggregated result.

The aggregated classes' result of USC GE 2011 with a zoom-in can be viewed on Fig. 10. Here it is quite clear that a misclassification of pixels inside the Nile River occurred, categorizing water as urban. As explained in case of the RE 2012 satellite image, water turbulence caused by the wind led to this unusual spectral response that is similar to the urban response inside the Nile River.

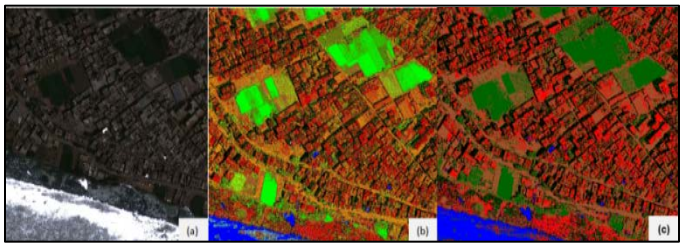


Fig. 10: GE 2011: (a) original; (b) USC result with misclassification of urban areas on as ground; (c) USC aggregated result with zooming-in.

With this, it must be stated here that also the unsupervised classification of the high resolution GeoEye satellite images was not flawless. Another example is shown in Fig. 11 where buildings are classified as ground.

To find out how much urban encroachment and urban expansion occurred between 2010 and 2011, the total area values of the USC aggregated classes of year 2010 and those of year 2011 are differentiated in Table 3 although the existing misclassification.

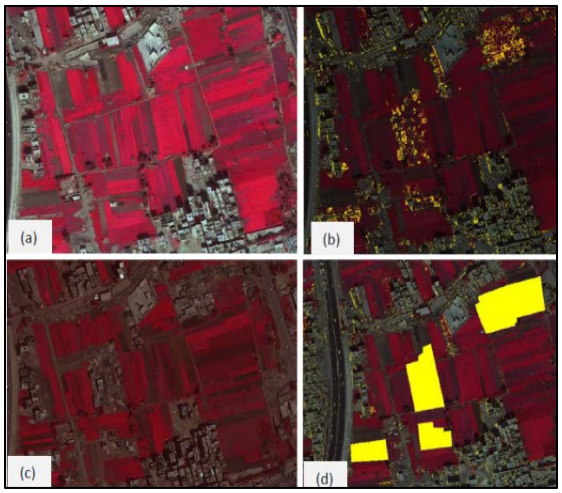


Fig. 11: a-d GeoEye USC image differencing (location of change) showing urban encroachment in orange between years 2010 and 2011, sample area 1: (a) original GE 2010; (b) original GE 2010 overlaid with change in orange; (c) original GE 2011; (d) original GE 2011 with digitized polygons of urban change.

TABLE 3 DIFFERENCE OF TOTAL AREAS BETWEEN AGGREGATED USC GE2010 AND AGGREGATED USC GE 2011

Landuse Type	Area (hectares) SC GE 2011	Area (hectares) SC RE 2010	Difference
Vegetation	2174.61	3819.86	-1645.25
Urban	1642.59	1136.59	506.00
Water	200.69	270.58	-69.89
Ground	2904.51	2397.90	506.61
Shadow	1770.78	1894.35	-123.57

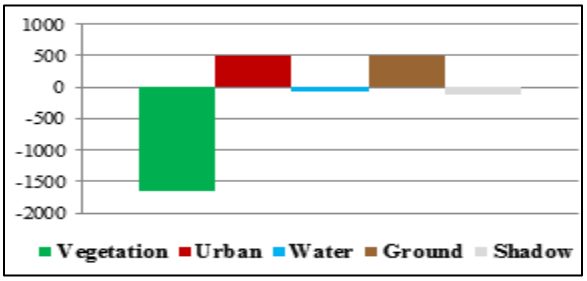


Table 3 reveals a loss of vegetation area of approximately 1645.25 hectares between 2010 and 2011, just before the revolution of January 2011. At the same time there was an urban growth, as encroachment or as expansion, with a value of 506 hectares. The decrease in the value of the shadow can be explained when examining the metadata of the satellite images, where it was found out that the two images have different sun azimuth and sun elevation angles. The increase of the ground total area can

be interpreted as agricultural land that is being 'on its way' to be urban or a case of misclassification as in Fig.10. And finally the decrease in the water total area is because of the wind current that changed the spectral response pattern of the water as explained previously, which is considered also as a misclassification.

Having the aggregated USC result of the GE satellite images 2010 and 2011 respectively produced, it is time to subtract the result of GE 2010 from GE 2011 using EI Model Maker, to be able to identify where specifically urban encroachment occurred. The result of subtraction is then overlaid on the GE 2010 image and on the GE 2011 image. From this overlay operation it is quite visible on this high resolution satellite images where the agricultural land was infringed. This is made clear in Fig.11-d, where four rectangles were digitized showing agricultural areas where direct urban encroachment occurred. The combined total area of these four rectangles is approximately four hectares of agricultural land that was infringed. This number can be considered as a rough estimation of the loss of agricultural land in favor to urban built-up areas in this region.

In order to assess the status of urban encroachment with high precision, every building that was built on agricultural land was digitized manually. This operation is performed by the image analyst to obtain exact building areas with respect to former agricultural land. As an example for this approach, which is considered the most precise among all the procedures discussed in this study, another sample area where clear urban encroachment occurred is taken under investigation.

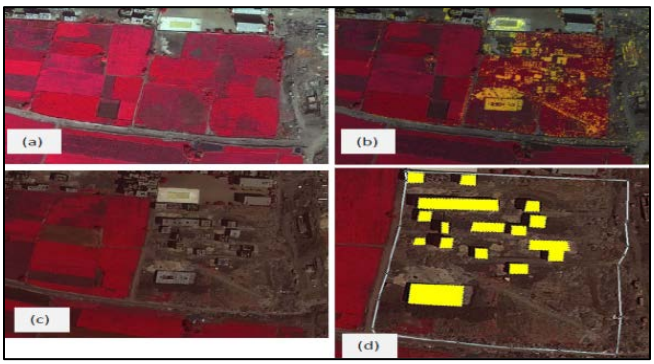


Fig. 12: GeoEye USC image differencing (location of change) showing urban encroachment in orange between years 2010 and 2011, sample area 2: (a) original GE 2010; (b) original GE 2010 overlaid with change in orange; (c) original GE 2011; (d) original GE 2011 with digitized polygons of urban change.

Fig. 12 shows sample area five, which is surrounded by a white polygon, which was previously an agricultural field area of 3.1118 hectares (measured on GE 2010 satellite image). After digitizing all the 15 buildings inside of it, shown in yellow, their respective area was accumulated to be 0.3623 hectares. This outcome states that an infringement of approximately 12 % of the formerly agricultural sample area was realized. With this digitizing approach the most reliable results can be achieved especially when using high resolution satellite imagery.

4.2. Supervised classification results

After the unsupervised classification was completed on the satellite images of Landsat, RapidEye and GeoEye, it is convenient now to test the supervised classification. Here it is important to be reminded with the aim of this study, which is to portray the urban encroachment and expansion in the time frame between 2010 and 2013, regarding the Egyptian revolution of January 2011 as an additional prime trigger for this phenomenon. It should be mentioned here that, in case of the Landsat, supervised classification will be conducted on the satellite images of years 2002 and 2013 as they have no deficiency.

The supervised classification benefits from the former unsupervised classification by adopting the spectral signatures of its defined classes from the generated signature file. After that, the chosen training regions add new spectral signatures which aim to complement a complete signature library and thus to perform a more accurate classification.

Landsat supervised classification

A total of 96 training regions were drawn on the LS-8 2013 original image, 33 for vegetation, 27 for urban, 10 for water and 26 for sand and then all these signatures were aggregated and the result is shown in Fig.13. The supervised classification for Landsat, apart from the signature selection part, was a very quick procedure.

For LS-7 2002, a total of 64 training regions were drawn, 18 for vegetation, 17 for urban, 9 for water, 16 for sand and 4 for clouds. The next step was to aggregate these signatures and the result is shown in Fig. 14.

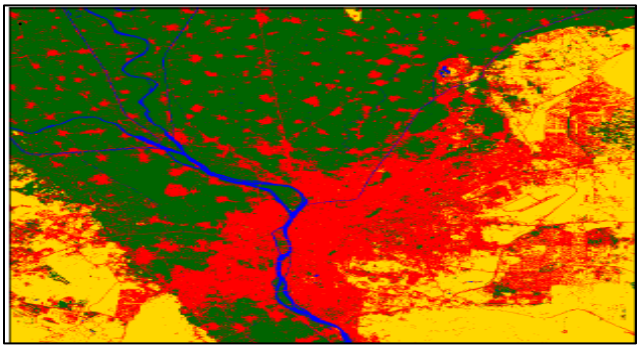


Fig. 13: Aggregated classes for Landsat-8 2013 supervised classification final result.

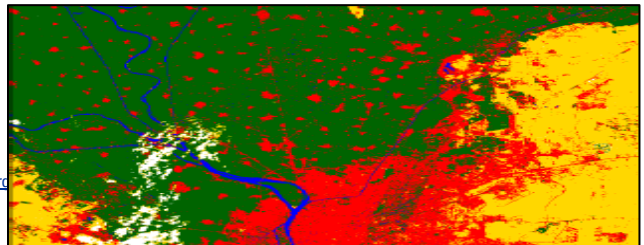
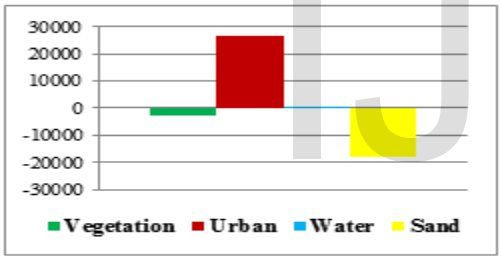


Fig. 14: Aggregated classes for Landsat-7 2002 supervised classification final result.

In order to calculate the change in total area of each class from year 2002 to year 2013, Table 4 was created, from which it is clear that the vegetation class lost 3008 hectares while the urban class gained 26790 hectares. Also the sand (desert) class lost 17990 hectares within the period between 2002 and 2013.

TABLE 4 DIFFERENCE OF TOTAL AREAS BETWEEN AGGREGATED SC LS-7 2002 AND AGGREGATED SC LS-8 2013

Landuse Type	Area (hectares) SC LS 2013	Area (hectares) SC LS 2002	Difference
Vegetation	100809.00	103817.00	-3008.00
Urban	79762.60	52972.50	26790.10
Water	3955.41	3654.36	301.05
Sand	62283.00	80273.00	-17990.00



The image difference result (using the SC outcome of each of the two dates) shown in Fig. 15 portrays the location where urban expansion/encroachment occurred highlighted in orange color. While, this result is quite satisfactory in terms of an overview for the change that occurred in the whole scene.

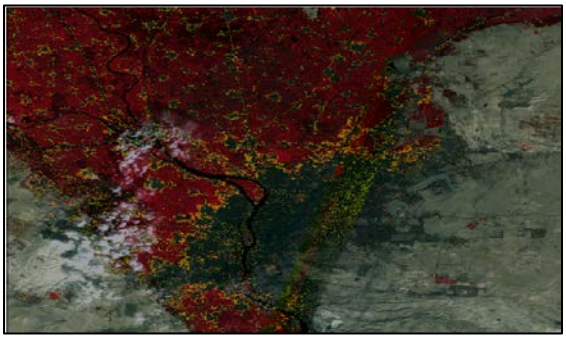


Fig. 15: Landsat SC image differencing: Urban encroachment/expansion in orange between years 2002 and 2013.

RapidEye supervised classification

By looking at the available RapidEye images of 2011 and 2012, the different sensor settings led to having a distinct illumination for each image, thus an operation of histogram matching using EI was necessary before the SC. The histogram matching operation brought the DN values of the RE 2012 in the same range of DN values of the RE 2011 image. The SC for RE 2011 satellite image required two iterations to yield acceptable results. In the first iteration a total of 71 training regions for 3 classes (vegetation, urban and water) were implemented. To improve the result, 7 additional signatures for the vegetation class were added, beside that the introduction of a forth class for sand was essential. The result of SC-2 with a total of 85 classes (47 for vegetation, 24 for urban, 7 for water and 7 for sand) was aggregated and the product is displayed in Fig.16.

Likewise, RE 2012 satellite image was investigated and SC with a total of 141 collected signatures was executed. Only a single iteration was implemented, yet the number of training regions was increased in comparison to the RE 2011 satellite image. 59 signatures for vegetation, 20 for urban, 25 for water and 17 for sand were inserted to train the algorithm. Water required 25 signatures, as the water turbulence in the RE 2012 led to an unstable signature for water areas, thus to represent all water response, more samples were captured. The result of the SC for RE 2012 is shown in Fig.17.

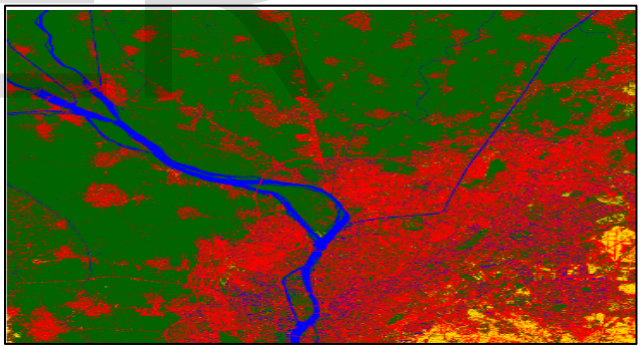


Fig. 16: RapidEye 2011 supervised classification final result.

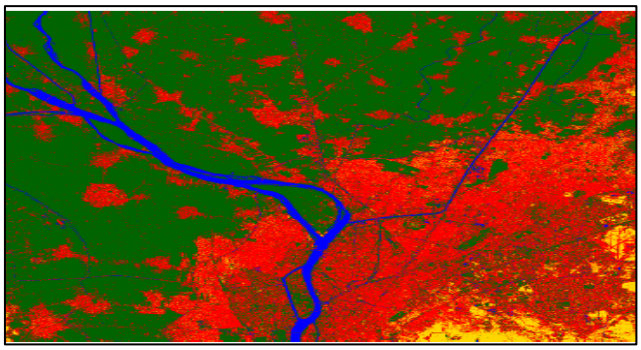


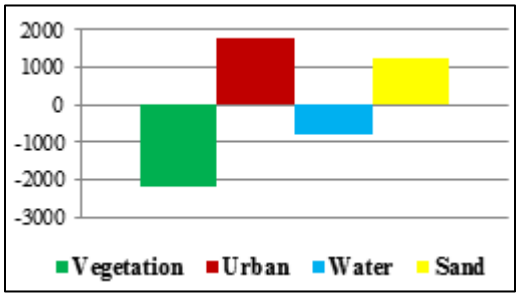
Fig. 17: RapidEye 2012 supervised classification final result.

It can be viewed from Table 5 that the vegetation, mainly the agricultural fields, decreased between years 2012 and 2011 with about 2212 hectares in the area covered by the RE

satellite images. On the other hand, an expected urban growth is estimated to be approximately 1788 hectares. The water turbulences explain the difference in water total areas between 2012 and 2011.

TABLE 5 DIFFERENCE OF TOTAL AREAS BETWEEN AGGREGATED SC RE 2011 AND AGGREGATED SC RE 2012

Landuse Type	Area (hectares) SC RE 2012	Area (hectares) SC RE 2011	Difference
Vegetation	39477.70	41689.70	-2212
Urban	22902.00	21113.30	1788.7
Water	2458.83	3244.83	-786
Sand	2841.17	1631.88	1209.29



The approach of image differencing was executed, in order to define the location where the urban encroachment happened, yet it was rather difficult, more precisely not achievable, compared to image difference using SC results of the GeoEye satellite. As the image differencing result was unsatisfactory for SC of RE, it is not included anymore.

GeoEye supervised classification

Satellite images were investigated and supervised classification procedures for the GE 2010 satellite image and the GE 2011 are classified. A total of 249 signatures were collected, 131 for vegetation, 56 for urban areas, 30 for water, 10 for ground, 7 for shadow and 15 for street. It can be stated that the third SC iteration mad a significant classification improvement which was sufficient as the overall SC GE 2010 final result as shown in Fig. 18. On the other hand, two iterations were required to receive the final result which is shown in Fig. 19.

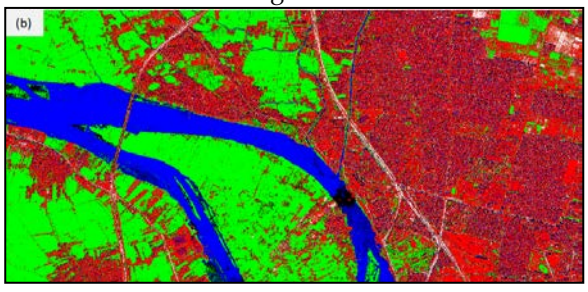


Fig. 18: GeoEye 2010 supervised classification final result.

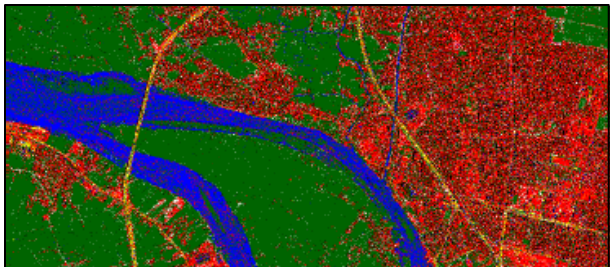
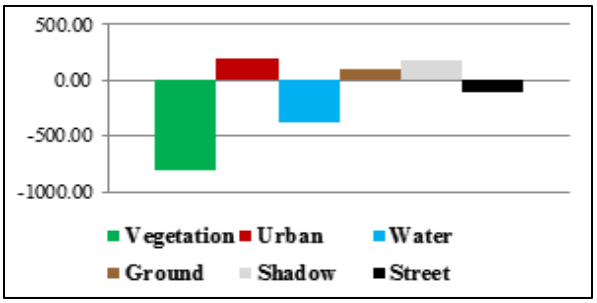


Fig. 19: GeoEye 2011 supervised classification final result.

It can be read from Table 6 that the vegetation, which is mainly represented by agricultural fields, has decreased from 2011 to 2010 by approximately 808 hectares in the area covered by the GE images. In parallel, the urban areas increased by approximately 185 acres. The reduction of the water area can be explained by a slight classification error due to water turbulence, which made the water class has a similar signature to the urban class. This error cannot to be corrected, even by taking training regions especially to compensate it in iteration two.

TABLE 6 DIFFERENCE OF TOTAL AREAS BETWEEN AGGREGATED SC GE 2010 AND AGGREGATED SC GE 2011

Landuse Type	Area (hectares) SC GE 2011	Area (hectares) SC RE 2010	Difference
Vegetation	3636.74	4444.97	-808.23
Urban	2852.94	2667.32	185.62
Water	375.10	750.13	-375.03
Ground	499.02	398.72	100.30
Shadow	1163.46	985.75	177.71
Street	165.95	269.50	-103.55



Also using the SC outcome of the two GE satellite images, image differencing was applied. From this operation areas where agricultural fields are lost in favor of urban areas are highlighted in yellow pixels, and do not represent significant class change most of the time. In Fig. 20 a sample area is shown where the GE 2011 satellite image is put in comparison with the GE 2010 image which was overlaid with the change in yellow.

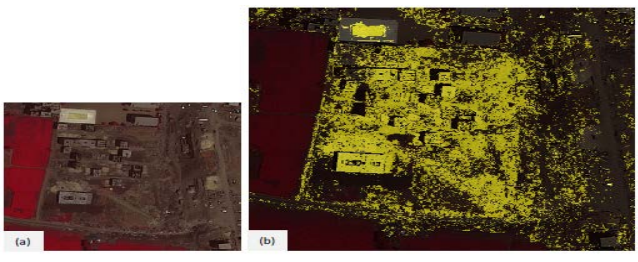


Fig. 20: GeoEye SC image differencing (location of change) showing urban encroachment in orange between years 2010 and 2011: (a) original GE 2011; (b) original GE 2010 overlaid with change in orange.

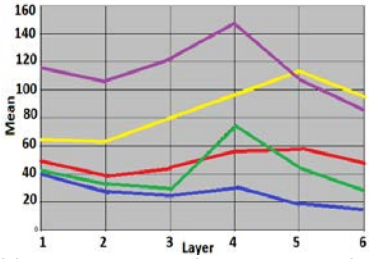
5. Accuracy Assessment

The accuracy assessment for the unsupervised classification includes 3 steps, first, a band signature separation calculation based on Euclidean distance between the class centers. Second, signatures mean plot is drawn, including the mean of each class with respect to each band of the satellite image. And third, a scatter diagram based on the red and near-infrared band is plotted with a standard deviation of 2,0 to show the relation of the classes to each other and which classes overlap and which stand-alone. For the supervised classification the error matrix and kappa statistics are calculated in addition to these three steps of accuracy assessment.

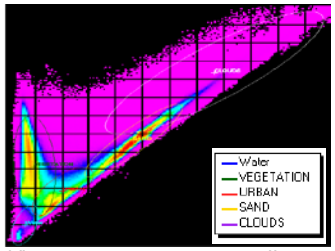
For the signature mean plot it can be generally said that if one class's line intersects another class's line they are harder to split and that the classification can include errors. Similarly with a scatter diagram, if the error ellipse of one class overlaps the error ellipse of the other class, this means that their separation during classification is rather difficult. As for the error matrix from which we get an overall accuracy of the classification process, it can be generically said, the higher the percentage the better. Also, the higher the overall kappa statistics value, the more it comes nearer to the value of 1, the better.

5.1. Unsupervised classification results assessment

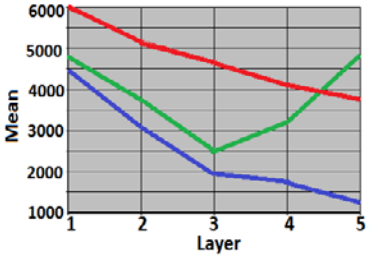
Fig. 21 represents a sample of the accuracy assessment of the unsupervised classification of the image of the three satellite images. Where figures (a,b,c) represent the signature mean plot while figures (d,e,f) illustrate the scatter diagram.



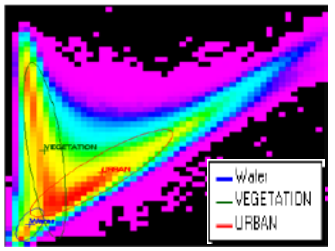
(a) LS-7 2002 USC signature mean plot



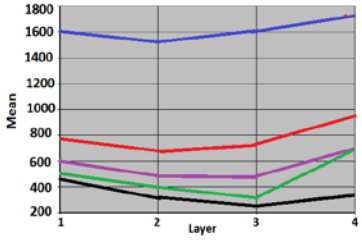
(d) LS-7 2002 USC scatter diagram



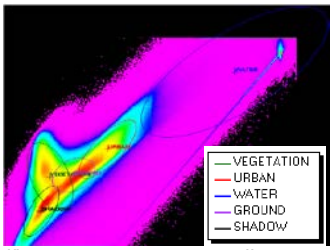
(b) RE 2011 USC signature mean plot



(e) RE 2011 USC Scatter diagram



(c) GE 2011 USC signature mean plot



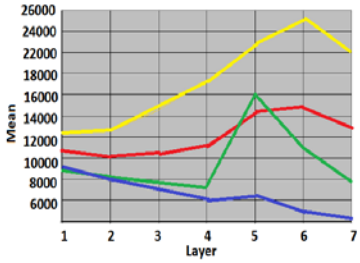
(f) GE 2011 USC scatter diagram

Fig. 21: The accuracy assessment of the unsupervised classification.

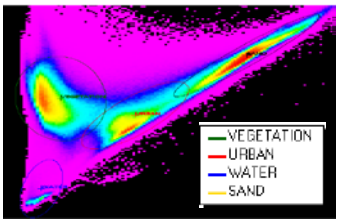
It is clear from the signature plot in Fig. 21 (a,b,c) how the class vegetation (green) overlaps with the class of urban (red) / ground (purple), where the most intersection between classes, the hardest during classification. From the scatter diagram in Fig. 21 (d,e,f) it can be read that the classes are separated in a quite acceptable manner, and that only the vegetation and urban classes have a significant overlap.

5.2. Supervised classification results assessment

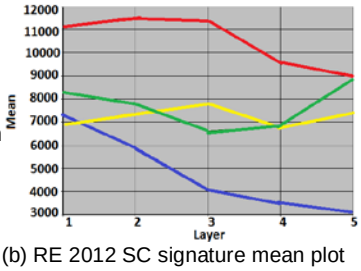
It is shown in Fig. 22 the accuracy assessment of the supervised classification of the two pairs of the three satellite images. Where the left column of the figure (a,b,c) represents the signature mean plot while the right column (d,e,f) illustrates the scatter diagram.



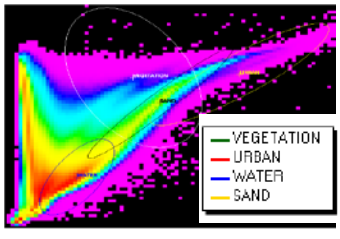
(a) LS-8 2013 SC signature mean plot



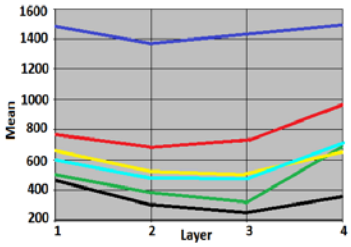
(d) LS-8 2013 SC scatter diagram



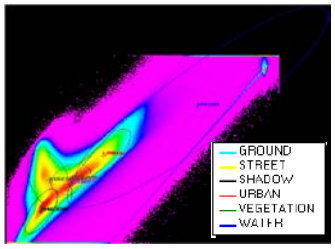
(b) RE 2012 SC signature mean plot



(e) RE 2012 SC Scatter diagram



(c) GE 2011 SC signature mean plot



(f) GE 2011 SC scatter diagram

Fig. 22: The accuracy assessment of the supervised classification.

In Fig. 22 (a,b,c) it is obvious how the vegetation line intersects the urban line in the signature mean plot in most causes that cause difficulties during the classification process.

The scatter diagram in Fig. 22 (d,e,f) shows the main overlap between urban and vegetation again. Table 7 represents the results of the error matrix of the accuracy assessment and kappa statistics.

TABLE 7 THE RESULTS OF THE ERROR MATRIX OF THE ACCURACY ASSESSMENT

Satellite	Over all classification Accuracy	Overall Kappa Statistics
LS-7 2002	96.19%	0.94
LS-8 2013	93.33%	0.90
RE 2011	98.10%	0.97
RE 2012	88.61%	0.84
GE 2010	80.00%	0.73
GE 2011	79.33%	0.71

6. CONCLUSIONS

Interpretation of satellite imagery is to take into account the date of image acquisition, the area coverage, image resolution, spectral bands, weather condition, cloud cover and other criteria that subsequently have significant influence on the overall classification process. Pre-processing is necessary for geometric and radiometric corrections, where special attention is requested for geo-referencing HR satellite imagery to avoid false image differencing results.

Moreover, satellite imagery need to be checked for wind influence on water surfaces, as wind causes water turbulences that may lead to misclassification. Using a water mask may help here, yet it was not entirely sufficient to compensate for this natural occurrence because the imagery of the first and second date, in case of RE and GE respectively, were captured with only few months difference.

For unsupervised classification, an important decision to make is the choice of the number of classes. The higher the image resolution the more classes are needed. For the LS unsupervised classification 10 classes were sufficient while for HR satellite imagery of GeoEye 16 classes had to be generated. In the urban expansion and encroachment analysis the main classes under consideration are the urban and vegetation class, nevertheless allocating classes of water, sand, ground, shadow and other was essential in order to obtain the correct total area values of the class that represent shrinkage or expansion respectively. Due to the nature of unsupervised classification, assigning proper class names to the classes (clusters) may fail. As to be expected the unsupervised classification with Landsat 30m has led to the most satisfactory results. In case of RapidEye (5m) and GeoEye (0,5m) the outcome was rather poor and unsatisfactory.

In case of supervised classification, which was started by using the signatures from unsupervised classification, the collection of further spectral signatures has big impact on the classification result. Shadow areas of buildings and vegetation have to be separately collected. In general, with an increased number of signatures and classes with subclasses better classification outcome has been achieved but at the expense of computer time-consuming process. Furthermore the higher the image resolution the more signatures are recommended. Accordingly, the LS supervised classification process was quicker than with GE. On the one the other hand, in case of LS and RE satellite imagery urban areas signature collection was done by drawing polygons over several neighboring buildings/pixels. On the other hand, in case of the HR satellite imagery of GE separate standing buildings, which are easily identified on the image, were used to collect the urban signature. It should also be taken in consideration that some vegetation fields have a spectral response pattern that is very similar to urban constructions or ground. In such cases, it might be helpful to collect even more signatures specifically for these areas which should be assigned as vegetation.

Finally, it must be pointed out that for supervised classification an overall classification accuracy (93-96%) for LS, (88-89%) for RE and (79-80%) for GE was achieved.

Change detection was carried out by image differencing the classification result of two dates of imagery. In terms of a regional overview the Landsat images indicated the location of the urban expansion and urban encroachment most successfully. That is achieved by using the supervised classification result of each date and subtracting the former date from the latter. Pixels or pixel groups where change occurred, either because of urban expansion or urban encroachment, agglomerate and are easily identifiable,

where other pixels where insignificant change occurred (maybe just a modest DN change) are standing alone. When the image differencing is applied using the SC results of the GE images, it was possible to define exactly which buildings are newly created and where precisely urban encroachment or urban expansion has taken place.

In conclusion, it can be advised to use supervised classification outcome of LS for an overview of urban change and to use supervised classification outcome of GE for exact determination of locations of new buildings. RE failed to achieve any of these two targets completely. Thereafter, the location of change has been determined, the calculation of the extent or total change in area of each class between date one and date two. Table 8 summarizes the performance of the used classification procedures in accordance to urban change detection. The (✓) symbol indicates a successful operation while the (x) means that the operation did not result in a satisfying outcome.

TABLE 8 SUMMARY OF THE CLASSIFICATION AND CHANGE DETECTION PROCEDURES PERFORMANCE FOR URBAN ENCROACHMENT/EXPANSION ASSESSMENT USING DIFFERENT SATELLITE IMAGERY RESOLUTION

Satellite	Unsupervised Classification	Supervised Classification	Change Detection (Location)	Change Detection (Extent)	Quality Assessment (SC)
Landsat (30m)	✓	✓	Overview SC ✓ USC ✓	SC ✓ USC ✓	93 - 96%
RapidEye (5m)	X	✓	SC x USC x	SC ✓ USC x	88 - 98%
GeoEye (0.5m)	X	✓	Buildings SC ✓ USC ✓	Buildings SC ✓ USC x	79 - 80%

REFERENCES

1. Ahmad, A., Quegan, S. (2012): Analysis of Maximum Likelihood on Multispectral Data Classification, Applied Mathematical Sciences, Vol. 6, 2012, no. 129, 6425 - 6436
2. Awad, M. (2010): An Unsupervised Artificial Neural Network Method for Satellite Image Segmentation, The International Arab Journal of Information Technology, Vol. 7, No. 2, April 2010
3. Congalton, R.; Green, K. (2009): Assessing the Accuracy of Remotely Sensed Data, Boca Raton: CRC Press, Taylor & Francis Group.
4. Jensen (1996): Unsupervised Classification Algorithms. [Online], Available: http://www.yale.edu/ceo/Projects/swap/landcover/Unsupervised_classification.htm [03 Feb 2014].
5. Khamis, Rami Y. (2014): Assessing the urban encroachment phenomenon in Egypt after the 2011 revolution using satellite imagery, Master thesis, Stuttgart University of Applied Sciences, Geomatics, Computer Science and Mathematics Department.
6. Kipper, R., Fischer, M. (ed.) (2009): Cairo's informal areas between urban challenges and hidden potentials, Norprint SA Portugal: Deutsche Gesellschaft für Technische Zusammenarbeit (gtz).
7. Lefebvre, H. (2003): The Urban Revolution, translated by: Bononno, R., Minneapolis: University of Minnesota Press.
8. Lillesand T., Kiefer, R., Chipman, J. (2004): Remote Sensing and Image Interpretation, 5th ed., Hoboken, NJ: John Wiley & Sons.
9. Masroor Hussaina, Dongmei Chena, Angela Chenga, Hui Weib, David Stanleyb (2013): Change detection from remotely sensed images: From pixel-based to object-based approaches, ISPRS Journal of Photogrammetry and Remote Sensing 80 91–106.
10. Mekawy, H., Yousry, A. (2012) Cairo: The Predicament of a Fragmented Metropolis,[Online],Available:http://s3.amazonaws.com/academia.edu.documents/29664402/Cairo_Te_Predicament_of_a_Fragmented_Metropolis.pdf [06 Oct 2013].
11. Morans, E. (2010): Land Cover Classification in a Complex Urban-rural Landscape with Quickbird imagery. [Online], Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3105321/ [05 Oct 2013].
12. Oruc, M.; Marangoz, A.; Buyuksalih, G. (2001): Comparison of Pixel-Based and Object Oriented classification Approaches Using Landsat-7 ETM Spectral Bands. [Online], Available: http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/510.pdf [03 Oct 2013].
13. Paul M. Dare (2005): Photogrammetric Engineering & Remote Sensing Vol. 71, No. 2, February 2005, pp. 169–177. 0099-1112/05/7102–0169/$3.00/0 © 2005 American Society for Photogrammetry and Remote Sensing
14. Sejourne, M. (2009): The history of informal settlements, in Kipper, R. and Fischer, M. (ed.) Cairo's informal areas between urban challenges and hidden potentials, Norprint SA Portugal: Deutsche Gesellschaft für Technische Zusammenarbeit (gtz).
15. Schneider, A., Friedl, M. A., McIver, M. A. and C. E. Woodcock (2003): Mapping Urban Areas by Fusing Multiple Sources of Coarse Resolution Remotely Sensed Data, December 2003.
16. Sims, D. (2000): Residential Informality in Greater Cairo: Typologies, Representative Areas, Quantification, Valuation and Causal Factors. Institute for Liberty and Democracy, Lima, and the Egyptian Center for Economic Studies, Cairo (unpublished).
17. Sims, D. (2012): Understanding Cairo - The logic of a city out of control, Cairo: American University Press.
18. Sims, D.; Sejourne, M.; El Shorbagi, M. (2003): The Case of Cairo, Egypt. [Online],Available:http://www.ucl.ac.uk/dpuProjects/Global_Report/pdfs/Cairo.pdf [24 Sep 2013].
19. State Information Service, Egypt (2013): Development of agriculture in Egypt.[Online],Available:[04Oct2013].http://www.sis.gov.eg/Ar/Templates/Articles/tmpArticles.aspx?ArtID=61672