Method for detection of small water bodies developed for purpose of cage fish farming in Eastern Africa

Petra Hesslerová[1,2], Martin Šíma[†]

[1] ENKI, o.p.s.
Dukelská 145
379 01, Třeboň, Czech Republic

2 Department of Applied Geoinformatics and Spatial Planning, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 1166,
165 00 Praha, Czech Republic
hesslerova@fzp.czu.cz

Abstract. This paper introduces a remote sensing method, based on multispectral satellite data that was developed to assess potential water bodies for regional development of aquaculture in Eastern Africa, as one of the tasks of BOMOSA project. The method will be used directly by institutions in Eastern Africa and therefore has to meet the criteria of simplicity, accessibility, low costs, high accuracy and reliability. The paper introduces two approaches of small water bodies detection. One is based on principal component analysis and the second on application of Normalized Difference Snow index and subsequent thematic recoding (density slicing). Final results were composed from the water masks of both approaches. The method was originally developed for Landsat data, with the most satisfactory results; however its application on Terra Aster and ALI EO-1 images was tested as well, due to the termination of Landsat data reception for East African region in 2001.

Keywords: aquaculture, Normalized Difference Snow Index, Principal Component Analyses, multispectral satellite data

1 Introduction

Development of the method, based on remote sensing data, for identification of small water bodies for aquaculture production for rural communities in Africa is as a very important milestone in the region's effort to turn the natural potential into food fish and provide animal proteins. Integrating BOMOSA[1] cage fish farming system in reservoirs, ponds and temporary water bodies in Eastern Africa is a project co-funded by the European Commission within the Sixth Framework Programme (no. 032103) that aims to validate this farming method in selected localities, and later promote it in similar aquatic ecosystems places in Eastern Africa (mainly in Kenya, Uganda and Ethiopia). BOMOSA Cage Fish Farming System is by design based on small water bodies which are at least 1 hectar. Most of the African countries lack reliable, accurate and up-to-date geographic and cartographic data, but there is a solution with considerable potential - remote sensing and development of the method for small water bodies detection. The method will be used directly by institutions in Eastern Africa and therefore has to meet the following criteria:

- Except for the preparation of data, all the stages involved have to be accomplished through a simple, easily accessible and user-friendly software for multispectral RS data processing and interpretation
- All the obligatory and "case" procedures involved in the data analysis and processing have to be simple and universal
- The results provided by the method have to be as concise and accurate as possible

Application and usability of remote sensing BOMOSA method for identification of water bodies, its dissemination and upscalling is highly dependent on the satellite data availability.
The satellite data suitable for BOMOSA purposes had to meet the following requirements:

- relatively high spatial resolution (with respect to minimum size of fish farming plots, 30 meters spatial resolution is acceptable)

[1] BOMOSA is the project acronym, composed from the first letters of three main partner´s institutions – **Bo**denkultur Universität in Vienna, **Mo**i University in Eldoret, Kenya and **Sa**gana Aquaculture Centre-Kenyan Ministry of Livestock and Fisheries Development, Kenya. BOMOSA also means a pioneering small-scale fish farming system in Eastern Africa, realized by establishing fish cages. For more information see www.bomosa.org

- sufficient spectral resolution to detect and analyse water bodies
- fair price (free) and easy way of obtaining data
- data coverage in as large areas as possible, in several time horizons

Despite some limits, it is the Landsat Satellite System that meets the above requirements best, so the method of small water bodies detection was originally developed for this data type. However, during the project emerged that the CSIR Satellite Application Centre (SAC) that used to serve as regional African receiving station had discontinued reception of Landsat data by 19th August 2001. It means that there are no later Landsat data available. As a potential source of new and up-to-date data Terra Aster multispectral satellite data has been tested. The Aster data are on-demand which limits the available scene archive and data coverage. Another data alternative is ALI sensor on board EO-1. The Advanced Land Imager (ALI) was designed to demonstrate improved Landsat spatial and spectral resolution and has nearly the same spectral and spatial characteristics. The data have been available since 2001.

2 Method

Of all categories of land cover discernible through multispectral satellite data, water is normally considered to be the easiest to detect. The multispectral classification is one of the most used methods for discrimination of water bodies. It is based on the different features of objects on the images which are attributed certain thematic contents based on various statistical rules. However, in order to identify the target objects correctly, it is necessary to provide a clear definition of these features called spectral signatures, using what we call training fields. It is good to know at least something about the target area when selecting the training fields. Apart from their homogeneity and complexity, the potential training fields should be assessed in terms of their size and shape. In order to avoid the inclusion of mixels from the margins of the area and prevent distorting final results, the minimum size of the training field was set at 3 x 3 pixels. As a consequence, line-shaped objects such as water streams and tiny water bodies cannot be used as training fields as they fail to meet the condition. However, the size of water bodies required for the BOMOSA project is often less than 3x3 pixels. This prevents from choosing suitable training fields and thus defining and calculating statistical features called spectral signatures whose choice is crucial to the success and accuracy of the entire classification process (Jensen 2005). Therefore, these basic processing methods could not be applied in our research and a new approach had to be developed.

To distinguish small water bodies, from size 1 pixel, we have tested several approaches. From all of the tested, two have provided satisfactory results – Principal Component Analysis (PCA) and Normalized Difference Snow Index (NDSI). Both complementary variants are based on radiometric enhancement of Landsat data. The principal components method is able to detect water bodies as such and also areas which contain a bigger amount of vegetation (wetland etc.). The method based on the NDSI index can only detect water bodies; however, it is able to detect even straight water streams which are mostly ignored by the first alternative.

The method and all the supporting procedures are designed for a free software system for digital data processing called MultiSpec (Purdue University, Indiana, USA). Its scheme is universal and can therefore be used in commercial software systems.

2.1 Water Bodies Detection through Principal Component Analysis

The Principal Component Analysis (PCA) is a statistical method, which is also used to enhance multispectral images. The major principle is to reduce the original number of spectral bands and suppress correlation between them without losing any information. The values of the principal components (of newly established bands) are defined as linear combinations of the original bands. Most of the information from the multispectral image is usually concentrated in the first three components. The first component tends to carry the information about the topography, others provide details about humidity, mineral composition etc., depending on the nature of the area. In other words, it is a “scene-wise” tool, the use and results of which are influenced by the area under scrutiny. As a result, each and every method based on the principal component analysis is a “case” method in fact, its wider use in the particular form depending on the similarity of areas in terms of how many classes of land cover, of what type and in what ratio the areas contain. Yet the principle remains the same, and the method is not difficult to adapt even for very different areas. So, the result of principle component analysis will depend on the choice of the area, on its heterogeneity in terms of land cover and relief. In brief, the PCA method offers a huge advantage for homogeneous areas; when applied to large and heterogeneous areas; it will obviously be less reliable and accurate.

The method proposed here is based on the general fact that virtually all relevant information about the structure and state of land cover is hidden in the data of the first five spectral channels from the Landsat TM / ETM+ sensor systems. Therefore, the parameters entered into the principal component analysis included five new, radiometrically highlighted “component images” called PC1 – PC5. Of the five images, PC1 contains the biggest amount of information. The higher the number of the component image, the less information and more noise areas it contains.
When the new components were analyzed, it was found that it is the PCA2 component that carries information about the water potential of the area in question. This means it is fairly easy to identify in the new image the individual water bodies which stand out in stark contrast to the rest of the land cover. It is also possible to detect correlation with humidity in the other categories of the land cover.

After separating PCA2 from the original image thematic recoding (tresholding) follows. The purpose of that is to split the image into two categories only: water bodies (target object) and background. This makes it necessary to define a threshold, or a boundary running between the two categories, which will be different for each model area, though.
In order to define or fine-tune the thresholds, we use the specific ability of the MultiSpec software to open the image in a different mode. The PC2 multispectral component image is normally shown as black-and-white, but MultiSpec can open it as a whole spectrum of colours rather than in several shades of the grey. The main advantage of this “thematic” image is that it can be recoded, which the multispectral image can't. The recoding is an interactive process in which all the pixels are analyzed in terms of what water objects they belong to. This way, the final threshold is defined which distinguishes water bodies from the rest of the land cover.
The supporting interpretation and tresholding is based on colour compositions (which are generally considered as a type of radiometric enhancement) and on interactive techniques providing a spectral enhancement of the visualized image. The image recoding and the identification of localities falling into the category of water bodies require the concurrently visualized RGB synthesis of the original multispectral bands of the image. The combination of the bands (usually TM4-TM5-TM7) makes it easier to define the threshold.

In some cases, the mask of the water bodies may contain pixels that belong to a different class of land cover and cannot be eliminated through the introduction of thresholds. Therefore, other methods need to be employed which will possibly eliminate these noise pixels from the mask. For example, these methods may include supervised classification, arithmetic operations with images, correction to the influence of relief using DEM (digital elevation model) etc. Unfortunately, no precise and universal method can be defined, and each area will need a different method to eliminate the noise pixels. The choice of the most efficient method and the threshold definition will mainly depend on the user's expertise.

2.2 Water bodies detection through image rationing

The first alternative was designed to detect any type of water bodies no matter what the quality of the water, whereas alternative two aims to detect water bodies of high quality with special relevance to the aquaculture technologies researched in the BOMOSA project.
The method proposed here is based on the radiometric highlighting of satellite data using a process called image rationing. It is especially used to reduce the influence of different exposures caused by variable relief structure and shadows cast on some places. But in this approach, carefully selected spectral images of low correlation are provided with various indices distinguishing and evaluating all the different classes and states of land cover. All of these approaches are “pixel-wise”, meaning that the pixel is categorized based on its own value, i.e. completely independent on the size and quality of the area in question.
In order to distinguish water bodies relevant to East-African landscape, we used the Normalized Difference Snow Index (NDSI) which is calculated using data from spectral channels Landsat TM / ETM+ in the following way:

$$NDSI = \frac{TM2 - TM5}{TM2 + TM5} \quad (1)$$

This index is primarily designed to detect and evaluate the snow cover (Dozier 1989, Dozier and Marks 1987, Salomonson and Appel 2004). At the same time, it is able to detect any water bodies whose reflectance is closer to that of snow than in Eastern Africa which abounds in turbid, overgrown and muddy waters. The index values are within an interval <-1; 1>. However, MultiSpec software is unable to operate these decimal numbers; the result would only contain values 0 and 1. Therefore multiplying the expression (1) with coefficient (according to rule of proportion) will stretch the data into the required 256 (8bit data) values scale. After index calculation the procedure of tresholding (water bodies mask) follows (see section in 2.1).

In order to gain as complex information as possible about water bodies in the area, it is advisable to add the masks resulting from both variants together. The method was tested on five areas in Kenya (The Rift Valley, Machakos, Thika, Kisumu and Bomet regions), three areas in Ethiopia (Lake Koka, Alagae and Awasa regions) and two in Uganda (Jinja-Kampala region, Kamuli district). The regions were selected with regards to the localization of BOMOSA plots, potential areas of interest and characteristics of Landsat data.

2.3 Application on another satellite data

Future application of remote sensing BOMOSA method for small water bodies detection, its dissemination and upscalling is highly dependent on the satellite data availability. The method was originally developed for Landsat satellite data. The discontinuity of Landsat mission in African region has appeared as crucial constraint in future use of method. The only solution is to complete archive data analysis by field observation and GPS measurements, or try another satellite data. As a substitution, Terra Aster and EO-1 ALI satellites were considered as another data sources.

Application of BOMOSA remote sensing method on Terra Aster data

Both approaches of the small water bodies detection – Principal Component Analysis and image rationing (Normalized Difference Snow Index) were tested on Terra Aster data. Processing of five scenes has revealed some limitations and constraints of Bomosa method application. Both approaches are generally characterized by lower ability to detect small water bodies. The use of NDSI index seems to be more universal and accurate. It was applicable in most cases (processing 6 Aster scenes); however in a comparison with the results of Landsat images analyses, the smallest water bodies are often not detected. Principal Component Analysis is “scene-wise” approach. The results are highly dependent on the nature of the area – heterogeneity of land cover types and their ratio in selected area. In most analyzed Aster images mixing water bodies with other land cover types – urban areas (especially cities) and forest vegetation is very often phenomenon. Therefore new methodological approach was tested, based on thematic recoding (tresholding) of AST 3N band. Aster band 3N records reflected near-infrared radiation (0.76 - 0.86μm) and is essential mainly in vegetation analysis. Dense and green vegetation reflects in, this spectral interval, significant amount of incident radiation (usually displayed by light hues of grey scale). Water objects appear dark. By very cautious recoding of this band, it is possible to extract water bodies. The principle of recoding is the same one, as in case of NDSI and PCA approaches.

Application of BOMOSA remote sensing method on EO-1 ALI data

Another data alternative could be ALI sensor on board EO-1. The Advanced Land Imager (ALI) was designed to demonstrate improved Landsat spatial and spectral resolution and has nearly the same characteristics. The scene size is 185 x 37 km, so nearly six times narrower, in a comparison with Landsat and data have been available since 2001.

Both approaches of small water bodies detection were tested. Because of 16-bit radiometric resolution of the data, Principal component analysis approach could not been used. The NDSI approach, based on thematic recoding of Normalized Difference Snow Index appeared as applicable. Instead of using channels TM2 and TM5 as in case of Landsat data, the formula was adjusted to EO-1 ALI data in following way (2), with respect to wavelengths of the channels:

$$NDSI = \frac{C4 - C9}{C4 + C9} \quad (2)$$

3 Results

Following figures display land cover distribution to provide basic information about analysed region and the final water bodies masks. Presented are all three satellite data types – Landsat, Terra Aster and EO-1 ALI.
Landsat data were provided for free by Global Land Cover Facility (GLCF), University of Maryland, USA, and EO-1 ALI by U.S. Geological Survey. The data have been used solely for education and research purposes of the BOMOSA project. Terra Aster was acquired from Land Processed Distributed Active Archive Centre data archive (USGS-NASA) and purchased by ENKI, o.p.s. Třeboň.

Landsat scene
Lake Koka region (Fig. 1.) is situated approximately seventy kilometres south-eastwards of Addis Abeba , Ethiopia and is situated on Landsat scene 168-054. Date of acquisition was on 5 December 2001. , scene size covers the area (62,7 x 74,1 km), 4646 km^2 .
For extraction of water bodies following parameters were used:
- PC2 component   Recoded (tresholded) mask of water bodies for digital numbers = <1; 74>
- NDSI   Multiplication coefficient C = 260
  Recoded mask of water bodies for digital numbers = <46; 255>

Some regions in Ethiopia (especially the north ones) are characterized by rugged topography. This is a limiting factor for accurate scene processing; steep slopes and deep canyon valleys cause the noise in the image – overshadowed parts of relief often appear as water bodies. Such topographic effects could be removed by methods, using digital elevation data and spatial analysis that is beyond the MultiSpec processing potential.

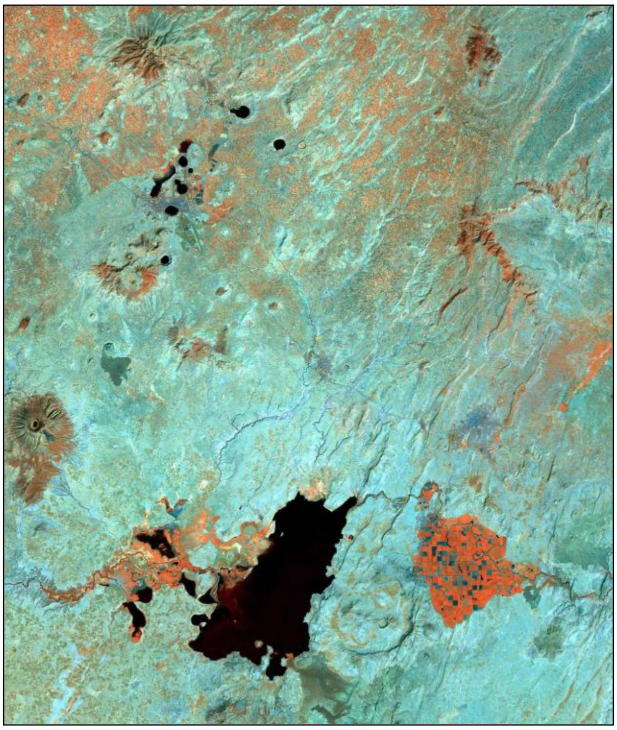

Fig. 1. RGB synthesis of Landsat channels 4-5-7 enhancing water bodies (dark objects), vegetation cover (red) and bare grounds (cyan)

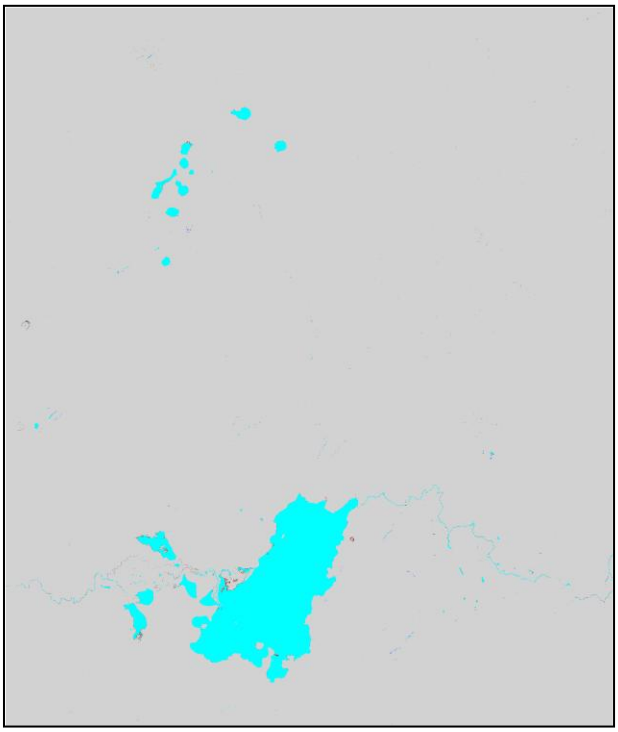

Fig. 2. Landsat scene - final mask of water bodies in Lake Koka region. Cyan indicates water detected by both approaches, blue – by NDSI approach and red- principal component approach.

Terra Aster scene

As a model area for Terra Aster data processing, Machakos area (seventy kilometers of southeast from Nairobi, Kenya) was selected (Fig. 3.). The scene was acquisited on 26 August 2004, scene size 60 x 60 kilometers. In case of processing Machakos scene nearly neither of both methods could be used. PC3 highly correlated not only with water, but nearly with all land cover types (mainly with humid vegetation and urban areas, in some cases with wet bare grounds as well). NDSI index detected only some of water bodies. Therefore the water mask was derived as a combination of NDSI approach and thematic recoding of AST 3N band.

The recoding thresholds are following:

- NDSI Multiplication coefficient C = 255
  Recoded mask of water bodies for digital numbers = <26; 255>
- AST 3N band Recoded mask of water bodies for digital numbers = <1; 53>

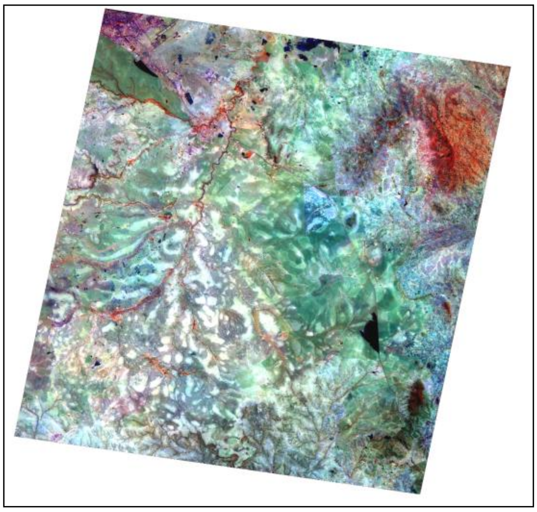

Fig. 3. RGB synthesis of Terra Aster channels 3-4-7 enhancing water bodies (dark objects), vegetation cover (red) and bare grounds (cyan)

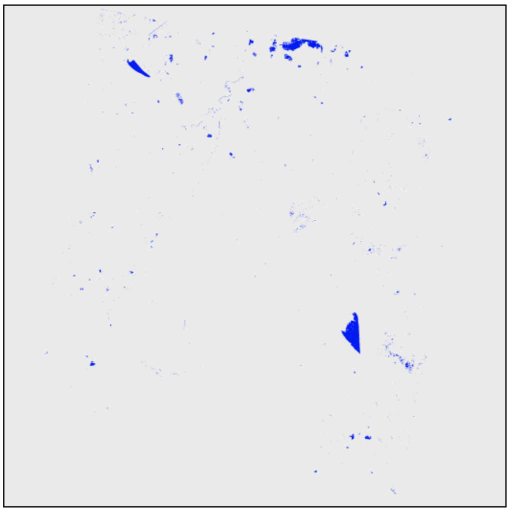

Fig. 4. Terra Aster scene - final mask of water bodies in Machakos region. The final mask is a combination of NDSI approach and tresholding of channel AST 3N.

EO1-ALI scene

As it was stated above, one of the possibilities to get up-to-data data for African region is using EO-1 ALI data, obtained from U.S. Geological Survey data archive. The main disadvantage of these data is that are resampled by cubic convolution. This interpolation method produces slightly smoother image appearance than the others, keeps the original geographical position, and nevertheless changes the original measured values of image. The scene of central part of Lake Kyoga was chosen as a model area (Fig. 5). The date of acquisition was on 11 June 2005, scene size is 185 x 37 km.

Used parameters: Multiplication coefficient C = 255
Recoded mask of water bodies for digital numbers = <55; 255>

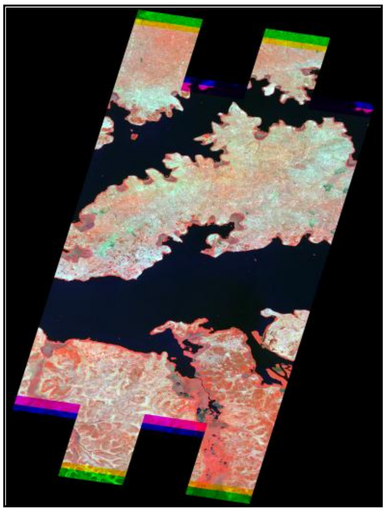

Fig. 5. Land cover of central part of Lake Kyoga in 11 June 2005. RGB synthesis of EO-1 ALI channels 6-9-4.

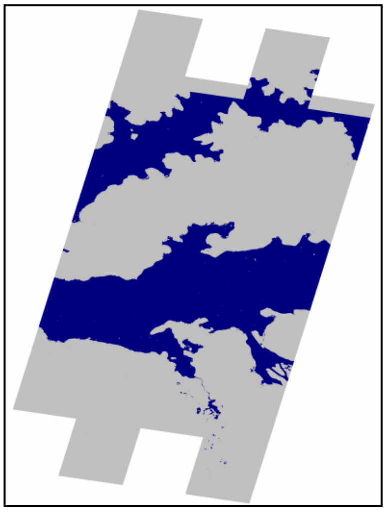

Fig. 6. Water bodies derived from EO-1 ALI data by NDSI approach.

4 Discussion / Conclusion

The method for detecting small water bodies was – with minor constraints – designed for Eastern Africa and the BOMOSA project. Its universal use and application to other areas cannot be guaranteed (as each type of land cover has a specific nature; similarly, different water bodies have different features and qualities). So what are the constraints of the method? The size of the water bodies that can be potentially detected by this method depends on the spatial resolution of the satellite data. In case of Landsat and ALI EO-1 data, the smallest discernible areas had the size of 0.08 ha. Clouds are another potential problem, especially the shadows cast by the clouds which are then mistaken for water bodies. But shadows are also a problem when it comes to the relief. Even some very steep slopes and bottoms of canyon valleys might be mistaken for water bodies, which make it rather difficult to detect small water bodies in areas with highly variable relief. Another factor potentially distorting the results is thick vegetation. The multispectral satellite systems scan the land cover only and cannot penetrate below the thick cover of plants. This makes it very difficult to detect water bodies in areas like this. Urban units present a potential problem as well as they normally contain a huge amount of mixels, and in order to eliminate them, other correction and more sophisticated procedures need to be employed. The method in any case is not able to evaluate water quality, temperature, water volume, or the shape of pond basin, cannot categorize water bodies. Therefore it is still essential to implement field work. The limits of usage of BOMOSA remote sensing method in the future can be seen in data availability For this moment if there is continuity of ALI EO-1 mission as a compensation for Landsat data, since December 2012 the Landsat Data Continuity Mission (LDCM) is the future.

The method was tested in ten regions in three African states. Nearly for each of the regions all of three different data types were available. This fact enabled to compare the sensitivity and accuracy of the method for water bodies detection. As it was mentioned, the best results were provided by Landsat data. The verification was firstly done by comparison with multispectral image. Taking into account the printing possibilities, it was not reasonable to present the same regions on all of three data types. The main restriction was the discriminability of the results in a hard copy form. However it was also essential to verify the method in field. Verification, based on these satellite water maps and their integration with GPS technology, enabled to track in terrain small water bodies (smaller than 1ha), identified by BOMOSA method. Main field works were held in Machakos and Kisumu area in Kenya and Kamuli District in Uganda, with satisfactory results. The field work was slightly limited by the access possibilities to water objects.

The method (Hesslerová, Šíma, Pokorný 2009) developed for the purposes of BOMOSA fish farming should mainly serve as a mapping tool for water bodies detection. The contribution consists in possibilities of evaluation “water potential” of selected regions. It was above the frame of the project to evaluate and process the whole areas of involved countries – Kenya, Uganda and Ethiopia and map all water bodies. However the method showed the way how it is possible to create water maps that could serve as a background and inputs into water and fishery management. The method uses all advantages that generally provide remote sensing data – is able to detect water bodies in huge areas (several thousands of sq km) in one moment of data acquisition, by one objective method. Can observe the remote areas, normally not accessible by road. And in addition, due to the large data archives and repeated regime of images acquisition is able to look back and monitor changes-over-time.

The outputs of the method are full-valued GIS layers (Geographical Information System). Integration of water bodies masks with both relevant GIS layers and GPS data field collection (e.g. road network, shopping centres, ethnical issues, etc) and use of basic GIS analyses (SQL querying, overlaying, selecting, calculating, summarizing,..) would enable better BOMOSA up-scaling. The relational database, gathering the information about water quality, fry, parasites occurrence, socio-economics, etc., would be an invaluable tool in plots management.

5 Acknowledgements

Many thanks to all African partners, involved in BOMOSA project for their support during fieldwork, feedbacks, comments, suggestions and reviews - Kenya Marine & Fisheries Research Institute, Sangoro Aquaculture Station, Kenya Ethiopian Institute of Agricultural Research, NFLARR, Sebetha,

Ethiopia, Egerton University, Egerton, Kenya, Ministry of Agriculture, Animal Industry and Fisheries, Dept of Aquaculture, Entebbe, Uganda, Moi University, Dept of Zoology, Eldoret, Kenya, Kenyan Ministry of Livestock and Fisheries Development, Sagana Fish Farm, Kenya, as well as our European partners from University of Natural Resources & Applied Life Sciences, Vienna, Austria, University of Bologna, Italy, Austrian Academy of Sciences, Institute for Limnology, IPGL-Office, Mondsee Austria.
Supported by: BOMOSA EC no. 032103, and Ministry of education, youth and sports of the Czech Republic, (NPVII, project, no. 2B06023).
The author also acknowledges the Global Land Cover Facility (GLCF), University of Maryland, USA as a main source of the dataset, as well as the U.S. Geological Survey. The data have been used solely for education and research purposes of the BOMOSA project.

6 References

Dozier, J., Marks, D. Snow mapping and classification from Landsat Thematic Mapper data. *Annals of Glaciology*, 9, 1989, p. 97-103.
Dozier, J. Spectral signature of Alpine snow cover from the Landsat Thematic Mapper. *Remote Sensing of Environment*, 28, 1989, p.9-22.
Hesslerová, P., Šíma, M., Pokorný, J. *BOMOSA remote sensing method for small water bodies detection.*ENKI, o.p.s., 2009, Třeboň, ISBN 978-80-254-5611-8.
Jensen, J.R. *Digital Image Processing*. *A Remote Sensing Perspective.* Prentice Hall, 2005, New Jersey.
Salomonson, V.V., Appel, I. Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment* , 89, 2004, p. 351–360.

Earth Explorer USGS http://edcsns17.cr.usgs.gov/EarthExplorer/
Global Land Cover Facility (GLCF), University of Maryland. http://glcfapp.umiacs.umd.edu/index.shtml
Earth Observing System (EOS) Data Gateway - data products by NASA and affiliated centres http://edcimswww.cr.usgs.gov/pub/imswelcome/.