NEAR REAL TIME FLOOD MONITORING TOOL

A. Ajmar [a], P. Boccardo[b], F. Disabato [b], F. Giulio Tonolo [a]

[a] ITHACA, via P.C. Boggio 61, 10138 Torino, Italy - (andrea.ajmar, fabio.giuliotonolo)@ithaca.polito.it
[b] DITAG - Politecnico di Torino, Corso Duca degli Abruzzi 24, 10135 Torino –franca.disabato@polito.it

KEY WORDS: Remote Sensing, Flood, Monitoring, MODIS

ABSTRACT:

Risk reduction and disaster respond experts working on flood related issues could greatly benefit from the availability of information such as the real extent of past flood events and the monitoring of flood extent during ongoing events. To respond to the afore mentioned needs, ITHACA is developing a tool devoted to the automated detection of flooded areas with a worldwide coverage, basically trough the identification of water bodies on the ground and the comparison to a reference water extent based on the analysis of a 10 year remotely sensed data archive. The main input data is satellite imagery, specifically the MODIS/Terra Surface Reflectance Daily products (MOD09GQK and MOD09GQ), characterized by a spatial resolution of 250 meters. These products allow to have a daily estimate of the surface spectral reflectance investigating two bands of the e.m. spectrum, the red one (620-670 nm) and the IR one (841-876 nm). The main advantage of this product is that it contains the surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering or absorption. The main disadvantage is that, being MODIS a passive sensor, the cloud coverage can be a relevant obstacle to the analysis. Therefore a cloud, cloud-shadow, and hill-shade masking procedure has to be used. Furthermore a multi-temporal approach has to be adopted in order to reduce the presence of no-data cells due to the cloud coverage. The temporal compositing algorithm is based on daily images representing water bodies on the ground, generated by the use of band ratios and threshold techniques. For each cell of the image, we calculated the number of days during which the relative area on the ground is classified as flooded. Defining a number of days threshold is therefore possible to state if the area is covered by water, as well as to provide a reliability parameter based on the cloud cover persistence. The comparison of the classified water bodies with average values related to the same period will finally allow to state with high probability if the area is flooded or normally covered by water. Preliminary results related to specific case studies will be shown.

1. INTRODUCTION

ITHACA is a non-profit association, envisioned as a centre of applied research developing IT products and services in support of emergency management, focused especially on early-warning and damage assessment phases.
Specifically for early-warning, ITHACA has developed and implemented an alert system based on precipitation analysis and related to historical flooded areas detection. The purpose of the tool, operational with a nearly global coverage and in near real-time, is to monitor actual precipitation rates, derived from a satellite platform, and to compare those with rainfall thresholds corresponding with historical flood events occurred on the field, detecting in that way extreme events.
In the field of early-impact and damage assessment, main objective of ITHACA activities is to provide georeferenced information (mainly in form of map products) about affected areas and population. Main final user of those elaboration is United Nations World Food Programme (WFP). Since the beginning of their activity (January 2007), ITHACA has been activated in occasion of more than 50 events, most of them related with flood caused by intense precipitations, cyclones, hurricanes, etc. More than 200 rapid mapping products were delivered.
Timely availability of geographic information about water bodies extension and, possibly, depth at an appropriate scale is fundamental for all response phases of the risk cycle. Both historical, to correctly identify the reference conditions or to produce scenarios based on historical events, and near real-time data, for the actual and precise identification of water covered areas, are thus essential. For that reason, ITHACA decided to activate a research activity finalized to the automation of the classification of water bodies from remote sensing images, with a dual objective:

- to process an huge volume of historical data, in order to obtain dynamic reference data, related to different climatic periods/seasons;
- to develop a near real-time monitoring system for flood events.

2. CLASSIFICATION OF WATER BODIES

2.1 Literature review

Actually, water bodies identification on a scene acquired by optical sensors installed on satellite platforms is based on simple but effective histogram threshold techniques; those techniques exploit the behaviour of water in the infrared bands, where those surfaces have high absorption rates. Similar approach is valid also for processing radar images, due to the fact that, in case of calm water hit by an incident microwave beam, the specular response dominates the returned signal.
As highlighted in previous articles, those methods are simple to apply but several disadvantages are also evident. Shadows, due to the local morphology or to the presence of clouds, are classified as water bodies and threshold values cannot be defined uniquely but adapted to the conditions at the moment of the acquisition.
For those reason, an extensive review of water bodies classification techniques presented in literature was conducted, in order to identify those capable to solve or minimize the above mentioned problem. The analyzed techniques are mainly based on indexes derived from differential band ratios, to make threshold values independent from image acquisition parameters. In particular, according to literature the use of the Normalized Differential Water Index (NDWI) is commonly used to identify and classify flooded areas. Unfortunately, a unique definition of this index was not found, probably due to its adaptation to the different characteristics of spectral sensors

mounted on satellite platform normally used for those applications.

The most diffuse definition of NDWI, described by Chowdary et al. (2008), McFeeters (1996), Chatterjee et al. (2005), Jain et al. (2006), Purba et Al. (2006), Hui et al (2008), is based on reflectivity in the green and near-infrared bands:

$$NDWI = (\rho_{GREEN} - \rho_{NIR})/(\rho_{GREEN} + \rho_{NIR})$$

This index reduces commission errors during classification, due to vegetation and bare soil classes. Zhuowei (2007), Hui et al. (2008), Fengming et al (2008) highlight the low reliability of this index in urban areas, proposing a Modified Normalized Difference Water Index (MNDWI) to minimize also errors due to the presence of shadows:

$$MNDWI = (\rho_{GREEN} - \rho_{SWIR})/(\rho_{GREEN} + \rho_{SWIR})$$

Huggel (2002) propose san NDWI definition based on reflectivity in the blue and near-infrared bands, especially conceived to the identification of mountain lakes:

$$NDWI = (\rho_{NIR} - \rho_{BLUE})/ (\rho_{NIR} + \rho_{BLUE})$$

Phadil (2006), Islam et Al. (2009), Sakamoto et al. (2007), Mori et al. (2009) use red and short-wavelength infrared (SWIR), to highlight the residual influence of humid soils:

$$NDWI = (\rho_{RED} - \rho_{SWIR})/(\rho_{RED} + \rho_{SWIR}) + 1$$

Gao (1996), DeAlwis et al. (2007) propose the below NDWI definition, for the identification of water saturated soils:

$$NDWI = (\rho_{NIR} - \rho_{SWIR})/(\rho_{NIR} + \rho_{SWIR})$$

Finally, literature review shows a vast heterogeneity of NDWI definition, with as a common component the use of differential ratios based on those bands exalting relative reflectivity differences of water spectral signature (visible and infrared bands).

After several experimental tries, we decided to adopt the afore mentioned MNDWI index, because of the higher capacity of shadow effects removal:

$$MNDWI = (\rho_{GREEN} - \rho_{SWIR})/(\rho_{GREEN} + \rho_{SWIR})$$

In the next paragraphs we describe the experimental procedures to verify MNDWI effectiveness, in several geographic areas and on the basis of available satellite remote sensing data; results are compared with methodologies based jointly on NDVI index and near-infrared thresholding, the approach initially used by the authors.

2.2 Available datasets

In order to perform an analysis on the extension of flooded areas on a regional scale, low spatial resolution data have been used.
Among available datasets, MODIS (Moderate Resolution Imaging Spectroradiometer) data have been chosen, because their spatial resolution (250-500 m) allows to detect flood event effects at a regional scale, while their available historical archive , which lasts from 2000 until present, and their temporal frequency (daily) are suitable to create an archive of flood events.

Two different MODIS products are used for the procedure:

- MODIS Daily Surface Reflectance products, generally available after 4 days from the acquisition and suitable for historical analysis;
- MODIS near real time products, available some hours after the acquisition and suitable for near-real time analysis.

MODIS Daily Surface Reflectance products, deriving from images acquired by Terra and Aqua satellites, have a spatial resolution from 250 (MOD09GQ and MYD09GQ) to 500 meters (MOD09GA and MYD09GA). 250 m spatial resolution products provide a daily estimate of the surface spectral reflectance for two bands, the red one (620-670 nm) and the IR one (841-876 nm), while MOD09GA and MYD09GA provide this information for 7 bands, 3 in the visible (620-670 nm, 459-479 nm, 545-565 nm) and 4 in the IR (841-876 nm, 1230-1250 nm, 1628-1652 nm, 2105-2155 nm) . The advantage of this product is that it contains the surface spectral reflectance as it would be measured at ground level, in the absence of atmospheric scattering or absorption.
The product is provided in a grid format, in particular in a ISIN grid (Integerized Sinusoidal grid, Figure 1). Each data file corresponds to a square of the grid. These data are provided in a hdf format and need a reprojection in a proper coordinate system (generally Geographic WGS84 system).

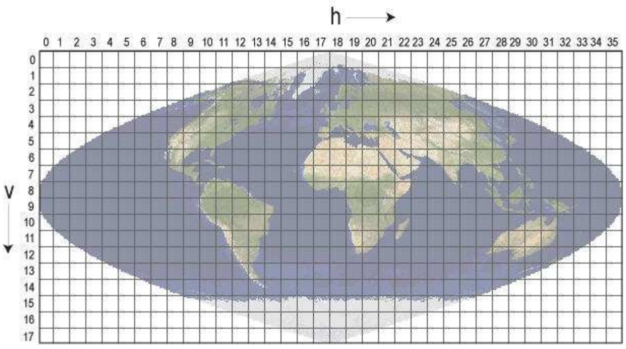


Figure 1 – MODIS ISIN grid projection and tiles.

The Rapid Response System is a web-based service making available to the users MODIS data in near real-time; it’s especially conceived to support emergency management, specifically for floods and fire events. MODIS Rapid Subsets grants daily availability of two images acquired on the same geographical area by Terra and Aqua satellite platforms. These images are geometrically corrected for some selected areas of the globe (**Figure 2**), with a spatial resolution up to 250m, and georeferenced in a geographic reference system referred to WGS84 datum. Those characteristics strongly reduces project pre-processing time.
These data are distributed in a compressed lossy raster format, badly affecting the classification process; nevertheless, the near real-time image availability is fundamental to built-up an operational emergency management process.
Relatively to spectral characteristics, the following additive synthesis are available, suitable for the extraction of clouds and snow masks and for the classification of water bodies:

- true colors, MODIS bands 1, 4, 3 (670 nm: 565 nm: 479 nm);
- false colors, MODIS bands 7, 2, 1 (2,155 nm: 876 nm: 670 nm).

For a limited subset, also the following synthesis is available:

- false colors, MODIS bands 3, 6, 7 (479 nm: 1,652 nm: 2,155 nm).

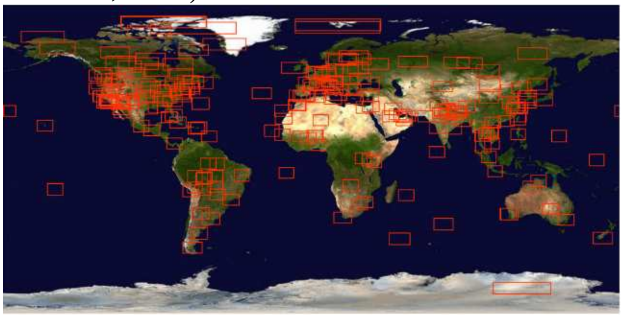


Figure 2 - The MODIS Rapid Response System available subsets.

2.3 Test results

The main test area for the procedure has been Bangladesh, one of the most flood-prone countries in the world, due to its geographical location, exposed to monsoons and cyclones, and morphology: about 60% of its territory is less than 6 m (19.7 ft) above sea level. Climate change is expected to create a 39% increase in flood-prone areas. According to several studies, once-in-20-year floods already are occurring about every four years. Rising salinity levels as brackish water inundates cropland could mean the loss of 659,000 m^3 of annual rice production. Meltwater from the Himalayas provides water to most of Bangladesh. Rapid glacier melt will mean more water flowing down the Ganges and Padma rivers in the monsoon months, causing more devastating floods. In the long term, as the water in the rivers disappears, the result will be more severe droughts.
On the whole country territory we performed an historical analysis aimed at the identification of frequently water covered areas, due to tidal effects or agricultural practices (rice fields), and consequently of the areas more subjected to floods.
With the objective to produce an historical analysis as complete as possible, we utilised MODIS Surface Reflectance products; due to the availability of an archive of images available since year 2000, water bodies distribution and extent were classified for a period of nearly 10 years.
The previous mentioned analysis has been executed following two different approaches, in order to have a comparison dataset.
The first methodology includes:

- MNDWI index calculation, using MODIS spectral bands 4 and 6, with a spatial resolution of 500 m;
- Masking of areas covered by clouds or snow, through thresholding procedures in the visible bands.

The second methodology includes:

- NDVI index calculation, using MODIS spectral bands 2 and 1, with a spatial resolution of 250 m;
- Masking of bare soil areas, through thresholding procedures in the near-infrared bands;
- Masking of areas covered by clouds or snow, through thresholding procedures in the visible bands.

Furthermore, cloud shadows represent a big problem for the classification of the flooded areas, because they could be interpreted by an automated algorithm as water areas, as very dark shadows have spectral characteristics quite similar to those of water bodies, while lighter shadows can be easily confused with dark vegetation. (Luo et al., 2008). For this reason it has been decided to first classify each daily MODIS file and then create a temporal composite of the classification results. The product of the compositing process is a file reporting the number of times that a pixel has been classified as covered by water. In order to create an automated process capable to provide effective information for the operator (flooded and not flooded areas), it has been assumed that the probability that a pixel is covered by the shadow of a cloud for more than a threshold-percentage of the compositing period (for example for more than 1 day over a 10 days compositing period) is extremely low. Because of this all the pixels that result to be covered by water for more than this threshold period (≥ 2) will be considered as areas affected by water. The same approach may be used for the cloud coverage.

Both applied procedures produced comparable results, in terms of classification accuracy, over Bangladesh flat areas. For this reason, as the objective was to develop a monitoring system of floodable areas, we preferred to adopt the procedure that uses higher spatial resolution data sources, the one based on thresholding techniques using both NDVI index and infra-red bands.

The same methodologies for a near real time use have been also tested in Mozambique and in Pakistan, areas having a morphology significantly different from Bangladesh. The objective was to verify not only the correct execution of the process but also its applicability on several different geographic areas. In those cases we analyzed 73 scenes for south-eastern Africa (period from 1 January 2007 to 16 March 2007) and 93 for Pakistan (period from 1 June 2005 to 9 September 2005).
We tested several MODIS band combinations for the NDWI calculation, according to the different definitions found in literature:

- bands 1-7 (Red-NIR);
- bands 2-7 (NIR-SWIR2);
- bands 4-6 (Green, SWIR1);
- bands 3-6 (Blue-SWIR1).

The first two combinations were tested because they require to download just the MODIS Rapid Response System additive synthesis 7-2-1, significantly reducing pre-processing times; but, to be effective, they need the integration with NDVI index to mask vegetated areas. Bands 4-6 combination resulted providing better results than 3-6 combination, having an higher discrimination capacity for shadowed areas. Pakistan complex morphology and high and persistent cloud coverage requested to modify both threshold value and composite period to minimize the commission errors while classifying water bodies.

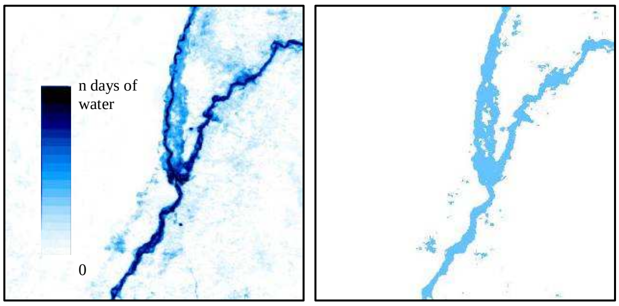
2.4 Implemented algorithm

In the case study of Bangladesh, in order to detect the extension of flooded areas, an automated procedure has been implemented. This procedure performs the following steps:

- automated detection of available satellite imagery, according to date and area of interest;
- eventually data mosaicking and reprojection;
- data classification;
- data compositing;
- extraction of flooded areas.

The first step of the procedure is to select MODIS Surface Reflectance satellite data from a local archive given a geographical extent and the date of the past flood event (input).
After the selection of the data, these are eventually mosaicked (if the area of interest fall in more than one square of the grid) and reprojected. These actions are performed by means of the

use of the batch version of the MODIS Reprojection Tool (MRT), the software provided by the Land Processes (LP) Distributed Active Archive Center (DAAC), in order to mosaic and reproject MODIS data.
The classification and time compositing of the flooded areas are performed by techniques described in the previous paragraph (Figure 3). The whole procedure is developed in the IDL (Interactive Data Language).



a) b)

Figure 3 – Time composition (a) and result of classification errors masking using the temporal threshold approach (b).

As already mentioned, aiming to produce an automated tool in support of emergency management, the capability of providing timely information is a crucial factor. With the aim of evaluating processing times over significantly extended geographic areas, we simulated the processing workflow of 4 adjacent MODIS scenes (each one having a size of 67 MB), for a period of 108 subsequent days. Times for geometrical correction were also estimated, in order to simulate the worst case scenario, where no Real Time Subsets are available. Results shows that downloading (5 hours) and georeferencing (4 hours) phases are the most critical, whether classification (1.5 hours) and temporal composite production (0.5 hours) require significant lower processing times.

3. FLOOD MONITORING TOOL

3.1 Objective

Encouraging results of water bodies classification test, both in terms of quality and performances, even while processing huge datasets as the ones needed to cover a whole country, brought us to the decision of orienting the research activity to provide an operational application.
We developed an operative tool to provide extension and localization of water bodies at continental scale and in near real-time, discriminating between potential flooded areas and reference water, those estimated in function of the specifying analysed season.

3.2 Near real-time monitoring tool

The classification methodology described in the above paragraphs, including the utilization of temporal composites, allows to identify spatial and temporal distribution of water bodies on the processed area and for the n days considered in the analysis.
This output is not sufficient for an operational instrument finalized to flood monitoring, as a reliable reference water body extent, dynamically adapted to the specific analyzed season, is mandatory for identifying flood effects in an effective way.
Then, a pre-processing phase finalized to collection, archive and elaboration of historical data has to be developed; the elaboration produces a database table where, to each portion of the discretized area of interest (generally referred to the single pixel of the original image) are associated values related to presence of water or cloud coverage for the entire analyzed period.
Subsequently, statistical analysis performed on previously produced data allow to extract value-added information, such as mean and standard deviation values for any portion (pixel) of the area, calculated on the entire historical dataset and, if relevant, subdivided per climatic and agricultural seasons, in order to consider trends related with precipitations and human activities.
Finally, the availability of such information allows a near real-time comparison of actual (last n days) ground conditions in relation with historical reference means (Figure 4), to distinguish between real flooded areas and normal or managed water extents.

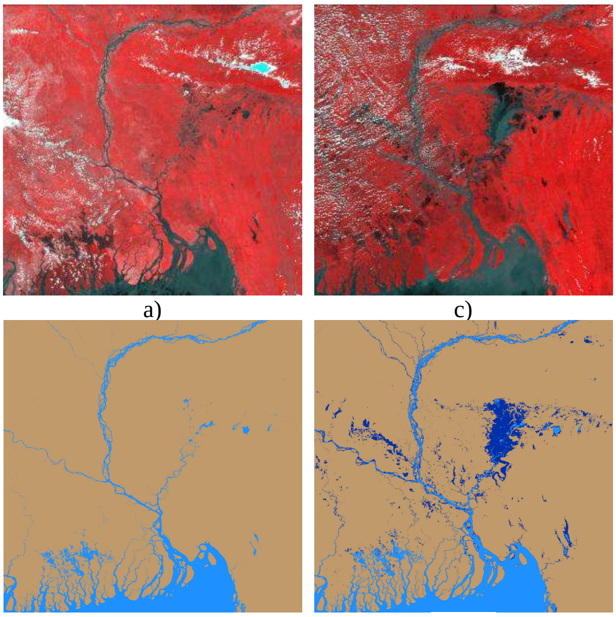


Figure 4 – Bangladesh: historical reference situation (a, MODIS false colour) and classification of water bodies (b), compared with near real-time situation during a flood event (c, MODIS false colour) and classification of water bodies (d, in light blue the reference water and in dark blue the additional water present during a flood event).

3.3 Case-study: Bangladesh

A beta version of the monitoring system has been implemented and validated for the entire dataset covering Bangladesh. Results has been qualitatively validated, being suitable ground-truth data difficult to find.

4. CONCLUSIONS

Research activity object of this paper is based on a previous activity of the same working group, finalized to the identification of the historical maximal extent of water bodies, related with past flood events. Implemented updates, based on the experimentation previously described, allowed to set up an application with suitable characteristics for emergency management, in particular for monitoring in near real-time wide geographic areas in order to detect water bodies and to alert for possible floods.

Before the system can be considered fully operative, some open issues must be solved:

- MODIS Rapid Subsets not always have a geographical coverage suitable for covering wide geographic areas, such as big watershed or entire continents. That makes necessary the use of Real-Time products that are not geometrically corrected and that require additional pre-elaboration time;
- Mainly in case of wide geographic areas, and for a continuous monitoring service, the design and implementation of a grid computing system is considered the solution for reducing significantly processing times and thus providing timely results;
- We must quantitatively evaluate the influence of the lossy compression of MODIS Rapid Response System product on classification accuracy;
- Tests on areas with complex morphology, such as the one executed on Pakistan territory, highlighted the necessity of modeling the effects of shadows due to altimetry variation, in order to minimize commission errors.

Once consolidated the procedure and solved the open issues, we plan to develop a web-GIS application dedicated to the visualization of system results. The application will include some basic geoprocessing tools, such as the possibility to estimate potentially affected population.

References

Ackerman, S., Strabala, K., P. Menzel, P., Frey, R., C.Moeller, C.,Gumley, L., Baum, B., Wetzel Seemann, S., Zhang, H., 2006. Discriminating clear-sky from cloud with modis algorithm theoretical basis document (MOD35), Version 5.0.

Chowdary, V.M., Chandran, R. V., Neeti, N., Bothale, R.V., Srivastava, Y.K., Ingle, P., Ramakrishnan, D., Dutta, D., Jeyaram, A., Sharma, J.R., Singh, R., 2008. Assessment of surface and sub-surface waterlogged areas in irrigation command areas of Bihar state using remote sensing and GIS. *Agricultural Water Management*, Elsevier, vol. 95(7), pp. 754-766.

Disabato, F., 2008. Classificazione automatica di aree alluvionate. *XII Conferenza Nazionale ASITA, L'Aquila*.

Easton, Z. M., Dahlke, H. E., Philpot, W. D., Steenhuis, T. S, DeAlwis, D. A., 2007. Unsupervised classification of saturated areas using a time series of remotely sensed images. *Hydrology and Earth System Sciences*, Volume 11, Issue 5, 2007, pp.1609-1620.

Fadhil, A.M., 2006. Environmental Change Monitoring by Geoinformation Technology for Baghdad and its Neighboring Areas. *Proceeding of the International Scientific Conference of Map Asia 2006: The 5th Asian Conference in GIS, GPS, Aerial Photography and Remote Sensing. Bangkok, Thailand*.

Fengming, H., Bing X., Huabing H. , Qian, Y., Peng G., 2008. Modelling spatial-temporal change of Poyang Lake using multitemporal Landsat imagery. *International Journal of Remote Sensing*, Volume 29, Issue 20, pp. 5767 – 5784.

Gao, B.C., 2006. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, Volume 58, Issue 3, pp. 257-266.

Huggel, C., Kääb, A., Haeberli, W., Teysseire, P., Paul, F., 2002. Remote sensing based assessment of hazards from glacier lake outbursts: A case study in the Swiss Alps. *Canadian Geotechnical Journal*, Volume 39, Number 2, pp. 316–330.

Islam A.S., Bala, S.K., Haque, A.,2009. Flood Inundation map of Bangladesh using MODIS Surface Reflectance data. *Proceedings of the 2nd International Conference on Water and Flood Management (ICWFM-2009)*, vol-2(ISBN: 984-300-000303-6).

Luo, Y., Trishchenko, A., Khlopenkov, K.V., 2008. Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment*, Volume 112, Issue 12, pp. 4167-4185.

Sakamoto, T., Nguyen, N. V., Kotera, A., Ohno, H., Ishitsuka, N., Yokozawa, M., 2007. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sensing of Environment*, Volume 109, Issue 3, pp. 295-313.

Thenkabail, P. S., Schull, M., and Turral, H., 2005. Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. *Remote Sensing of Environment*, Volume 95, Issue 3, pp. 317-341.

Zhuowei, H.U., Gong, H., Zhu, L., 2007. Fast flooding information extraction in emergency response of flood disaster. *ISPRS Workshop on Updating Geo-spatial Databases with Imagery & The 5th ISPRS Workshop on DMGISs*.

Aknowledgements

MODIS data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (lpdaac.usgs.gov).