INUNDATED AREA DELINEATION USING MODIS DATA: TOWARDS A GLOBAL SCALE GEO-DATABASE OF FLOOD EVENTS

P. Villa\textsuperscript{a,b}, M. Gianinetto\textsuperscript{a,c}

\textsuperscript{a} Politecnico di Milano University, DIIAR Department – Remote Sensing Laboratory, Piazza Leonardo da Vinci 32, 20133 Milano, Italy – (paolo.villa, marco.gianinetto@polimi.it)

\textsuperscript{b} Commission IV, WG IV/8, Commission VII, WG VII/5, Commission VIII, WG VIII/2

\textsuperscript{c} Commission III, WG III/1, Commission VII, WG VII/6, Commission VIII, WG VIII/2

KEYWORDS: Natural hazards, Flood maps, MODIS, Geo-databases, Less developed countries, Low resolution data

ABSTRACT:
The availability of global and accurate information is the primary factor affecting the possibility of planning and managing effective disaster response strategies, above all in less developed countries. The second determinant factor that avoids the full spreading of remote sensing technologies is cost-effectiveness and steadiness of results. This paper illustrates a straightforward method for rapid retrieval of inundation maps at regional and global scale by processing MODIS data with the Spectral-Temporal Principal Components Analysis and Digital Terrain Model filtering. Case studies are presented for three different vulnerable regions in developing countries struck by a severe river flood during the last year (2005, from spring to fall): India, Pakistan and Romania. For all the events studied it was obtained an overall accuracy greater than 95% and a kappa coefficient grater than 0.70, demonstrating this methodology is very accurate in mapping inundated areas. Moreover, the integration with vector data (such as roads, railways or urbanized areas) may be used to fast detect infrastructure damages at regional and global scale. This work is the first step to develop a global geo-database of flood-affected areas, a basic tool for helping public administrators in efficiently managing natural hazards. This is especially useful for less developed countries, which unfortunately suffer the heaviest damages because of the high density of population and the scarcity of prevention and rapid response strategies.

1. INTRODUCTION

During the last 20 years a group of determining factors has made flood events more and more dangerous for economy and human life, both for developed countries and, above all, for developing and less developed countries. Just think at the recent hurricane Katrina that struck New Orleans in September 2005, in the healthy U.S. (Kiage et al., 2005), or the Asian Tsunami that devastated the Indochinese islands in December 2004 or, again, the cyclic monsoonal floods affecting Bangladesh almost every year (Sanyal and Lu, 2004). These events pose a global treat primarily for less developed countries, due to the high density of population, the lack of response strategies and the difficulties in prevision and prevention that are made worse by the scarcity of international cooperation (again think about the delay in the tsunami alert and the controversies thus aroused). All this facts bring to the conclusion that, both for developed and non developed countries, there is a strong need for a global coverage of flood related events. If not, the rise in flood casualties and flood damages amount, which unfortunately is very manifest in the data of the last 20 years, could became too high a cost to pay for international coordination and cooperation chronic inadequacy. Remote sensing flood monitoring capabilities are well known for prevention issues, but are not yet fully exploited for flood mapping and post flood damage assessment, especially at regional and global scale (Gianinetto and Villa, 2006; Villa and Gianinetto, 2006).

Today, the increasing monitoring capabilities of low-resolution satellites (Barton and Barthols, 1989) could help the post flood problem handling through the creation of a global geo-database of flooded areas, constantly updated and available for the whole globe. With the new generation of low-resolution sensors operative from 1999, among which Terra and Aqua MODIS (250 m to 1,000 m ground resolution), NOAA/AVHRR (1,100 m to 4,400 m ground resolution), SPOT/VEGETATION (1,000 m ground resolution) and Envisat/MERIS (300 m ground resolution), satellite multispectral data are now available on a daily basis (twice a day for MODIS data) all over the globe, thus helping to solve the problem of data availability. The role of research is therefore to guarantee the practical usefulness of this huge amount of data, that is to provide the institutional users with prompt and efficient tools to handle the post crisis phase of natural hazard (among which a large part is related to floods) in order to help the displacement of people and to minimize the effects of damages to the economy of the areas involved.

2. FLOODS AND REMOTE SENSING

The basic idea of assessing flood damages through the use of satellite’s observations is founded on the well known mapping capabilities of multispectral sensors (Bryant and Gilver, 1999). In the last decade several kind of remote sensing flood analysis have been performed using data collected with a wide range of sensors. On the basis of past experiences it is possible to outline the ideal sensor’s characteristics as:

- high revisit time;
- spectral richness (visible, near-infrared, shortwave-infrared, thermal-infrared);
- multi-resolution capabilities;
- accurate radiometric and geometric correction;
- low cost;
- ready availability.
The MODIS sensor, on board of NASA’s Terra and Aqua EOS satellites, fulfills all those requirements. The Terra satellite, formerly EOS/AM-1, has a sun-synchronous circular orbit with descending antemeridian nodal crossing and daily revisit. The Aqua satellite, formerly known as EOS/PM-1, also has a sun-synchronous orbit with daily revisit and afternoon equatorial crossing. Therefore, the result is the daily acquisition of two images of the same place: one in the morning and one in the afternoon. This imaging capability becomes very useful when analysing natural hazards, such as floods, with relatively fast response and the steadiness of analysis.

Summarizing, MODIS data are:

a) available two or three times a day for each point of the Earth (depending on the coverage area needed);
b) their 36 spectral bands span over a large part of the electromagnetic spectrum with a good spectral resolution (see Table 2);
c) are available at different spatial resolution (from 250 m to 1,000 m);
d) present a very high quality radiometric correction and are provided with Ground Control Points (GCPs) data to perform an accurate geocoding;
e) are free of charges;
f) are promptly available from File Transfer Protocol (FTP) web servers.

3. FLOODS AND DEVELOPING COUNTRIES: CASE STUDIES

The main objective of this paper is to show how it is possible to enrich remotely sensed derived products with vector Geographic Information System (GIS) dataset for developing a global geo-database of flood events at small scale (Brivio et al., 2002). Obviously, the main disadvantage of such a tool is the low detail level achievable, but, on the other hand, it has the great advantage of the fast response and the steadiness of analysis.

### Table 1. Global statistics for floodings damages in the years 1985-2005 (from Dartmouth Flood Observatory, 2006).

<table>
<thead>
<tr>
<th>Year</th>
<th>Average Duration (days)</th>
<th>Casualties</th>
<th>Total Affected Region (sq km)</th>
<th>Total Damage (USD)</th>
<th>Average Flooded Area (ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>6.70</td>
<td>3,034</td>
<td>5,831,709</td>
<td>$2,168,237,782</td>
<td>1,136,432,24</td>
</tr>
<tr>
<td>1986</td>
<td>13.28</td>
<td>1,554</td>
<td>8,862,982</td>
<td>$2,157,098,463</td>
<td>1,824,274,8</td>
</tr>
<tr>
<td>1987</td>
<td>8.46</td>
<td>3,247</td>
<td>1,459,016</td>
<td>$2,667,700,000</td>
<td>336,298,77</td>
</tr>
<tr>
<td>1989</td>
<td>8.95</td>
<td>9,838</td>
<td>8,620,018</td>
<td>$1,673,442,200</td>
<td>725,397,27</td>
</tr>
<tr>
<td>1990</td>
<td>7.88</td>
<td>4,159</td>
<td>14,931,686</td>
<td>$4,704,224,478</td>
<td>1,111,468,46</td>
</tr>
<tr>
<td>1991</td>
<td>6.10</td>
<td>10,180</td>
<td>17,685,033</td>
<td>82,365,740,000</td>
<td>259,173,63</td>
</tr>
<tr>
<td>1992</td>
<td>11.12</td>
<td>8,720</td>
<td>13,078,925</td>
<td>$24,929,269,900</td>
<td>540,323,90</td>
</tr>
<tr>
<td>1993</td>
<td>8.12</td>
<td>7,453</td>
<td>35,102,006</td>
<td>$2,190,860,000</td>
<td>866,234,84</td>
</tr>
<tr>
<td>1994</td>
<td>8.64</td>
<td>6,401</td>
<td>8,574,572</td>
<td>$13,242,094,000</td>
<td>327,628,33</td>
</tr>
<tr>
<td>1995</td>
<td>12.10</td>
<td>7,625</td>
<td>48,573,631</td>
<td>$40,860,092,000</td>
<td>579,300,00</td>
</tr>
<tr>
<td>1996</td>
<td>8.54</td>
<td>6,210</td>
<td>12,798,803</td>
<td>$12,196,866,000</td>
<td>222,528,57</td>
</tr>
<tr>
<td>1997</td>
<td>9.06</td>
<td>7,819</td>
<td>6,115,340</td>
<td>$10,369,926,000</td>
<td>130,011,22</td>
</tr>
<tr>
<td>1998</td>
<td>9.88</td>
<td>23,896</td>
<td>43,324,701</td>
<td>$28,746,681,175</td>
<td>1,053,818,27</td>
</tr>
<tr>
<td>1999</td>
<td>10.96</td>
<td>33,792</td>
<td>57,129,721</td>
<td>$55,222,677,688</td>
<td>609,772,19</td>
</tr>
<tr>
<td>2000</td>
<td>8.74</td>
<td>10,664</td>
<td>59,228,261</td>
<td>$943,196,83</td>
<td>609,772,19</td>
</tr>
<tr>
<td>2001</td>
<td>8.35</td>
<td>5,623</td>
<td>37,973,069</td>
<td>$13,217,061,635</td>
<td>943,196,83</td>
</tr>
<tr>
<td>2002</td>
<td>9.31</td>
<td>4,422</td>
<td>20,536,442</td>
<td>$196,165,56</td>
<td>130,011,22</td>
</tr>
<tr>
<td>2003</td>
<td>12.98</td>
<td>4,554</td>
<td>21,667,574</td>
<td>$460,730,23</td>
<td>196,165,56</td>
</tr>
<tr>
<td>2004</td>
<td>13.68</td>
<td>173,474</td>
<td>51,467,552</td>
<td>$7,553,518,00</td>
<td>460,730,23</td>
</tr>
<tr>
<td>2005</td>
<td>13.29</td>
<td>10,140</td>
<td>19,367,962</td>
<td>$82,249,650,00</td>
<td>207,870,84</td>
</tr>
</tbody>
</table>

### Table 2. MODIS spectral bands specifications.
Thus, a real time updated global database could be very useful and effective especially for developing countries, where the priority is a fast response to an hazard rather than the local detail of the analysis. Furthermore, a detailed local study can always be performed in a subsequent step using higher resolution data.

The flood mapping methodology proposed has been already tested by the authors on Landsat/TM, Landsat/ETM+ and NOAA/AVHRR data. Now the same method was applied applied to low-resolution MODIS data collected during year 2005: this is the first step for evaluating the concrete feasibility of the global geo-database cited above.

Three tests sites struck by a severe flooding during 2005 were chosen, all of them located in less developed countries. They are:

a) Uttar Pradesh region, India (flood event of September 2005, rivers Gange and Ramganga);
b) Kabul river valley, Pakistan (flood event of June 2005, river Kabul);
c) Timis county, Romania (flood event of April 2005, river Timis).

3.1 Uttar Pradesh flood

During the last days of September 2005, the Northern part of India suffered the consequences of monsoon rains, that caused the flooding of river Gange and tributaries, besides landslides especially in Muzaffarnagar district (Uttar Pradesh), resulting in a total of 23 casualties and 35 villages involved over an area of 308,700 km² (Anderson and Brakenridge, 2005a).

3.2 Kabul river valley flood

The event regarding Pakistan took place near the border with Afghanistan, in the North of the country, where the river Kabul reached peak flow levels in consequence of the unusually strong snowmelt on Hindu Kush mountains and heavy rains; the inundation of the plains surrounding Peshawar city caused over 50,000 people displaced and more than 1,100 houses destroyed (Anderson and Brakenridge, 2005b).

3.3 Timis county flood

The last event studied is the one that involved the Timis county (Southern part of Romania, near the border with Serbia and Montenegro), were the worst flood in the last 50 years was experienced. Extensive damage were reported: crop fields surrounding Timisoara city were destroyed, about 350 bridges were damaged, over 3,700 people homeless and economic losses valued over 500 million U.S. Dollars (Anderson and Brakenridge, 2005c).

3.4 Flood mapping

For each of the regions cited above, a pair of Terra/MODIS Calibrated Radiiances Level 1B scenes, provided by U.S. National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Data Gateway, were used for the retrieval of inundation maps, Spectral bands 1 and 2 (resampled from 250 m to 500 m ground resolution) and spectral bands 3, 4, 5, 6 and 7 (500 m ground resolution), showing a spectral content similar to that of Landsat/TM and Landsat/ETM+, were used (Table 2).

The closest pre-flood cloud free image and the first useful post-flood image were used for the detection of changes in land cover before and after the flood.

The following data were collected:

a) Uttar Pradesh dataset:
   1) Pre flood scene: August 28th, 2005
   2) Post flood scene: September 27th, 2005

b) Kabul river valley dataset:
   1) Pre flood scene: June 17th, 2005
   2) Post flood scene: June 26th, 2005

c) Timis county dataset:
   1) Pre flood scene: April 5th, 2005

Further available Digital Terrain Model (DTM) derived from the resampling of the Shuttle Radar Topography Mission (SRTM version 2) 3 arcsec DTM were supplied by United States Geological Survey (USGS).

Pre-flood and post-flood MODIS scenes were first georeferenced in the UTM-WGS84 reference system using the GCPs enclosed in the original HDF files and then atmospherically corrected using a low resolution MODTRAN model combined with aerosol retrieval (Berk et al., 1998).
Ground-reflectance derived images were then generated and after a visual inspection it was decided to exclude from the successive processing the MODIS spectral band nr.5 (1,230 – 1,250 nm), due to its poor radiometric quality (low Signal-to-Noise Ratio) and abundance of missing scan lines. Therefore, for each of the test sites, a synthetic 12-band file (6 each from the pre and post-event scenes) was created and to it the Spectral-Temporal Principal Components Analysis (ST-PCA) was applied (Fung and LeDrew, 1987; Ingebritsen and Lyon, 1985). For a detailed description of the data processing see Gianinetto et al. (2006).

The results of PC rotation were inspected to find the PC band that better separated inundated areas from non-inundated areas. Finally, logical filtering was applied with the aid of a slope chart derived from the DTM, labelling pixels as inundated if both PC band and terrain slope fell into a pre-determined range of threshold values, as shown below:

a) Uttar Pradesh dataset:
1) PC band chosen: PC3
2) PC threshold: -1300.00 (maximum value)
3) DTM slope threshold: 5.0 %

b) Kabul river valley dataset:
1) PC band chosen: PC5
2) PC threshold: +300.00 (maximum value)
3) DTM slope threshold: 5.0 %

c) Timis county dataset:
1) PC band chosen: PC2
2) PC threshold: +1250.00 (minimum value)
3) DTM slope threshold: 5.0 %

3.5 Geo-database integration

In the recent years two new techniques have completely changed the approach to geographic information delivering: 1) Global Positioning System (GPS); and 2) Web-Geographic Information Systems (Web-GIS).

For the purpose of providing maps (both thematic and metric) to developing countries, GPS is devoted to the acquisition of vector and GIS information to integrate the spatial databases (DBs). This task can be easily carried out by a GIS data logger palm receiver, which allows to collect georeferenced features (e.g. points, lines or polygons).

The cost of such a GPS system (data-logger, related facilities and software for data processing) is in the order of a few thousands of Euros and their accuracy is in the range of 0.5 – 5 meters, depending upon the type of receiver (single frequency, double frequency, phase measurement) and the complexity of signal post-processing (e.g. signal differentiation with respect to that acquired by a master station, broadcast differential correction).

Collected vector layers and inundation maps derived from satellite’s observations have been thought as the initial data constituting a spatial DB, which will be then integrated by adding up further information. Each object in the database is linked to an attribute table, specifying some important characteristics of it. The attribute table is made up by a set of attributes which are common to all possible features. Then specialized attributes are introduced for particular kinds of features (e.g., in case of roads, attributes describing the class of the road, which kinds of vehicle can run on it and the like should be introduced).

Regardless the data distribution, Web-GIS techniques have changed the modality to deliver geographic information because they allow to access data to a huge number of people from everywhere a Web connection is available. Their use in developing countries is really important for two purposes: 1) the possibility of maintaining and delivering data with a very low cost of infrastructures (only some workstations, Web-GIS server software, a few operators and an office to guest them are strictly needed) and the real time communication with the other Countries of the World.

Figures 4 and 5 show an example of spatial DB for the Uttar Pradesh region (India) test site. Here the DB is composed of satellite derived flood maps (yellow colour) and GIS vector data (e.g. state boundaries, disctrict boundaries and main cities, important railroads and main roads) already available.

4. ACCURACY ASSESSMENT AND DISCUSSION

Flood maps were produced for all the three test sites and validated using as ground truth the flood extension maps provided by NASA’s sponsored Dartmouth Flood Observatory at Dartmouth College (Hanover, USA) (Anderson and Brakenridge, 2005a; Anderson and Brakenridge, 2005b; Anderson and Brakenridge, 2005c).
The validation assessed an overall accuracy greater than 95% for all the three dataset (97.87% for the Uttar Pradesh dataset, 95.76% for the Kabul river valley dataset and 98.78% for the Timis county dataset) and a kappa coefficient greater than 0.70 for all the three dataset (0.77 for the Uttar Pradesh dataset, 0.73 for the Kabul river valley dataset and 0.70 for the Timis county dataset). Tables 3, 4 and 5 show the error matrices calculated for the three test sites. As expected, non-flooded areas present very few errors, both in commission and in omission, while the flooded areas reach omission error values of 25 to 34 percent, that is to say that, according to the aim of the study and to the resolution of input data, this methodology has performed very well on a regional scale, though slightly underestimating the maximum inundation area.

5. CONCLUSIONS

The rising amount of damages, social and economic, due to natural hazards demand the implementation of a prompt and effective response strategy from scientific research and technology. In the case of floods, the creation of a global geo-database could help managing more effectively the post crisis phase of displacement and reconstruction, above all in less developed countries.

The ST-PCA mapping methodology proposed, already successfully applied for local scale analysis, has proven to be accurate also at regional scale using low resolution MODIS data. This can be the first step to develop a global geo-database which integrates satellite derived mapping information with land-cover, infrastructural, social and economic data, in order to perform a deep analysis of flood events dynamics and its consequences.

Last but not least, this mapping technique could become in the near future a global tool for environmental monitoring (not only flood hazards), thus helping public administrators in efficiently managing natural hazards.
ACKNOWLEDGEMENTS

The present work owes a debt of gratitude to NASA EOS Data Gateway, which provided the FTP transfer of MODIS data; thanks to the work of the webmaster Chao-Hsi Chang and the Responsible NASA Official Medora Macie. Special thanks to the always very kind and helpful Dartmouth Flood Observatory at Dartmouth College, especially to Robert Brakenridge and Elaine Anderson, for providing their flood maps used for accuracy evaluation. In the end we would like to thank the USGS, which provided the SRTM-v2 DTM.

REFERENCES

References from Journals:

References from Other Literature:

References from websites: