Mapping Extent of Floods: What We Have Learned and How We Can Do Better

Yong Wang

Abstract: Much attention has been given to mapping the extent of a flood by using optical, radar, digital elevation model (DEM), and river gauge data. The mapped extent is often supported and verified by ground observations. The popularity of methods that use these data sets has arisen due to effectiveness, availability, and low cost. This paper summarizes the strengths and weaknesses of the individual and combined use of optical, radar, DEM, and river gauge data to map flood extent. The 1999 flood associated with Hurricane Floyd in eastern North Carolina is used as an example throughout the paper.

DOI: 10.1061/(ASCE)1527-6988(2002)3:2(68)

CE Database keywords: Flood plains; Mapping; Rain gauges.

Introduction

It is essential for effective response, recovery, and mitigation that the fullest extent of flooding be measured as a function of time. Mapping the extent of a flood by using optical, radar, digital elevation model (DEM), and river gauge data, supported and verified by ground observations, has received increased attention because of the availability of these data sets and the effectiveness of using the data for emergency response. In addition, the low cost of these data has helped to facilitate their use. Success stories can be found in a number of recent references [for example, Imhoff et al. (1987); Corbley (1993); Hess et al. (1995); Correia et al. (1998); Jones et al. (1998); Melack and Wang (1998); Dartmouth Flood Observatory (1999); Colby et al. (2000); Wang et al. (2002)]. During the 1999 flood that involved eastern North Carolina after Hurricane Floyd, Landsat 7 TM (thematic mapper), DEM, and river gauge data, coupled with field observations, were used to map the areas of flooding in Pitt County, North Carolina (Wang et al. 2002). Since then, substantial insights about flood extent mapping in floodplains have been gained through in-depth investigation of using these as well as other types of data. The objective of this paper is to discuss and summarize the individual and combined value of using optical, radar, DEM, and river gauge data to map flood areas, what has been learned from this experience, and how one might more effectively map the extent of future floods.

Mapping Extent of Flood

One goal in flood mapping is to identify areas that are flooded or not flooded. There are two steps in this process: (1) identifying water versus nonwater areas before and during the flood event, respectively; and (2) comparing the areas classified as water or nonwater before and during the flood event to determine which areas have been flooded. Once an assessment of water versus nonwater areas is made on both dates (before and during a flood), one can determine whether an area is flooded or not (Table 1). It should be noted that it is possible to have an area that is classified as water in the preflood data set and classified as nonwater in during-flood or postflood data set. Possible explanations for this apparent discrepancy include change of land use and land cover type between the dates that the data sets were acquired or created. Other factors such as cloud shadows in the preflood optical remotely sensed data (Wang et al. 2002) or simply an analyst error can contribute to this apparent discrepancy.

Mapping Flood Extent using Optical Data

Optical remote sensors measure solar reflectance from objects on the ground. The sensors depend on solar radiation for their measurement. Optical data collected on a clear and sunny day work the best. Optical data include satellite data (such as Landsat TM data) and aerial photographs. Because dry surfaces and wet/water surfaces have distinctly different reflectance characteristics, the TM data or aerial photographs can easily identify the surfaces. After Hurricane Floyd of 1999, there was massive flooding in eastern North Carolina. The flooding was clearly evident when comparing a pair of TM images acquired on July 28, 1999, before the flood, and on September 30, 1999, during the flood (Fig. 1). The floodwater overflowed the Chowan, Roanoke, Tar/Pamlico, and Neuse rivers, which created much wider river channels.

Table 1. Change Detection Criteria in Identifying Flooded and Nonflooded Areas

<table>
<thead>
<tr>
<th>Before flood</th>
<th>During flood</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry</td>
<td>Wet</td>
<td>Flooded areas</td>
</tr>
<tr>
<td>Wet</td>
<td>Dry Errors</td>
<td>Errors possibly caused by data sets or other factors</td>
</tr>
<tr>
<td>Wet</td>
<td>Wet Water</td>
<td>Water bodies (ponds, rivers, and so on)</td>
</tr>
</tbody>
</table>
There are several methods for locating/classifying water or nonwater areas using TM data [for example, Jensen (1996)]. An addition of two TM bands (TM4 + TM7) from July 28 and September 30, 1999, respectively, is one method of providing additional spectral information on the event (Wang et al. 2002). After locating water or nonwater areas on both images, determination of areas that were flooded could be easily made by using a change-detection method (Table 1). During the 1999 Pitt County, North Carolina, flood, of a total county area of 1,697 km$^2$, 13.3 km$^2$ were normal water bodies (such as river channels and ponds), 288.8 km$^2$ were flooded, and 1,453.7 km$^2$ were not flooded on September 30 (Wang et al. 2002). In summary, using TM data in flood extent mapping is

1. Reliable and accurate: The flood extent map was partially verified using aerial photographs taken during the flood and from field observations after the floodwater receded;
2. Simply applied: Georeferencing a pair of TM images, identifying water or nonwater areas on the pair of TM images, and carrying out a change-detection analysis were the major steps involved; and
3. Efficient and cost-effective: A typical Landsat 7 TM image covers an area of about 180$\times$180 km, or 32,400 km$^2$; a flood extent map for such a large area can be easily developed.

Currently each Landsat 7 TM image costs about $600 when ordered from the U.S. Geological Survey (USGS) EROS data center. Georeferencing images and analyzing changes in the images did not require specific computer hardware and software.

However, there are limitations in the use of TM data. First, due to dense or continuous canopy coverage in forested areas (such as bottomland forest/hardwood swamps in the floodplain and some dense southern yellow pine stands near river banks) and due to the lack of canopy penetration of the TM data, flooded areas under the canopy cannot be detected. This obviously leads to the underestimation of flooded areas. In our study, this underestimation was verified and quantified through ground truthing and visual interpretation of low-altitude oblique aerial photos taken during the 1999 flood. The undetected flooded areas most frequently showed up as “patches or holes” within the primary floodplain near riverbanks on the derived flood extent map (Wang et al. 2002). It is important to point out this underestimation, because floods in the floodplains of North Carolina, the entire East Coast, and the coast of the Gulf of Mexico often occur from midsummer to fall, and trees in the floodplain are almost fully leafed during this period of time.

Second, the 15$\times$15 m spatial resolution of TM’s panchromatic band 8 and 30$\times$30 m spatial resolution of TM’s bands 1 to 5 and 7 may be too coarse for identifying small areas or patches of flooded terrain in vegetated areas, agricultural fields, or commercial and residential areas.

Third, since a satellite’s orbit patterns in space and ground tracks are predetermined, this limits the temporal observation (to a target or study area) from the satellite. For example, it takes the Landsat 7 satellite 16 days to revisit the same area. Thus, the satellite data may not contain the information needed to map the (full) extent of a flood. In the worst case, if a flood occurs just after the satellite’s overpass and the water quickly recedes before the satellite’s next overpass, from the satellite’s point of view, nothing (no flooding) has occurred. Also, during the time a satellite passes over an area, the area being imaged should be cloud-free or at least not dominantly under cloud cover. If it is a cloudy day, optical remotely sensed data might not be usable. In studying the 1999 flood of eastern North Carolina, TM data acquired on the July 28 were used as the preflood data set because of (almost) cloud-free conditions, and TM data collected on September 30 were used as the during-flood data set. However, the Tar River in Pitt County crested on September 21 (Colby et al. 2000; Wang et al. 2002). Thus, the image taken on September 30 may not portray the full extent of the flooding.

### Mapping Flood Extent using Radar Data

The basic principles of using radar data to map flood extent are similar to those used with the optical data discussed earlier: identifying water or nonwater areas before and during the flood event, respectively, and comparing the areas classified as water or nonwater before and during flood event to determine which areas represent flooding. Once the locations of water or nonwater areas in the data sets on both dates (before and during a flood) are made, the decision rules (Table 1) can be used again to determine whether an area is flooded or not.

The advantage of using synthetic aperture radar (SAR) data over optical data is the ability of radar microwaves to penetrate cloud cover and forest canopies (to some extent). (SAR is a radar system in which a fine azimuth resolution is achieved by storing and processing data on the Doppler shift of multiple return pulses and a fine cross-track resolution is obtained by frequency modu-

---

**Fig. 1.** TM images of northeastern part of North Carolina showing (a) preflight condition on 7/28/1999, and (b) during-flood condition on 9/30/1999. Portions of floodplain in eastern North Carolina were flooded, as seen in comparison of (b) to (a). Four major river systems were identified. Three river gauge stations on the Tar/Pamlico River are shown in (a): T= Tarboro, G= Greenville, and W= Washington.
nominal area of 100 km starts from $2,750 per image (http://www.rsi.ca/storefront/radsat/rsat_img_usd.htm). For ERS SAR data, it costs from $1,940 to $2,810 to order a raw dataset or map-oriented digital image covering an area of $\sim 100 \times \sim 100$ km (http://www.auslig.gov.au/flows/prod_ser/ersprice.htm).

**Flood Mapping using Topographic and River Gauge Data**

USGS has created DEM data or digital representations of ground surface topography for the U.S. and its territories, as part of the National Mapping Program. DEMs are available at 7.5-min, 15-min, 2-arc-s, and 1-degree units. The 7.5-min DEM data correspond to the USGS 1:24,000 scale topographic quadrangle map series. Each 7.5-min DEM is based on $30 \times 30$ m data spacing in x and y directions. The vertical accuracy of DEMs depends on product levels. Level 1 data are created for high-relief areas, and Level 2 data are created for low-relief areas (such as eastern North Carolina). The RMS error of Level 2 data for eastern North Carolina is half of the contour interval or $\pm 1$ m (that is, most errors occur within this value). The elevation interval is 0.5 m.

The USGS and the U.S. Environmental Protection Agency operate a network of gauge stations on rivers in the U.S. There are 647 river gauge stations in North Carolina, at which a river’s water surface height, discharge, and so on are measured daily and are available to the public at the USGS Web site (http://sg1.dnrcrlg.er.usgs.gov). In studying the 1999 flood extent in the Greenville area, river water surface heights at three gauge stations on the Tar/Pamlico River coupled with DEM data in the area were used. The three stations [from upstream to downstream, Fig. 1(a)] were (1) Tarboro, in Edgecombe County, about 16 km upstream along the Tar River off the northwest corner of Pitt County; (2) Greenville, located near the center of Pitt County; and (3) Washington, in Beaufort County, about 5 km off Pitt County in the east. Along the river channel, the distance between the Tarboro and Greenville stations is about 44 km, and the distance between the Greenville and Washington stations is about 34 km. Due to heavy rain caused by Hurricanes Dennis and Floyd between September 6 and 15, 1999, the Tar/Pamlico River steadily rose. The preflood data chosen for this analysis was September 5, 1999, when the river was at its regular flow and water surface height. The flood date was September 21, when the Tar/Pamlico River crested between Tarboro and Washington. The water surface heights (above mean sea level) at the three gauge stations on the two dates are tabulated in Table 2.

<table>
<thead>
<tr>
<th>River gauge station</th>
<th>Before flood (m)</th>
<th>During flood (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tarboro</td>
<td>4.7</td>
<td>15.4</td>
</tr>
<tr>
<td>Greenville</td>
<td>1.8</td>
<td>8.3</td>
</tr>
<tr>
<td>Washington</td>
<td>1.1</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Since both SAR systems are of short wavelength, their penetration into tree canopies is limited, and in some cases they may not penetrate forested areas with dense canopies at all. If the sensor fails to penetrate the canopy, detection of flood/nonflooded conditions beneath tree canopies becomes questionable. Also, because these SAR sensors are on board satellites, they will have similar temporal and spatial resolution limitations, as does Landsat TM. Last, it may be expensive to purchase SAR data. The price of a Radarsat image of standard beam mode covering a nominal area of $100 \times 100$ km starts from $2,750$ per image.
ments, and each segment was about 3 km in river channel length. The surface water heights (Table 2) at the pre-flood time period between the Tarboro, Greenville, and Washington stations were linearly interpolated for each segment. The same procedure was used to determine the during-flood surface water heights. Since the increment in the DEM was 0.3 m, the interpolated water surface height value was rounded to the nearest multiples of 0.3. Once the water surface heights for the pre-flood and during-flood time periods were derived for each segment between Tarboro and Washington, the decision rules shown in Fig. 3 were used to identify nonflooded versus flooded areas:

1. Areas with elevations above the interpolated water surface height of September 21 were considered to be not flooded;
2. Areas of elevation between the interpolated water surface heights of September 5 and 21 were considered to be flooded; and
3. If the surface elevation of an area was less than or equal to the interpolated water surface height on September 5 (preflood date), then that area was considered to be a regular water body.

Since the DEM data are formatted by rows and columns and with north pointed up (Fig. 4), an easy way to map the inundated area using the DEM and river gauge data is (1) to segment the DEM data into small areas based on stream length; (2) to recode each segment based on the interpolated water surface heights at the preflood and during-flood periods to create a thematic layer of nonflooded, flooded, and regular water body areas (Fig. 3); and (3) to mosaic the thematic layers for all segments to create the final inundation map for the study area. Fig. 4 shows the flood extent on September 21, 1999, for the Greenville study area (about 22.0 km west-east and 20.0 km north-south, or an area of 440.0 km²). The black area was the regular river/stream channel and water body (17.0 km²); the shaded area the flooded area during the flood (66.4 km²); and the white area the nonflooded area (356.8 km²).

There are several advantages of using the DEM and river gauge data to map the flood extent:
1. These data are reliable and accurate. The derived flood extent map along the Tar River banks was verified by aerial photographs taken during the flood and by field observations after the floodwater receded in the study area.

2. The methodology is simple, efficient, and economical. Interpolation of river water surface heights between gauge stations and recording the DEM data were the major operations. DEM data are available for the entire U.S. and its territories. There are also many gauge stations along major rivers in the U.S. DEM and river gauge data are generally free to the public.

3. The data are easily updated. Since the DEM data do not vary in time, and the derived flood extent map is based on the combination of DEM and river gauge data, a model that uses DEM data to forecast inundation at different river water surface heights can be designed. Questions like “what will be the flood extent if the river water surface height is up another 1 m, 2 m, and so on?” can be posed and quickly answered. This information will be of great value in rapid response, recovery, and mitigation during and after a major flood.

However, this method also has limitations. Using river gauge data before and during a flood to forecast flood inundation only works where river surface height can be measured. To map flooding caused by a river’s tributaries or flooding in a tributary area, river gauge data on the tributaries are needed. Otherwise, significant error may occur. This was why the measured water-surface heights on the Tar/Pamlico River by three gauge stations were unable to capture the flooding that occurred inside the tributaries and the northern and southern areas away from the main river channel (Fig. 4). Scattered flood areas were observed from ground observations and analysis of TM data (Wang et al. 2002). Also, the river gauge/DEM method failed to identify water bodies and/or flooded areas at higher elevations. Caution should also be exercised when working in large areas with large topographic variations that include localized depressions. Water surface heights from other river gauge stations should also be incorporated, and efficient interpolation and/or extrapolation methods should be developed to provide a continuous representation of flood elevations upstream and downstream. In areas where few river gauge stations exist, estimates of water surface height might have to be made.

The Level 2 DEM data have an estimated accuracy of one-half the interval of the contour lines on the USGS topographic sheet. The contour interval of the topographic sheets in Pitt County is 2 m. The potential ±1 m variations in accuracy of the DEM data will affect the estimated flooded areas due to the low topographic relief in the study area. Wang et al. (2002) reported that the
flooded area varied from 95.3 km² (at 1 m less than the DEM value), 98.6 km² (at the DEM value), and 110.4 km² (at 1 m more than the DEM value) in their study of flooding in an area of 580 km² within Pitt County.

**Concluding Remarks**

**Significance**

Many advantages exist in mapping the extent of a flood using multiple data sets, such as optical, radar, digital elevation model (DEM), and river gauge data, due to their effectiveness, availability, and cost. The cost of a Landsat 7 TM image is low, and the inundation analysis is simple. Many gauge stations exist on U.S. rivers, and the river gauge data are available to the public. The USGS DEM data can be freely downloaded. There are limitations, however, when these data are used individually. For example, Wang et al. (2002) reported that due to the inability of TM data to penetrate vegetation canopies in bottomland forest and hardwood swamps, TM data alone cannot identify those flooded areas under the canopies and lead to an underestimate of flooding in those areas. However, using DEM and river water surface height data can indeed identify those flooded areas. Furthermore, an analysis of DEM and river surface height data fails to determine flooding at higher elevations or in areas away from primary river channels. The analysis shows that using TM data can delineate these flooded areas. Combining the flooded areas identified from both analyses using a simple “OR” logic can overcome these limitations. Therefore, the integration of optical, radar, DEM, and river gauge data into the flood mapping analysis should provide for more effective delineation of future flood extent.

To take full advantage of what the data can offer, the data should be acquired as close as possible in time, especially the acquisition of data during a flood period. However, for the pre-flood analysis of water versus nonwater, the timing is not as critical. In addition to identifying flooded versus nonflooded areas, the land use and land cover type data should be included to study how a flood affects each land use and land cover type. Damage caused by a flood in commercial, industrial, residential, or open-space areas will be vastly different. Also, property value information from county or city tax offices should be added. By incorporating these data, damage (in terms of dollars) to houses and real estate properties can be estimated.

Due to the variation of annual precipitation, a 100-year or 500-year flood map of a floodplain varies slightly. Updated and accurate 100-year and 500-year flood extent maps provide information essential to responding to a flood and identifying flood mitigation activities. A method that integrates optical, radar, DEM, and river gauge data should make the updating of this information easy and reasonably reliable and should create more accurate flood extent maps for the 100-year or 500-year flood event.

**Additional and Better Digital Elevation Model and Optical Data Sets**

The National Aeronautics and Space Administration and the National Oceanic and Atmospheric Administration (NASA/NOAA) have a LIDAR (light detection and ranging) sensor (http://www.csc.noaa.gov/crs/ctm/amt2.html). This involves sending out a laser beam and measuring the time that the laser echo is backscattered by an object. By doing so, the object height or surface height is measured. Currently, the primary mission of the NASA/NOAA LIDAR sensor is to map the topography of Greenland glacier. The sensor is also being used to map the topography of beaches and dunes of the U.S. coast on an annual basis and to map the topographic features of beaches after being hit by a hurricane. The topography of beaches and sand dunes of the North Carolina coast was mapped after each hurricane strike since 1997. Using the LIDAR data one can create DEM data with high spatial resolution (up to 1.5×1.5 m) and fine height measurement (about 15 cm). The DEM data generated from the LIDAR data have helped greatly in mapping the topography of beaches, dunes, and coastal lines and in studying the topographic changes in the immediate coastal areas.

The state of North Carolina is now funding a project to use LIDAR data to create a DEM dataset for the entire state as part of the statewide floodplain mapping initiative (Dorman 2000). DEM data of high spatial resolution (in meters) and high vertical accuracy (in decimeters) will be available soon.

The U.S. launched its Earth Observing System (EOS) AM satellite into space in 2000. On board the satellite is a suite of sensors. The addition of sensors (to the Landsat 7 satellite) in space increases the temporal resolution of remotely sensed optical data; also, several EOS sensors offer a fine spatial resolution data with a pixel size in 10×10 m to 20×20 m range.

The Advanced Land Observing Satellite (ALOS) program from the National Space Agency of Japan (NASA) (1999) will offer the opportunity to collect (simultaneously) optical, radar, and DEM data from a single satellite. The optical data, three visible and one near infrared bands, will have a resolution of 10×10 m. The L-band radar data will have resolutions ranging from 100×100 m to 10×10 m, depending on ground swath widths and polarizations. A pair of optical sensors, one forward-looking and one nadir-looking, will create the DEM data with a 2.5×2.5 m spatial resolution. The ALOS data should be available in late 2004.

All these data developments point toward increased use of remotely sensed data and will provide excellent flood maps in the future. Designing methods that integrate various types of data sets and taking advantage of the strengths of each data set and/or sensor remain the primary challenge. This paper provides an initial perspective on this task and addresses some issues in the effective use of remotely sensed data and other data sets to achieve useful and accurate floodplain maps.

**Acknowledgments**

The writer wishes to thank the three anonymous reviewers and the journal editor, whose comments have greatly improved the quality of the paper.

**References**


