IMPROVEMENT OF ITHACA EARLY WARNING SYSTEM FOR FLOODS

A. Albanese*[a], F. Giorgi[b], N.P. Premachandra[a], O. Terzo[c]

[a]DITAG, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy – (adriana.albanese, prasanna.nishshanka)@polito.it
[b]ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action), Via P.C. Boggio 61, Torino, Italy - fabio.giorgi@ithaca.polito.it
[c]ISMB (Istituto Superiore Mario Boella),Via P.C. Boggio 61, Torino, Italy - terzo@ismb.it

KEY WORDS: Precipitation, Satellite, Databases, Floods, Mapping, Real-time

ABSTRACT:

This paper describes an Early Warning system for floods based on the use of near real-time satellite rainfall data and describes some expected improvements that are being developed through the use of rain gauge data regarding Bangladesh.
This completely automated Early Warning System, developed by the Association ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action), works on river basins scale with global coverage by using satellite rainfall data from Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA). The 3-hourly 3B42 data from 1998 to 2007 are used to detect all historical flood events in the past ten years using a hydrological method based on Depth Duration Frequency curves. The 3-hourly real time 3B42RT data with some statistical adjustments are used to detect critical rainfall events and to make alerts in near real-time. The capability of monitoring all basins of the world at the same time is achieved by a grid computing system, consisting of a combination of computer resources to process large amounts of data. In order to improve the results obtained from this system, the use of gauge rainfall data are very useful. Bangladesh rainfall data are being used in order to perform some tests on the correctness of the flood alerts in this first test area, where floods are very frequent and occur almost every year.
At present, this early warning system is able to detect worldwide heavy rainfalls successfully and the alerts are automatically mapped through the web application system.

1. INTRODUCTION

Flood warning systems are an important part of a holistic approach to the reduction of both tangible and intangible damage due to flooding, including loss of life, damage to property and goods, and reduction of negative health and social impacts (Penning-Rowsell et al., 2000; Carsell et al., 2004).
Since flooding is one of the most disastrous hazards in many regions of the world (Schanze, 2009) and it is increasing in the last years, the flood risk management has become very relevant. Flood forecasting and warning is most often provided through government agencies; however, real-time flood forecasting systems are often limited to developed countries and difficult to find in developing countries where more than three quarters of the world's population are living. In this context, the organisation ITHACA has developed an early warning system for floods able to provide a global coverage and particularly devoted to monitor floods in less developed countries.
ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action, www.ithacaweb.org) is a non-profit organization founded by Politecnico di Torino and SiTI (the Higher Institute on Territorial Systems for Innovation, www.siti.polito.it); it conducts research in the field of Geomatics and particularly it devotes its activities to monitor, analyze and forecast natural disasters, in collaboration with the WFP (World Food Programme,www.wfp.org), the largest UN (United Nations) operational Agency.
This Early Warning system gives an alert for floods around the world monitoring the heavy rainfalls in near real-time; it was developed in order to increase efficacy in approaching emergency preparedness related to flood events by WFP or other humanitarian assistance organizations. Through this system, they can understand the potentially floodable areas where their assistance is needed.
The methodology of this system is here summarized only briefly, because it has already been described in detail in previous papers (Albanese et al., 2008a; Albanese et al., 2008b; Albanese et al., 2009).
The aim of the second part of this project, in course of development, is to verify the correctness of the flood alerts and to introduce some improvements in order to have always better results. To this end, the use of gauge rainfall data is very useful and, thanks to the collaboration of the WFP, gauge rain data of Bangladesh have been obtained. The second part of the paper describes the use of those rainfall data, in progress of elaboration, in order to perform some tests in this first test area.
At present, the early warning system is able to detect heavy rainfall worldwide and the alerts are automatically mapped through the web application system.

2. METHODOLOGY

This Early Warning System is based on precipitation analyses and it uses rainfall data from satellite at worldwide extent.
The first step in this project is the individuation of the different extreme meteorological events occurred in the past years; a series of ten years of data have been elaborated using a hydrological analysis, in order to individuate the different extreme meteorological events occurred and the related cumulated rainfall values during these events.
The same analysis is performed using near real-time data to monitor current rainfall conditions. This can be considered the proper "early warning system for heavy rainfalls".

2.1 Data

The rainfall dataset used in this project belong to Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA).

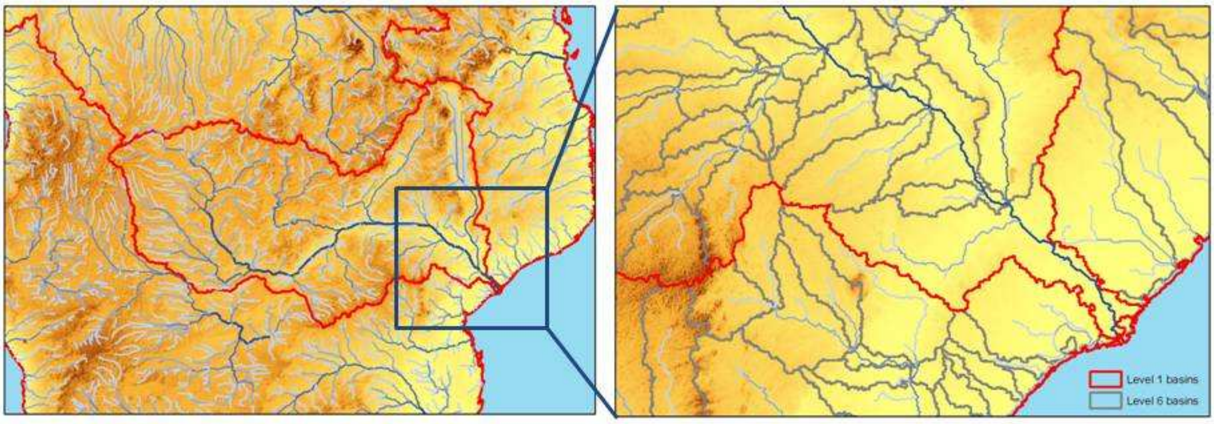

Figure 1. Stream line and drainage basin layers from HYDRO1k; the watershed layer is constituted by a territorial subdivision from level-1 to level-6

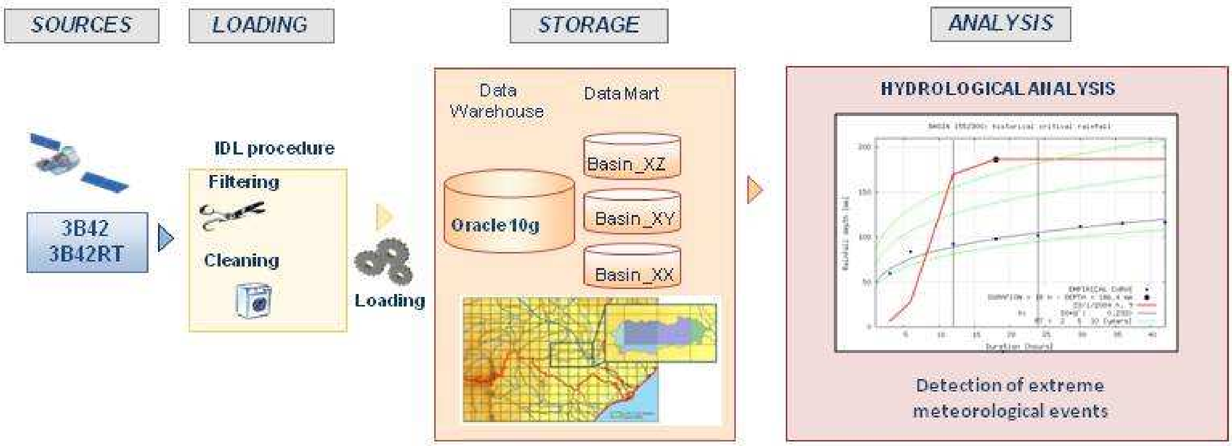

Figure 2. Early Warning System architecture

The TMPA provides a calibration-based sequential scheme for combining precipitations estimates from multiple satellites, as well as gauge analyses where feasible (Huffman et al., 1997; Huffman et al., 2007); in particular, in this project two products have been used: the 3B42 and the 3B42RT. These gridded estimates are on a 3-hour temporal resolution and 0.25 degree by 0.25 degree spatial resolution in a global belt, extending from 50 degrees South to 50 degrees North latitude.
The analysis of rainfall data are led on the basis of river catchment due to the strict correlation between the basin characteristics and flood effects. The GIS watershed layer of HYDRO1k, developed at the U.S. Geological Survey's Centre for Earth Resources Observation and Sciences (EROS), has been used. It is constituted by a territorial subdivision at different levels, increasing details from level-1 to level-6 (fig. 1); watersheds at maximum detail (level-6 basins) are used for the system.
Oracle 10g© has been used to load and elaborate the whole dataset of rainfall that was structured in a data warehouse architecture (fig. 2). This allows the management of this huge dataset by performing complex data analyses and queries.
A pre-elaboration of rainfall data is required: the rainfall values, originally associated to each TRMM grid cell (0.25 x 0.25 degree latitude and longitude) are calculated for each level-6 basin and it is the value used by the hydrological model to calculate critical rainfalls. A second pre-elaboration transforms the data into a suitable format for the database, from the primitive raster format of the gridded data to the ASCII format. Both pre-elaboration procedures are performed through a routine elaborated by ITHACA in a suitable environment for the raster elaboration, the IDL (Interactive Data Language).
Once the data are loaded, they become ready for the different analyses. In particular, the first analysis allows the detection of all historical flood events occurred in the past (see par. 2.2); subsequently, the whole system functions in near real time in order to find current critical rainfalls (see par. 2.3).

2.2 Analysis of historical data

The individuation of extreme meteorological events occurred in the past has been executed elaborating the whole dataset of 3B42 product, belonging to the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA), that covers the period from 1998 to the delayed present.
The 3B42 product is a post real–time research-quality product and it is available about 10-15 days after the end of each month, the time necessary to perform some calibration processes.
The elaboration of the whole dataset of 3B42 data is performed using the hydrological method of Depth Duration Frequency (DDF) curves (fig. 3), that allows the individuation of critical rainfalls occurred in the last ten years (from 1998 to 2007).
The DDF curves, calculated using Gumbel's extreme value distribution, represent the relation between the duration of rainfall d (measured in hours) and its associated depth of rainfall h (millimeters) and lead to the calculation of the recurrence interval of the extreme rainfall values, measured in years and called *Return Time (RT)*. These curves have been calculated for each level-6 basin at worldwide extent using a

complete automatic procedure developed in Octave (a Matlab porting open source software).

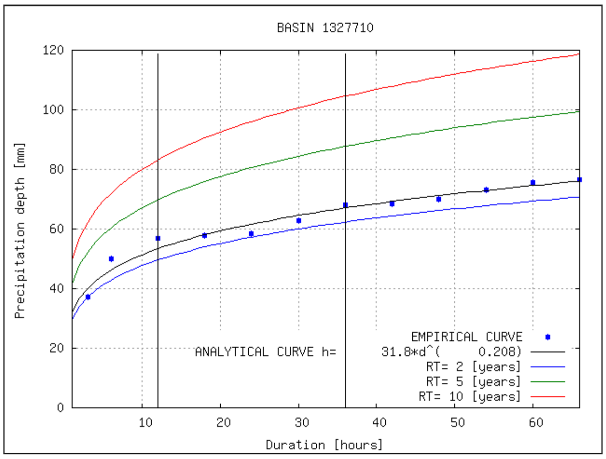

Figure 3. Depth Duration Frequency (DDF) curves

The entity of each rainfall event during the last ten years was determined comparing cumulated rainfall to the historical ones (represented by blue dots and interpolated by the black curve in fig. 3). This comparison was done using the frequency factor *(K_T)*, function of the mean of the annual maxima (Chow, 1954; Chow et al, 1988; ASCE, 1996) (fig. 4):

$$K_T = \frac{h(d)}{m(d)} \tag{1}$$

where K_T = frequency factor
$h(d)$ = depth of rainfall for the duration d
$m(d)$ = mean of the annual maxima for the duration d

The cumulated rainfall was considered as a critical rainfall event (represented by a black dot, in fig. 4) when the value of K_T is higher then a threshold of 0.9.

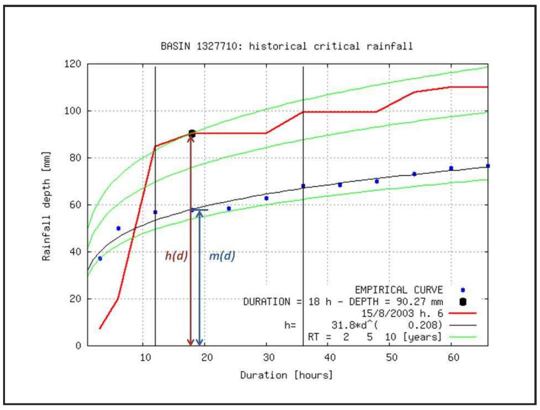

Figure 4. DDF curves with different return time are represented in green, cumulated rainfalls are represented in red and critical rainfall, detected through the calculation of the KT factor, are shown by the black dot

The estimation of K_T was performed for the durations of rainfalls considered critical for each river basin, able to gives peak discharge on the river (Alfieri et al., 2008). These rainfall durations are defined through the calculation of a specific parameter called *lag time* (t_l), using empirical formulas where morphometric parameters of the basin are required (e.g. longest flow path of the river, area and slope of the basin).

Automatic procedures are developed in order to perform the calculation of drainage basin parameters for worldwide river basins. After the definition of critical durations, the whole dataset of rainfall (more than 29.000 data for a single basin) was processed and the critical rainfalls for each level-6 basin are individuated.

2.3 Real time system

ITHACA has developed a Real Time System which analyze real time rainfall data to detect critical rainfalls and to present affected areas using maps. The real time rainfall data (the 3B42RT product) are compared with the DDF curves determined using historical data (see par. 2.2).

The 3B42RT data are posted onto the web about 6 hours after observation time, although processing issues may delay or anticipate this schedule; these estimates are to be considered as highly experimental, because the calibrator used for the research product (as instance 3B42), the TCI (TRMM combined instruments), is not available in real time and because for the RT system the monthly gauge adjustment is not possible.

Consequently, their use requires an adjustment in order to minimize the difference between the real-time data (3B42RT product) and the corrected product (3B42, validated and available after two months). For this purpose a statistical approach was implemented in order to compare the two series of data for the superimposition period, available from January 2003 to December 2007 (Albanese et al., 2009). This analysis creates a correction factor for each level-6 basin, in a completely automatic procedure. Then, it is possible to correct the real time satellite rainfall data product directly as it is downloaded.

The real system starts downloading 3B42RT from the FTP NASA server (ftp://trmmopen.gsfc.nasa.gov/pub/merged/mergeIRMicro) and finishes saving critical events detected in the database (fig. 5). It checks the server for the data availability every 14,400 seconds and downloads the data available. Once rainfall values are downloaded and transformed into .txt format, the files with rainfall data related to a single level-6 are generated. This is a very important step, because within few seconds and without passing through the Oracle database, files needed by the hydrological model are created. At this point these adjusted data can be used for the detection of critical events, by using the hydrological model above described (see par. 2.2).

The capability of monitoring all basins of the world at the same time is achieved by a grid computing system, consisting of the combination of computer resources to process large amounts of data (Albanese et al., 2009).

2.4 Use of rain gauge data to validate the methodology to detect heavy rainfalls

The precipitation data analyses, that allow to determine DDF curves used as comparison for rainfall in real time, shall be performed on a historical series of precipitation data adequately long.

In studying the climate of a particular territory, all the weather conditions over long periods of time (in the order of magnitude of decades) must be considered. In fact, climate is an abstract concept obtained by processing statistically a set of meteorological variables (e.g., daily or hourly data) in order to get appropriate climate indices (mean values and extreme values, cumulative, frequency, etc.).

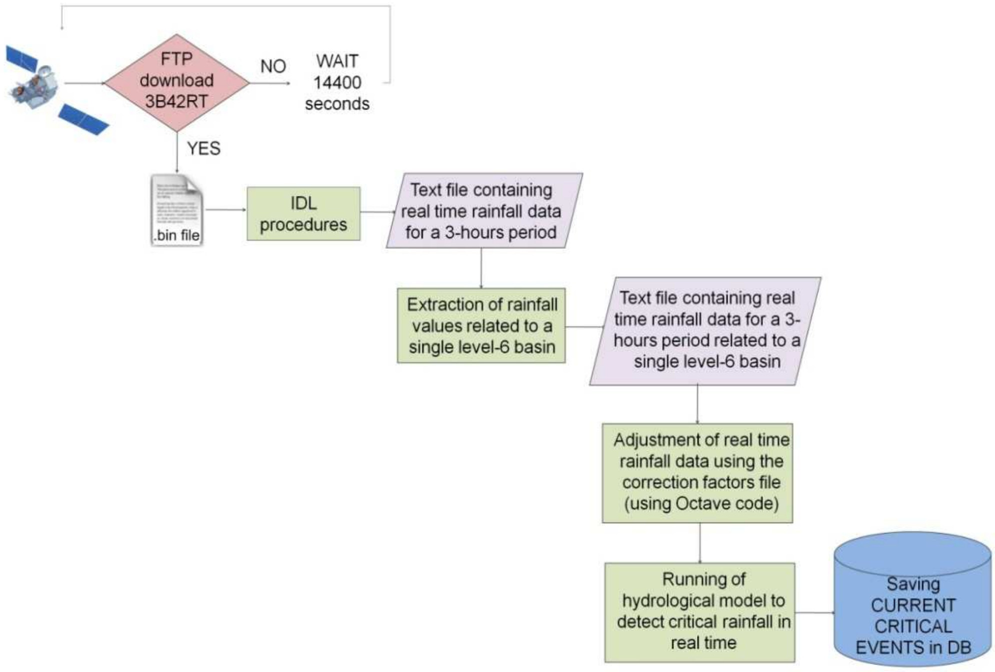

Figure 5. Work flow of the Real Time System: from the downloading of 3B42RT rainfall data in real time to the detection of critical rainfall events

According to the World Meteorological Organization (WMO) standards, the current climate of a territory used as reference to assess the degree of anomaly of current data is obtained by statistically processing the data of the last 30 years (Mariani, 2006).
Long series of satellite precipitation data are unfortunately not easily available for analysis; in the system developed by ITHACA historical series of ten years have been used. In this context, the use of 30 years of rain gauge data is particularly useful for improving the determination of DDF curves and to determine more accurately the threshold able to give alerts in this system.
Furthermore, the data set of ground network is used to estimate the difference between the satellite and the gauge precipitation data in order to make an appropriate adjustment in real-time satellite data (see par. 2.5).
The rain gauge data used are from Bangladesh and come from two different rain gauge networks (fig. 6): 54 stations are from FFWC (Flood Forecasting and Warning Centre, Bangladesh - http://www.ffwc.gov.bd/) and 35 are from BMD (Bangladesh Meteorological Department - http://www.bdonline.com/bmd/).
The rainfall series are from 1978 to 2008 for FFWC stations and from 1981 to 2008 for BMD, so for both networks the rainfall series cover about 30 years.
A procedure written in Octave has been developed to analyze all Bangladesh data series in an automatic way, in order to create the DDF curves using 30 years of data and in order to individuate critical rainfalls for each station in that period. The calculation of DDF curves, using a long series of data, allows to perform statistical analysis in order to define the recurrence interval of rainfall for a T-years, where T is higher than the length of the series. For instance in this case, starting from 30 years of data, the depth of precipitation until a return time of 100 years (that means, the rainfall with a 1 percent probability in any year, or chance of 1 in 100) has been defined, while starting from a series of 10 years of satellite data, this statistical extrapolation could be incorrect. So the graphs calculated using gauge data are more detailed both spatially and temporarily than the DDF curves using satellite data. The comparison is shown in figure 7: the DDF curves calculated using the satellite data (on the left of fig. 7) are related to the whole hydrographical basin (shown in green) and define the depth of rainfall until 10 years of Return Time; in the same area there are two rainfall stations on the ground, related to Bandarban and Lama town, and these two graphs (on the right of fig. 7 and fig. 8) define the depth of rainfall until 100 years of Return Time.
Through this analysis, the rainfall thresholds that define the alerts on the Bangladeshi territory will be defined more accurately.

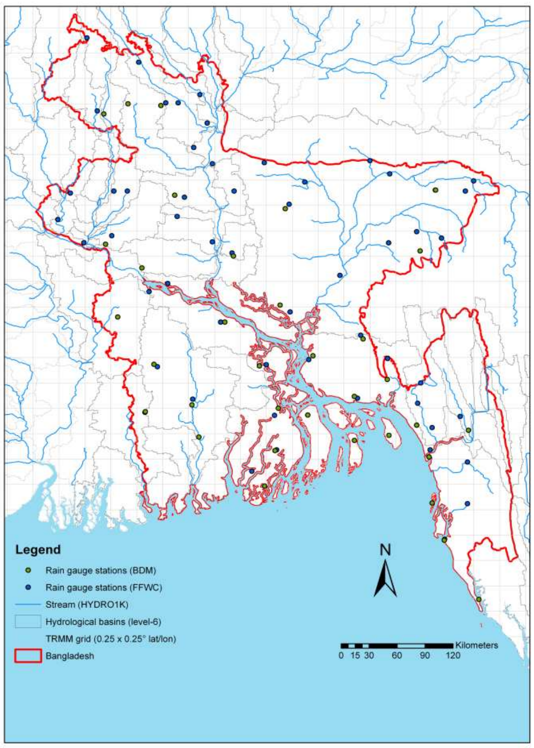

Figure 6. Rain gauge networks of Bangladesh

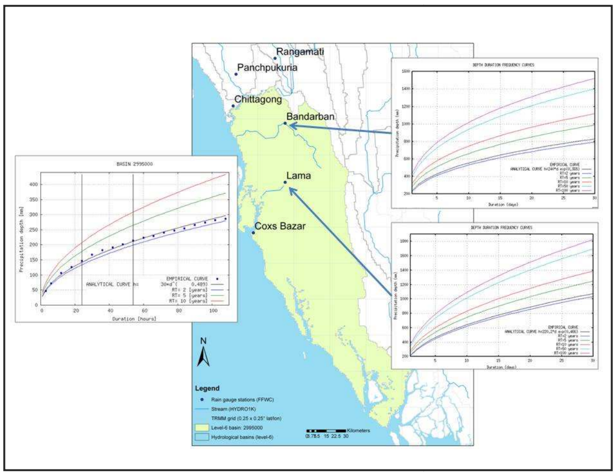

Figure 7. Comparison between DDF curves calculated using satellite rainfall data (graph on the left) and those calculated starting from rain gauge data (on the right)

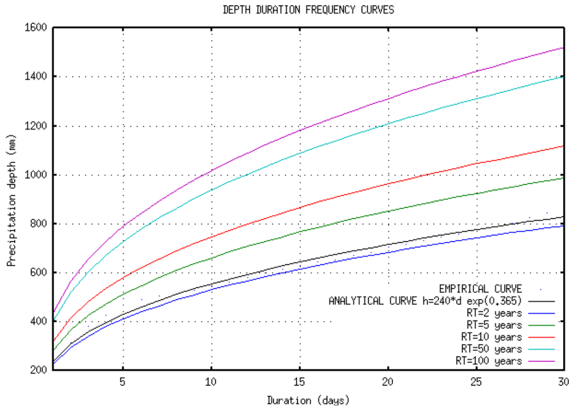

Figure 8. DDF curves calculated using rain gauge data: the statistical analysis of a long series of data allows to determine the rainfalls with a recurrence interval higher than the length of the series (so, in this case greater than 30 years).

2.5 Use of rain gauge data to validate ITHACA product

As this early warning system is designed for rain based flooding situations, in order to analyze the accuracy of TRMM data products and to analyze the effect of the ITHACA correction algorithm it is important to select a site where rain based flooding situations are very common. Bangladesh is very prone to rain based flooding situations. Therefore, to analyze the accuracy of TRMM data products and to analyze the effect of ITHACA correction algorithm, the TRMM data and ITHACA product from 2003 to 2007 were statistically compared with daily rain gauge data of Bangladesh for the same period.

As the normality tests showed that the data are not from normal distributions, the Wilcoxon signed rank test, which is one of Non-Parametric statistical tests, was used to compare daily paired data. The daily pairs were prepared for the period from 2003 to 2007 as Gauge data−3B42RT data, Gauge data−3B42 data and Gauge data−ITHACA product. The Wilcoxon signed rank test showed that out of 79 gauge stations at 63, 45 and 45 stations consequently the 3B42RT, 3B42 and ITHACA products are significantly different from gauge data at 95% confidence interval. In other words, at 20%, 43% and 43% of places the 3B42RT, 3B42 and ITHACA products respectively are not significantly different from gauge data. These results show that ITHACA product can achieve in near real-time a close accuracy to the most accurate product of TRMM (3B42) which has a latency of about 2 months.

3. RESULTS AND CONCLUSIONS

The alerts produced by this system are automatically mapped using the informative layers extracted from WFP SDI (Spatial Data Infrastructure) (Ajmar et al., 2008).

The overall situation can be visualized on the web application, in order to offer an easy access to the data during the emergency also to WFP's local offices. The web interface is developed with the most advanced javascript technologies in the geoweb field (Agosto et al., 2010).

The system shows the alerted basins with three different possible colours, according to the severity of rains (fig. 9); in particular, yellow colour means pre-alert (KT parameter is near to 1), orange indicates moderate alert (KT is between 1 and 2) and red is for heavy alert (KT higher than 2).

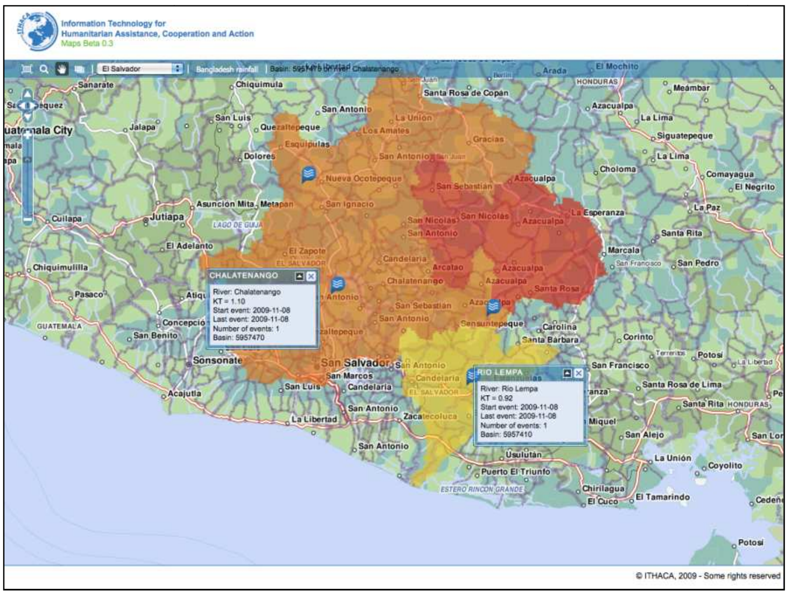

Figure 9. November 2009 El Salvador flood: example of flood event detected by ITHACA Early Warning System. The alerted basins are shown with 3 different possible colours, according to the severity of rains.

At the time being, ITHACA Early Warning System for floods is under testing for evaluation of the correctness of the alerts. The system is showing good results in monitoring heavy rainfalls and giving reliable alerts for floods. This first test was performed using information found on floods database available on the web (e.g. glidenumber, EMDAT, DFO) and gave good results.
The flood that hit El Salvador in November 2009 is an example of ITHACA's system successful recent results. This piece of news was reported by some of the world's most important media involved in natural disasters monitoring; for instance, Glide Number web site (www.glidenumber.net) showed this alert on November 9th (fig. 10). ITHACA Early Warning System indicated critical rainfalls on those territories starting from November 8th (fig. 9): the alert was therefore given one day in advance.

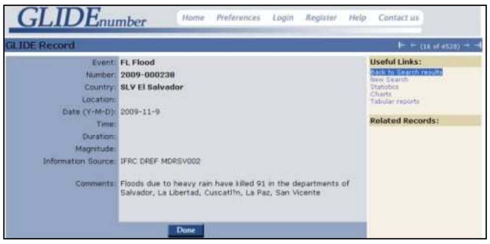

Figure 10. November 2009 El Salvador flood: news on Glide Number site

The results obtained from the system, already very good, have to be further validated and improved through the use of rain gauge data, coming from different parts of the world. The initial data used for this purpose were the rainfall data of Bangladesh; that country is suitable as a first test area, because floods are very frequent. As future developments, the use of rain gauge data from different parts of the world will be planned.
The rain gauge data of Bangladesh are been used for two purposes: firstly to validate the methodology for detecting extreme rainfalls, and secondly to analyze the accuracy of satellite data and to verify the effect of ITHACA's correction algorithm.
Regarding the first purpose, the results obtained using Bangladesh rainfall data shown an improvement in the determination of DDF curves both spatially and temporarily, in comparison to using satellite data; starting from these results, the thresholds of heavy rainfalls will be defined more precisely (in course of development).
As regards the second purpose, the results have shown that ITHACA product can achieve in near real-time a degree of accuracy close to the most accurate product of TRMM (3B42), which has a latency of about 2 months.

4. REFERENCES

References from Journals:
Agosto, E., Dalmasso, S., Pasquali, P., 2010. Ithaca world-wide Flood Alert System: the web framework, *Proc Gi4DM, Torino.*

Ajmar, A, Perez, F, Terzo, O., 2008. WFP Spatial Data Infrastructure (SDI) implementation in support of emergency management. *Proc XXI Congress ISPRS (Int Soc Photogrammetry Remote Sens)*, pp. 1097-1104.

Albanese, A, Disabato, F, Terzo, O, Vigna, R, Giardino, M, Perotti, L., 2008a. A preliminary approach to flood risk mapping and flood forecasting system for the LDCs. *Proceedings XXI Congress ISPRS (Int Soc Photogrammetry Remote Sens)*, pp. 1537–1542.

Albanese, A., Disabato, F., Terzo, O., Vigna, R., 2008b. Early Warning system for flood events and automatic classification of low resolution satellite imagery: Ithaca approach. *Proceedings 6th International Workshop on remote sensing for Disaster Management Applications*, Pavia, September 11-12, 2008, cd-rom.

Albanese, A., Giorgi, F., Premachandra, N.P., Terzo, O., Vigna, R., 2009. Application of an early warning system for heavy rainfall. *Proceedings 2th International Conference on Earth Observation for Global Changes (EOGC 2009).* Chengdu, China, 25-29 May 2009.

Alfieri, L., Laio, F., Claps, P., 2008. A simulation experiment for optimal design hyetograph selection. *Hydrol. Process.*, 22, pp. 813-820.

Carsell, K. M., Pingel, N. D., Ford, D. T., 2004. Quantifying the benefit of a flood warning system. *Natural Hazards Review*, 5, pp. 31–140.

Chow, V T, 1954. The log-probability law and its engineering applications. *Proc. ASCE*, 80, pp. 1-25.

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., Schneider, U., 1997. The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull Amer Meteor Soc*, 78, pp. 5–20.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., Wolff, D. B., 2007. The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. *J Hydrometeor*, 8(1), pp. 8-55.

Mariani, L., 2006. Alcuni metodi per l'analisi delle serie storiche in Agrometeorologia. *Rivista Italiana di Agrometeorologia*, 2, pp. 48 – 56.

Penning-Rowsell, E., Tunstall, S., Tapsell, S., Parker, D., 2000. The benefits of flood warnings: real but elusive, and politically significant. *Journal of the Chartered Institution of Water and Environmental Management*, 14, pp. 7–14.

Schanze, J, 2009. Editorial. *J Flood Risk Management*, 2, pp. 149–150.

References from Books:
ASCE, 1996. *Hydrology Handbook.* 2nd Edition, ASCE Manuals and Reports on Engineering Practice n. 28, New York.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. *Applied Hydrology.* McGraw-Hill, New York, USA.

References from websites:
HYDRO1K,
http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/hydro (accessed 20 Nov. 2009).

TRMM 3B42RT,
ftp://trmmopen.gsfc.nasa.gov/pub/merged/mergeIRMicro (accessed 20 Nov 2009).

5. ACKNOWLEDGEMENTS

The authors thank Dr. John McHarris (WFP, Dhaka) for providing the rain gauge data rainfall of Bangladesh.