I FACOLTA DI INGEGNERIA POLITECNICO DI TORINO

XXII Ciclo di Dottorato in Ambiente e Territorio

Final Dissertation April 2010

REMOTE SENSING DATA ANALYSIS FOR ENVIRONMENTAL AND HUMANITARIAN PURPOSES

The automation of information extraction from free satellite data

Tutor:
Prof. Piero Boccardo

Candidate:
Franca Disabato
Matr. 143511

AKNOWLEDGEMENTS

I would like to take the opportunity to thank many people for the support and contribution during this research project.

A major part of this study includes remote sensing work and data collection. The practical work of this study would never have been performed without facilities and logistics provided by the Land Processes Distributed Active Archive Center (LP DAAC), the image processing and archiving centre VITO and the Global Inventory Modeling and Mapping Studies (GIMMS) group.

I would like to thank my supervisor Piero Boccardo for giving me ideas and support during the planning and accomplishment of the thesis. My appreciation also goes to all my colleagues, for their company, help and discussions.

In particular, I owe my deepest gratitude to Francesca for helping me with my drafts, and above all for exchanging ideas and supporting me during my Ph.D experience.

I would like to express my gratitude to my parents and friends, for their moral support and affection all along. I would finally like to thank Siva, for encouraging me, for his useful advices and for his patience during the drafting of this thesis.

INDEX

INTRODUCTION

This work is aimed at investigating technical possibilities to provide information on environmental parameters that can be used for risk management.

The World food Program (WFP) is the United Nations Agency which is involved in risk management for fighting hunger in least-developed and low-income countries, where victims of natural and manmade disasters, refugees, displaced people and the hungry poor suffer from severe food shortages.

Risk management includes three different phases (pre-disaster, response and post disaster) to be managed through different activities and actions. Pre disaster activities are meant to develop and deliver risk assessment, establish prevention actions and prepare the operative structures for managing an eventual emergency or disaster. In response and post disaster phase actions planned in the pre-disaster phase are executed focusing on saving lives and secondly, on social economic recovery.

In order to optimally manage its operations in the response and post disaster phases, WFP needs to know, in order to estimate the impact an event will have on future food security as soon as possible, the areas affected by the natural disaster, the number of affected people, and the effects that the event can cause to vegetation. For this, providing easy-to-consult thematic maps about the affected areas and population, with adequate spatial resolution, time frequency and regular updating can result determining. Satellite remote sensed data have increasingly been used in the last decades in order to provide updated information about land surface with an acceptable time frequency. Furthermore, satellite images can be managed by automatic procedures in order to extract synthetic information about the ground condition in a very short time and can be easily shared in the web.

The work of thesis, focused on the analysis and processing of satellite data, was carried out in cooperation with the association ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action), a center of research which works in cooperation with the WFP in order to provide IT products and tools for the management of food emergencies caused by natural disasters. These products should be able to facilitate the forecasting of the effects of catastrophic events, the estimation of the extension and location of the areas hit by the event, of the affected population and thereby the planning of interventions on the area that could be affected by food insecurity. The requested features of the instruments are:

- Regular updating
- Spatial resolution suitable for a synoptic analysis
- Low cost
- Easy consultation

Ithaca is developing different activities to provide georeferenced thematic data to WFP users, such a spatial data infrastructure for storing, querying and manipulating large amounts of global geographic information, and for sharing it between a large and differentiated community; a system of early warning for floods, a drought monitoring tool, procedures for rapid mapping in the response phase in a case of natural disaster, web GIS tools to distribute and share georeferenced information, that can be consulted only by means of a web browser.

The work of thesis is aimed at providing applications for the automatic production of base georeferenced thematic data, by using free global satellite data, which have characteristics suitable for analysis at a regional scale. In particular the main themes of the applications are water bodies and vegetation phenology. The first application aims at providing procedures for the automatic extraction of water bodies and will lead to the creation and update of an historical archive, which can be analyzed in order to catch the seasonality of water bodies and delineate scenarios of historical flooded areas. The automatic extraction of phenological parameters from satellite data will allow to integrate the existing drought monitoring system with information on vegetation seasonality and to provide further information for the evaluation of food insecurity in the post disaster phase.

In the thesis are described the activities carried on for the development of procedures for the automatic processing of free satellite data in order to produce customized layers according to the exigencies in format and distribution of the final users.

The main activities, which focused on the development of an automated procedure for the extraction of flooded areas, include the research of an algorithm for the classification of water bodies from satellite data, an important theme in the field of management of the emergencies due to flood events. Two main technologies are generally used: active sensors (radar) and passive sensors (optical data). Advantages for active sensors include the ability to obtain measurements anytime, regardless of the time of day or season, while passive sensors can only be used in the daytime cloud free conditions. Even if with radar technologies is possible to get information on the ground in all weather conditions, it is not possible to use radar data to obtain a continuous archive of flooded areas, because of the lack of a predetermined frequency in the acquisition of the images. For this reason the choice of the dataset went in favor of MODIS (Moderate Resolution Imaging Spectroradiometer), optical data with a daily frequency, a spatial resolution of 250 meters and an historical archive of 10 years. The presence of cloud coverage prevents from the acquisition of the earth surface, and the shadows due to clouds can be wrongly classified as water bodies because of the spectral response very similar to the one of water. After an analysis of the state of the art of the algorithms of automated classification of water bodies in images derived from optical sensors, the author developed an algorithm that allows to classify the data of reflectivity and to temporally composite them in order to obtain flooded areas scenarios for each event. This procedure was tested in the Bangladesh areas, providing encouraging classification accuracies.

For the vegetation theme, the main activities performed, here described, include the review of the existing methodologies for phenological studies and the automation of the data flow between inputs and outputs with the use of different global free satellite datasets. In literature, many studies demonstrated the utility of the NDVI (Normalized Difference Vegetation Index) indices for the monitoring of vegetation dynamics, in the study of cultivations, and for the survey of the vegetation water stress. The author developed a procedure for creating layers of phenological parameters which integrates the TIMESAT software, produced by Lars Eklundh and Per Jönsson, for processing NDVI indices derived from different satellite sensors: MODIS (Moderate Resolution Imaging Spectroradiometer), AVHRR (Advanced Very High Resolution Radiometer) AND SPOT (Système Pour l'Observation de la Terre) VEGETATION. The automated procedure starts from data downloading, calls in a batch mode the software and provides customized layers of phenological parameters such as the starting of the season or length of the season and many others.

The thesis is divided in 7 sections. In the first two chapters an introduction to the applications of remote sensing for disaster management and a description of the datasets used for this work of thesis are provided. In chapter 3 are explained the context, the objectives and the methodology followed during the work. In chapter 4 and 5 a description of the activities performed for the automatic extraction of water bodies and of phenological parameters is provided. The chapter 6 is dedicated to the application of the produced dataset for the analysis of countries hit by natural disasters. Finally, in the conclusions, considerations on the performed applications are presented and future developments are discussed.

CHAPTER 1: INTRODUCTION TO REMOTE SENSING FOR NATURAL DISASTERS

1.1 INTRODUCTION

Nowadays, the world is facing disasters on an unprecedented scale: millions of people are affected by natural disasters globally each year and, only in the last decade, more than 80% of all disaster-related deaths were caused by natural hazards. When an event hits regions of the world where local authorities are not structured to deal with their complex effects, they normally ask the intervention of international organizations, such as UN agencies. Those organizations are in charge of the activation of emergency procedures in order to cover affected population's immediate needs and, with a long-term view, to prepare protracted relief and recovery operations.

The different events that may generate the need for a humanitarian intervention may be grouped in three major kinds:

- sudden disasters, which affect food access and/or cause population displacements;
- slow-onset disasters, as droughts and crop failures;
- complex emergencies, that can involve conflict, widespread social and economic disruption and large population displacements.

Risk management includes three different phases (pre-disaster, response and post-disaster) to be managed through different activities and actions (Figure 1). Pre-disaster activities are meant to develop and deliver risk assessment, establish prevention actions and prepare the operative structures for managing an eventual emergency event or disaster. In response and post-disaster phase, actions planned in the pre-disaster phase are executed, focusing on saving lives and, secondly, on social and economical recovery.

FIGURE 1.1: RISK MANAGEMENT PHASES FOR NATURAL DISASTERS

The main agency working in this field is the UN World Food Programme (WFP), which increasingly is being called on to respond to hunger in the context of a wide range of emergencies, from natural disasters like the tsunami tragedy in Asia to manmade crises such as the ongoing civil conflict in Darfur, Sudan.

In 2006, WFP emergency operations provided food aid for 63.4 million people caught up in the everwidening net of humanitarian disasters. To respond successfully to emergencies entailing food shortages WFP must be able to mobilize its response quickly and efficiently, making preparedness a critical component of any emergency operation (HORTON, 2008).

Useful information to anticipate food security crises and launch rapid needs assessments, especially in natural disasters, comes from the use of advanced Geographical Information Systems (GIS) and remote sensing. WFP started to use in an effective way these instruments by developing a network of partnerships with academic research institutions, governments and UN agencies working in applied technology (remote sensing, GIS).

1.2 CONTRIBUTE OF REMOTE SENSING TO THE MANAGEMENT OF NATURAL DISASTERS

Remote sensing plays a key role for the providing of information suitable for regional planners for the management of natural disasters in developing countries. It is useful not only in the planning processes in general, but also to detect and map many types of natural hazards. Most remote sensing studies dealing with natural hazards concern the vulnerability of the areas, the monitoring of events which could cause a disaster and the assessment of the damages.

In the pre-disaster phase, the collection and analysis of relevant data and of information is useful to identify causal factors (both natural and human) which contribute to the hazard and develops a sufficient knowledge of the processes and of the causal factors leading to the hazard. Moreover, remote sensing instruments can provide useful input data for early warning systems dedicated to the evaluation of the impacts of natural components variability (e.g. the climate) on natural disasters occurrences, constituting natural disaster predictions systems.

In the response and post-disaster phases, the systematic observation of the environment can provide support to monitor the effects of a natural disaster on the territory (estimate of the changes possibly occurred) and to identify the affected population. In this case, the extracted information is used to produce suitable cartographic outputs.

A great variety of remote sensing instruments is mounted on satellite platforms (see table 1), employing various measurement technologies and techniques and acquiring data in a wide range of the electromagnetic spectrum. Therefore, space technologies ensure the measurement of a large range of geophysical parameters and the observation of a lot of natural phenomena.

TABLE 1.1 - REMOTE SENSING INSTRUMENTS CATEGORIES.

REMOTE SENSING INSTRUMENTS CATEGORIES
Atmospheric chemistry instruments
Atmospheric Temperature and humidity sounders
Cloud profile and rain radars
Earth radiation budget radiometers
High resolution optical imagers
Imaging multispectral radiometers (VIS/IR)
Imaging microwave radars
Imaging multi-spectral radiometers (passive microwave)
Lidars
Multiple direction/polarization instruments
Ocean color instruments
Radars altimeters
Scatterometers
Gravity, magnetic fields and geodynamics instruments
Meteorological missions

In the next paragraphs a short review of the main remote sensing data and its applications in disaster management is provided.

1.2.1 REMOTE SENSING MISSION FEATURES

In order to define the effectiveness of a remote sensing system for the production thematic maps for disaster management activities, two main factors have to be considered:

- sensor features*:* spatial and spectral resolution of the sensor and area of coverage;
- system features*:* temporal resolution (or revisit time) of the system, data cost and data availability.

Spatial resolution refers to the discernible detail in the image. Spatial data resources need to be matched with the hazard characteristics, and for risk assessment, also with those of the elements at risk. Ideally, it would be advisable to use a "multi-scale" approach in order to carry out post-disaster assessment activities. Therefore, the use of high spatial resolution imagery and ground checks allows a more detailed knowledge of target sites, while the use of medium and low spatial resolution data allows performing synoptical analyses. Systems according to their spatial resolution can be classified as very high, high, medium and low resolution.

Spectral resolution needs to be taken into consideration for the selection of the type of data to be used, since the different sensors are designed to cover different spectral regions. Spectral resolution refers to the bands range or the bands width acquired by the sensor. In the work of thesis radar and optical systems (visible/infrared, panchromatic and multispectral) are considered. It is noticed that optical sensors have limitations in cloudy conditions as the objects on the ground surface may be obscured from view. In that case, radar systems provide a more reliable data: long wavelengths can pass through the cloud coverage, smoke and fog, ensuring useful acquisitions of the targets in any weather conditions.

The frequency, or **revisit time,** is the average time needed by the system to re-fly over the same ground point of the terrestrial surface. This system feature is linked to the orbit cycle, the number of satellites and the de-pointing capability. Hazard events can be sudden and of short duration (e.g. earthquakes or landslides), sudden but of long duration (e.g. a dam break leading to prolonged flooding), but can also show precursory signs (e.g. volcanic activity or hurricanes). Some events, such as earthquakes, may also show a repetitive pattern, where violent aftershocks may affect areas already destabilized by the primary event. Some effects may also be delayed, such as disease outbreak after a flood or earthquake. It is necessary to have a good understanding of the spatial, spectral and temporal characteristics of the hazard(s) under consideration, before deciding on a specific analysis type and data requirements.

In a contest of humanitarian assistance, a not negligible factor is the cost of the satellite data. Thus low cost and free of charge satellite data are preferred for the analysis of the areas that can be or have been hit by a natural disaster.

According to the described features, suitable satellite sensing systems, which are considered in this work of thesis, are classified as follows:

- **Very high to high spatial resolution sensors** provide detailed images of the Earth's surface. In general, these instruments have a spatial resolution in the range 0.5 to 2 m (very high resolution, VHR) or 2 to 10 m (high resolution, HR) and a kilometric swath width. High resolution imagers are panchromatic (PAN, they acquire data in a single band that covers the e.m. spectrum range from Visible to the Near Infrared) and multispectral (XS, they acquire simultaneously data in some spectral bands in the Visible and in the Infrared ranges) sensors. Panchromatic images are especially devoted to the extraction of the geometric content, while multispectral images, which hold a greater spectral information content, are more suitable for tasks such as land cover classifications. As aforesaid, the use of these sensors can be limited by weather conditions (they are unable to pass through cloud, rain or fog). Moreover, some systems have pointing capabilities which enable imagery of specified areas to be acquired more frequently. Some high resolution instruments can also provide stereo images by using data collected on a single orbit (along track). Stereo images allow 3D feature extraction. Common applications of high and very high resolution satellite data are map generation and updating and generation of digital elevation models.

- Visible (VIS) and Infrared (IR) imaging **multispectral** sensors acquire images of the Earth's atmosphere and surface providing accurate spectral information, with spatial resolution varying from 10 m up to several km (medium spatial resolution, MR and low resolution, LR) and with a swath width generally in the range several hundred to a few thousand km. Data are usually acquired in multiple narrow, precisely calibrated spectral channels. The images acquired by these instruments operating in VIS and IR ranges are suitable for environmental analyses, allowing to perform accurate land cover classifications and to extract a wide range of parameters useful in many research fields. Hyper spectral instruments are able to simultaneously acquire imagery in many tens of narrow bands, improving the quality of land cover and land use information derived.
- Generally, **low resolution** multispectral sensors are considered an important source of data for a lot of studies dealing with processes in the biosphere, since they provide information at a global scale. For example, classification and seasonal monitoring of vegetation types on a global basis allows modeling of primary production (the growth of vegetation). Such information is of great value in supporting the identification of drought areas and in providing early warning on food shortages.
- **Imaging microwave radars** instruments transmit at frequencies of around 1 to 10 GHz and measure the backscattered signals to generate microwave images of Earth's surface at a spatial resolution between 2 m and 100 m and swath width of around 100 km. Both synthetic aperture radars (SARs) and some real aperture side-looking imaging radar systems fall into this category. Radars operating in a variety of wavelengths (usually L-, C- and X- bands) are available. X-band radar systems are the most commonly offered by the commercial sector. Radars have the capability to

pass through clouds providing data in any weather condition, day/night basis. SARs also have the ability to pass through vegetation and to sample surface roughness and surface dielectric properties. Data provided by multipolarized SARs (such as ASAR on ENVISAT) allow accurate land cover classification of terrestrial surface and provide quantitative measurements of biophysical parameters, such as soil moisture and biomass. Therefore, SARs systems are suitable for a lot of resource monitoring applications. A technique known as interferometry is used to record the phase shift between two SAR images recorded at slightly different times, thereby providing accurate information on the motion of surfaces and targets and allowing 3D information extraction.

1.3 APPLICATIONS OF REMOTE SENSING TO NATURAL DISASTERS

According to mission features, remote sensing can have different applications for the risk management (see table 1.2). In this section a review of the uses of remote sensing for the management of the main natural disasters is provided.

TABLE 1.2 – SATELLITE IMAGERY FOR THE MITIGATION OF ONSET NATURAL DISASTERS

	FLOOD	EARTHQUAKE	VOLCANIC ERUPTION	LANDSLIDE	TSUNAMI and HURRICANES	Desertification and Drought
Information to be obtained	Affected and flooded areas maps and damage estimate	Damage estimate	Maps of areas vulnerable to lava flows, ash and debris fall; lava flows and fires maps and damage estimate	Slide area maps ad damage estimate	Flooded areas maps and damage estimate	Maps of areas in the course of desertification or subjected to drought
Sensor type and spectral band	Radar and optical (PAN, NIR and TIR bands)	Optical (PAN and MS)	Radar and optical (PAN and MS)	Radar and optical (PAN and MS)	Radar and optical (PAN, NIR and TIR bands)	Radar and optical (PAN, NIR and TIR bands)
Spatial resolution	VHR, HR, MR and LR	VHR and HR	VHR and HR	VHR and HR	VHR, HR, MR and LR	LR
Synoptic view	Yes	No	No	No	Yes	Yes

In the next part of this section a brief summary of the use of remote sensing for natural disaster is given, referring to 6 categories: floods, hurricanes, earthquakes, volcanic eruptions, landslides and desertification.

1.3.1 FLOODS, TSUNAMIES AND HURRICANES

Floods are the most common of natural hazards that can affect people, infrastructure, and the natural environment. They can occur in many ways and in many environments. Riverine floods, the most prevalent, are due to heavy, prolonged rainfall, rapid snowmelt in upstream watersheds, or the regular spring thaw. Other floods are caused by extremely heavy rainfall occurring over a short period in relatively flat terrain, the backup of estuaries due to high tides coinciding with storm surges, dam failures, dam overtopping due to landslides into a reservoir, and seiche and wind tide effects in large lakes. Occasionally an eruption on a glacier or snow-covered volcanic peak can cause a flood or a mudflow in which the terrain is radically changed and any agrarian development is totally destroyed, frequently with much loss of life.

Even if it is impossible to define the entire flood potential in a given area, remote sensing can provide the evidence for potential flood situations. The most obvious evidence of a major flood potential, outside of historical evidence, is the identification of floodplain or flood-prone areas which are generally recognizable on remote sensing imagery. The most valuable application of remote sensing to flood hazard assessments, then, is in the mapping of areas susceptible to flooding in a risk assessment phase and immediate information of the areas affected by a flood event in a phase of damage assessment.

The main used technique is the comparison between pre and post event satellite imagery, that can be derived both from optical sensors and radar ones. Optical images are of easy interpretation but are affected by the presence of cloud coverage. The radar imagery, instead, can satisfactorily penetrate the clouded sky and define many floodplain features. Moisture on the ground noticeably affects the radar return and, together with the textural variations emphasized by the sensor, makes radar a potentially desirable tool for flood and floodplain mapping, but the only disadvantage of the time-availability after the event.

Furthermore high resolution data coming from VHR satellite can help to understand the situation at a bigger scale for the damage assessment phase

Tsunamis and hurricanes are often related to flood events; therefore remote sensing needs for response activities are similar. The determination of hurricane paths in the area of interest can be derived from remote sensing systems specially designed for meteorological purposes (which are not here considered). Low resolution optical satellite data are also interesting because of their monitoring capabilities arising from a large swath width and a small revisit time.

1.3.2. EARTHQUAKES

An earthquake is a tremor of the earth's surface usually triggered by the release of underground stress along fault lines. This release causes movement in masses of rock and resulting shock waves. In spite of extensive research and sophisticated equipment, it is impossible to predict an earthquake, although experts can estimate the likelihood of an earthquake occurring in a particular region. The greatest danger of an earthquake comes from falling buildings and structures and flying glass, stones and other objects.

The planning of development in earthquake-prone areas is laden with problems. There are large human settlements already located in earthquake/prone areas.

As with other geologic hazards, the frequency of occurrence can fall in cycles of decades or centuries. Earthquakes are particularly difficult to predict at this time. Thus, mitigation emphasis is on land use planning (non-intensive uses in most hazardous areas), on building strength and integrity, on response planning, and on incorporating mitigation measures into reconstruction efforts. The main problem is the identification of the earthquake damage-prone zones. While in most areas of great earthquake activity some seismic information is available, it may not be sufficient for planning purposes. Remote sensing techniques and resulting data interpretation can play a role in providing additional information.

In order to identify earthquake hazards optical medium resolution imagery has been effectively and widely used since it is less expensive and more readily available than other remote sensing data. Airborne and satellite radar mosaics have been successfully used for the delineation of fault zones. Radar is applicable to delineate unconsolidated deposits sitting on fault zones-upon which most of the destruction occurs-to identify areas where an earthquake can trigger landslides. This is best accomplished on stereo models using adjoining and overlapping radar flight lines. Conventional aerial photography and high resolution satellite imagery are also used for this purpose.

1.3.3 VOLCANIC ERUPTIONS AND RELATED HAZARDS

During a volcanic eruption, lava, tephra (ash, lapilli, solid chunks of rock), and various gases, are expelled from a volcanic vent or fissure. Many hazards are associated with the conditions brought about by volcanic activity. Active volcanoes pose hazards which include the immediate release of expelled ash, lava, pyroclastic flows, and/or poisonous hot gases; volcanic earthquakes; and the danger of mudflows and floods resulting from the rapid melting of snow and ice surrounding the vent during eruption. Some secondary hazards may threaten during volcanic activity or during periods of dormancy. These include landslides due to unstable accumulations of tephra, which may be triggered by heavy rains or by earthquakes.

Remote sensing data interpretation allows the analysis of past catastrophic events associated with recently active volcanoes. This information, together with available historical data, can be used in some areas of interest for the assessment of potential volcano-related hazard.

The small area involved with volcanic activities permits the use of VHR and HR sensors for its analyses. Panchromatic stereo images are suitable to extract and to map geomorphic evidence of recent activity and associated hazards, while optical data (VIS and IR) are suitable to prepare volcanic hazard maps (based on the prediction of the path of potential mudflows or lava flows coupled with population density information) and to produce damage assessments.

1.3.4 LANDSLIDES

Landslides, or mass movements of rock and unconsolidated materials such as soil, mud, and volcanic debris, are a common and diffused hazard. The mass movement of bed rock and unconsolidated materials results in different types of slides, magnitudes, and rates of movement. VHR and HR images are required to map landslide areas and damages. In this case, stereoscopic coverage is helpful, coupled with VIS and IR data.

1.3.5 DESERTIFICATION AND DROUGHT

Desertification occurs when an ecosystem experiences a diminution or loss of productivity. This process can have a natural and an anthropic component, which may reinforce each other, creating a synergetic effect. The degree of desertification risk is directly related to certain natural conditions such as climate, topography, natural vegetation, soil, and hydrology, as well as to the intensity and type of anthropic activity in the area.

Drought is an extended period of months or years when a region notes a deficiency in its water supply. Generally, this occurs when a region receives consistently below average precipitation. It can have a substantial impact on the ecosystem and agriculture of the affected region.

Remote sensing, both spaceborne and airborne, provides valuable tools for evaluating areas subject to desertification and drought. Digital data can be used for the purpose of locating, assessing, and monitoring deterioration of natural conditions in a given area. Information about these conditions can be obtained from direct measurements or inferred from indicators (keys to the recognition of a desertification process). Generally low resolution data are used in order to provide a synoptic view of the areas that can be prone to these problems, during the first stages of a detailed desertification or drought study.

1.4 AVAILABLE FREE SATELLITE DATA

In the field of disaster management free satellite data play a key role. Nowadays the main used global data free of charge are:

- Google Earth. It is one of the most known and used satellite imagery is the one collected in Google Earth. In this system are collected the highest resolution and most recently available satellite images, but only shown as raster pictures. It is not possible to perform synthesis or analysis on the displayed data, but it can have different uses, for example as a reference for the interpretation of other available imagery at a lower spatial resolution such as Landsat or MODIS (see below) or as a base layer for GIS analysis.
- Global DEMs (Digital Elevation Models). There are two main sources for global DEMs: the older GTOPO30, or the more recent SRTM-based DEM. The GTOPO30 (http://edcdaac.usgs.gov/gtopo30/ gtopo30.html) is a coarse, global DEM, with grid cells of 1km across. The data are georeferenced, providing a good backdrop for large areas where fine detail is not necessary. A more detailed source for DEMs is the dataset collected in 2000 during the Shuttle Radar Topography Mission (SRTM). A radar pair mounted on a space shuttle mapped nearly the entire globe at 30m resolution (http://www.jpl.nasa.gov/srtm/). Data outside the US are degraded to 90m.
- Advanced Very High Resolution Radiometer (AVHRR). The AVHRR is operating in an older satellite mission, having already flown, with continuously replaced instruments, for over 20 years. It provides better than daily coverage, at a resolution of 1.1km per cell at nadir. It is an excellent tool for frequent mapping at regional scale. Because of the frequent observations a very large archive exists. More details about this dataset are provided in the next chapters.
- Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). ASTER, launched in 1999, carries a spectacular 15 channels, with 4 bands at 15 m resolution, 6 at 60m, and 5 at 90m. Furthermore, because of the structure of the satellite system, the data can be used to create DEMs. The best way to search for ASTER data is via the WIST Warehouse Inventory Search Tool (WIST,

https://WIST.echo.nasa.gov/), where also other data from NASA or NOAA-operated satellites can be found. Now they cost a nominal modest fee of 80US$ per scene. However, educational organizations can apply for free data (see http://lpdaac.usgs.gov/aster/afd/index.php).

- Moderate Resolution Imaging Spectroradiometer (MODIS). There are actually MODIS sensors on two different satellites, acquiring data at moderate resolution in a remarkable 36 channels. The resolution is variable, with some bands at 250m, some at 500, and some at 1,000m. The coverage of MODIS is 2,230km, thus very large regions can be monitored daily. More information is given about this dataset in the next chapter.
- Landsat MSS/TM data. One of the oldest and best known satellites missions is Landsat, which has been providing Earth surface data since 1972. Initially the data had a resolution of 60 m, which was later improved to 30m (lower in the thermal bands). The last satellite includes a 15m panchromatic band, and a thermal band that provides 60m data. For many years the data were also commercially sold, and at several thousand dollars per scene very expensive. Recently the US government decided to make all Landsat data, including the entire archive, available free of charge.
- SPOT VEGETATION. The SPOT program consists of a series of optical remote sensing satellites with the primary mission of obtaining Earth imagery for land use, agriculture, forestry, geology, cartography, regional planning, water resources and GIS applications. SPOT mission offers commercial data a different resolutions, but they render available free of charge VEGETATION data which are older than 3 months, with a spatial resolution of 1.1 km. They are distributed by VITO in Belgium (http://free.vgt.vito.be/). More information about these data is provided in the next chapters.

1.5 IMAGE PROCESSING AND OUTPUT FORMATS

In this work of thesis particular attention is focused on the automation of procedures of image processing, since it is necessary to elaborate huge quantities of data and produce in near real time information on the status of the ground.

Before the use of satellite data for the extraction of features of interest it is necessary to render images derived from different sensors and at different times comparable the ones with the others. The main operations developed on satellite images during the thesis are ones usually developed in Digital Image Processing:

- image restoration;
- image enhancement;
- image classification;
- Image transformation.

Image restoration is concerned with the correction and calibration of images in order to achieve as faithful a representation of the earth surface as possible. In this phase it is possible to include all the treatments necessary to render satellite images coming from different sensors and different time of acquisition comparable the ones with the others (this are geometric, radiometric calibration, atmospheric correction, georeferencing). Image enhancement is predominantly concerned with the modification of images to optimize their appearance to the visual system. Image classification refers to the computer-assisted interpretation of image. Finally, image transformation refers to the derivation of new imagery as a result of some mathematical treatment of the raw image bands.

In order to satisfy the necessities of final user of the extracted information, it is also important to consider the output format of the final product. Among available format it is possible to resume the possible outputs in three categories: vector data, raster data and analysis tools (such as web-GIS).

Vector data are composed of discrete coordinates that can be used as points or connected to create lines and polygons. Coordinates are typically provided in geographic format (latitude/longitude) or projected (for example UTM). The points, defined as discrete location on the surface of the planet, are represented by an x-y coordinate pair. Each point on the map is created by latitude and longitude coordinates, and is stored as an individual record in the vector file. The lines are formed by connecting two data points The computer reads this line as straight, and renders the line as a vector connecting two x-y coordinates (X = longitude, Y = latitude). The more points used to create the line, the greater the detail. The polygons can be defined as areas fully encompassed by a series of connected lines. Because lines have direction, the system can determine the area that falls within the lines comprising the polygon. All of the data points that form the perimeter of the polygon must connect to form an unbroken line. Because its limited dimension on the disk, this kind of data is useful for a rapid exchange of information on features and can be easily integrated in a GIS environment.

Raster data represent features as a matrix of cells within rows and columns in continuous space. These cells are formed by pixels of a specific dimension size. Each raster data layer represents one attribute. Most analyses combine these layers to create new layers with new cell values, as either continuous or discrete data. Continuous data types have gradations, such as temperature or elevation. Discrete data types have clearly delineated boundaries, such as a city boundary or specific vegetation type. The cell size used for a raster layer affects the results of the analysis and how the map looks. Using too large a cell size will cause some information to be lost. Using too small a cell size will significantly increase the storage space and processing time required, without adding precision to the map.

Among analysis tools, web-GIS applications assume particular relevance. A web-GIS application can be defined as a complex system with access to the internet for capturing, storing, integrating, manipulating, analyzing and displaying data related to locations without the need of having a proprietary GIS software. This kind of analysis tools can be an advisable system of analysis for users who do not have GIS software for the analysis of georeferenced data.

CHAPTER 2: BASE DATA

The aim of this chapter is to provide information about the datasets used in the work of thesis. Mainly data deriving from the acquisition of three different sensors are used: MODIS (Moderate Resolution Imaging Spectroradiometer), AVHRR (Advanced Very High Resolution Radiometer) AND SPOT (Système Pour l'Observation de la Terre) Vegetation datasets.

2.1 THE MODIS SYSTEM

The MODIS is an instrument operating on Terra and Aqua satellites, which make part of EOS (Earth Observing System), the first observing system to offer integrated measurements of the Earth's processes. It is composed by a science component and a data system supporting a coordinated series of polar-orbiting and low-inclination satellites for long-term global observations of the land surface, biosphere, solid Earth, atmosphere, and oceans.

The EOS was launched by NASA for the Earth-Sun System Mission, which aims at providing objective scientific data and analysis to advance the understanding of Earth-Sun processes, in order to improve prediction capability for climate, weather and natural hazards.

MODIS continues the lineage of the Coastal Zone Color Scanner (CZCS), the Advanced Very High Resolution Radiometer (AVHRR), the High Resolution Infrared Spectrometer (HIRS), and the Thematic Mapper (TM).

2.1.1 TERRA AND AQUA SATELLITE MISSION

The MODIS instrument is a high signal-to-noise instrument designed to satisfy a diverse set of oceanographic, terrestrial, and atmospheric science observational needs.

It was completed in 1995. Since then, two spaceflight units, the Protoflight Model (PFM, aboard the Terra Satellite), and the Flight Model 1 (FM1, aboard the Aqua Satellite) have been completed and launched. Terra was launched on December 18, 1999, and Aqua was launched on May 4, 2002.

Terra orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. Terra and Aqua spacecrafts are in a near polar, sun-synchronous orbit at an altitude of 705 km. The Earth scene is perpendicular to the ground track and subtends a scan angle of 110 degrees. The orbit is adjusted so that it covers the complete Earth every 16 days and will be maintained with periodic adjustments. Different instruments are located on each satellite (table 2.1).

TABLE 2.1: TERRA-AQUA CHARACTERISTICS

SATELLITE	TERRA	AQUA
Launch:	December 18, 1999	May 4, 2002
	Western Test Range, Vandenberg Air Force	Western Test Range, Vandenberg Air Force Base
Orbit:	Altitude: 705 km	Altitude: 705 km
	Inclination: 98.2 degrees	Inclination: 98.2 degrees
	Period: 98.9 minutes	Period: 99 minutes
	Repeat Cycle: Ground Track Repeat: 16 days	Repeat Cycle: Ground Track Repeat: 16 days
	Sun-Synchronous	Sun-Synchronous
Vital Statistics:	Weight: 5,190 kg	Weight: 3,117 kg
	Size: 3.5 meters	
	Power: 2,530 watts	Power: 4,444 watts
	Design Life: 5 years	Design Life: 6 years
Instruments:	ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer	AIRS: Atmospheric Infrared Sounder
	CERES: Clouds and the Earth Radiant Energy System	AMSR-E: Advanced Microwave Scanning Radiometer-EOS
	MISR: Multi-angle Imaging SpectroRadiometer	AMSU-A: Advanced Microwave Sounding Unit-A
	MODIS: Moderate Resolution Imaging Spectroradiometer	CERES: Clouds and the Earth Radiant Energy System
	MOPITT: Measurements of Pollution in The Troposphere	HSB: Humidity Sounder for Brazil
		MODIS: Moderate Resolution Imaging Spectroradiometer

Terra simultaneously studies clouds, water vapour, aerosol particles, trace gases, terrestrial and oceanic properties, the interaction between them and their effect on atmospheric radiation and climate.

Aqua data include information on water vapour and clouds in the atmosphere, precipitation from the atmosphere, soil wetness on the land, glacial ice on the land, sea ice in the oceans, snow cover on both land and sea ice, and surface waters throughout the world oceans, bays, and lakes. Aqua also provides observations on vegetation cover on the land, phytoplankton and dissolved organic matter in the oceans, and the temperatures of the air, land, and water.

2.1.2 MODIS SENSOR

MODIS is a passive, imaging spectroradiometer, arranged in 36 spectral bands, which cover the visible and infrared spectrum ranging from 0.412 m to 14.235 m.

The acquisition of the narrow-band radiance observations over 36 spectral regions is realized by a scan mirror which views the Earth and by three internal calibrators.

All bands are acquired and digitized at 12 bits. Spatial resolution depends on the considered band acquired: two bands (bands 1-2) are imaged at a nominal resolution of 250 m at nadir, five (bands 3-7) at 500 m and the remaining (bands 8-36) at 1000 m (see table 2.2). In the case of 1km bands, 10 detectors are arranged in an along track linear array for these. In the case of 500 m bands (bands 3-7), there are 20 detectors arranged in an along track linear array, while for 250 m bands (bands 1 & 2) have there are 40 detectors in the along track direction (see figure 2.1).

There are three calibrator systems inside the MODIS instrument: a Solar Diffuser (SD) with a Solar Diffuser Stability Monitor (SDSM), a Spectroradiometric Calibration Assembly (SRCA); and a Blackbody (BB). In addition there is a Space View (SV) port that is used to provide a zero reference.

TABLE 2.2: MODIS BANDS AND THEIR POSSIBILE USES

PRIMARY USE	BAND	BANDWITH [nm]	Radiometric resolution	Geometric resolution[m]
Land/Cloud/Aerosols Boundaries	1	620 - 670	12 bits	250
	2	841 - 876	12 bits	250
Land/Cloud/Aerosols Properties	3	459 - 479	12 bits	500
	4	545 - 565	12 bits	500
	5	1230 - 1250	12 bits	500
	6	1628 - 1652	12 bits	500
	7	2105 - 2155	12 bits	500
Ocean Color/ Phytoplankton/ Biogeochemistry	8	405 - 420	12 bits	1000
	9	438 - 448	12 bits	1000
	10	483 - 493	12 bits	1000
	11	526 - 536	12 bits	1000
	12	546 - 556	12 bits	1000
	13	662 - 672	12 bits	1000
	14	673 - 683	12 bits	1000
	15	743 - 753	12 bits	1000
	16	862 - 877	12 bits	1000
Atmospheric Water Vapor	17	890 - 920	12 bits	1000
	18	931 - 941	12 bits	1000
	19	915 - 965	12 bits	1000
Surface/Cloud Temperature	20	3.660 - 3.840	12 bits	1000
	21	3.929 - 3.989	12 bits	1000
	22	3.929 - 3.989	12 bits	1000
	23	4.020 - 4.080	12 bits	1000
Atmospheric Temperature	24	4.433 - 4.498	12 bits	1000
	25	4.482 - 4.549	12 bits	1000
Cirrus Clouds	26	1.360 - 1.390	12 bits	1000
Water Vapor	27	6.535 - 6.895	12 bits	1000
	28	7.175 - 7.475	12 bits	1000
Cloud Properties	29	8.400 - 8.700	12 bits	1000
Ozone	30	9.580 - 9.880	12 bits	1000
Surface/Cloud Temperature	31	10.780 - 11.280	12 bits	1000
	32	11.770 - 12.270	12 bits	1000
Cloud Top Altitude	33	13.185 - 13.485	12 bits	1000
	34	13.485 - 13.785	12 bits	1000
	35	13.785 - 14.085	12 bits	1000
	36	14.085 - 14.385	12 bits	1000

Scene radiant flux is reflected by the double sided, beryllium scan mirror, which is continuously rotating at 20.3 rpm with a period maintained to 0.001 sec so as to control scan to scan underlay, that is, one side of the mirror traverses 360 degrees every 1.477 seconds. It is oval shaped, 21 cm wide (the axis of rotation) and 58 cm long. The mirror is nickel plated and coated with silver for high reflectance and low scatter over the broad spectral range of the sensor (see figure 2.1, 2.2). The reflectivity of the each side of the scan mirror is a function of the angle of incidence.

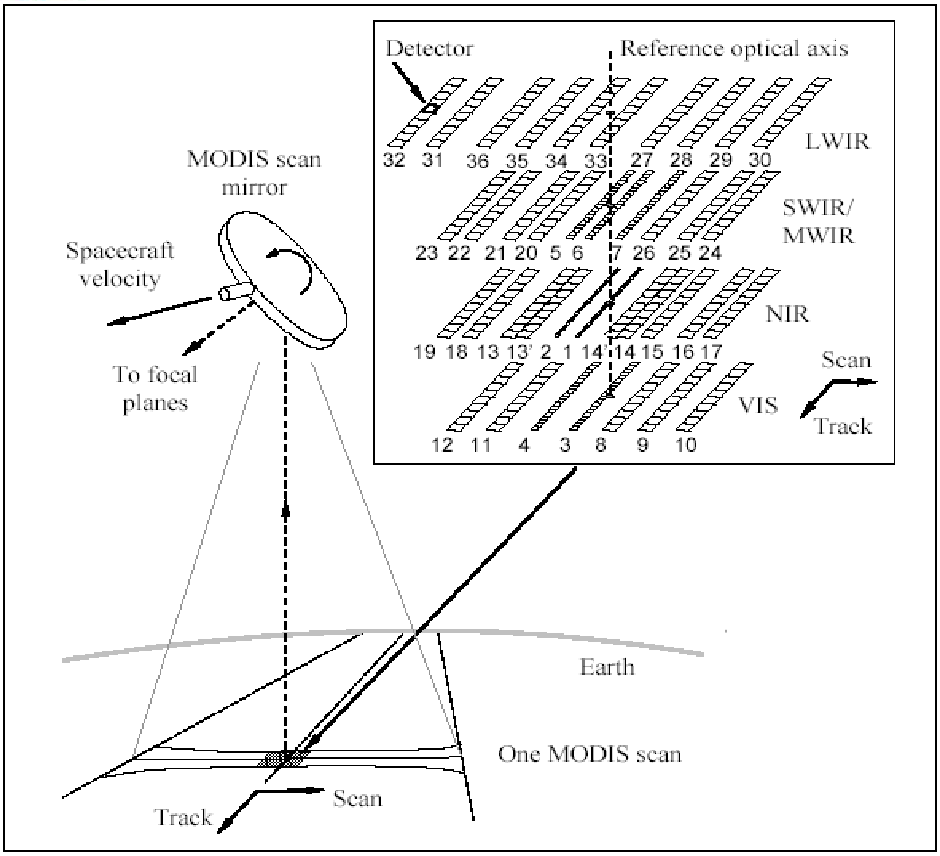

FIGURE 2.1: MODIS SCAN GEOMETRY

Immediately after the secondary mirror is a dichroic beamsplitter assembly (consisting of three beamsplitters) that directs the energy through four refractive objective assemblies and then onto the four focal plane assemblies with their individual bandpass filters. The beamsplitters are used to achieve spectral separation, dividing the MODIS spectral domain into four spectral regions: visible (VIS) (0.412 to 0.551 m), near infrared (NIR) (0.650 to 0.940 □m), short wavelength/medium wavelength infrared (SWIR/MWIR) (1.240 to 4.565 m), and long wavelength infrared (LWIR) (6.715 to 14.235 m).

Each spectral region has an objective lens assembly for imaging scene energy onto the corresponding focal plane. On each focal plane there are rows of detectors aligned (ten each for the 1 km bands, twenty each for the 500m bands, and forty each for the 250 m bands).

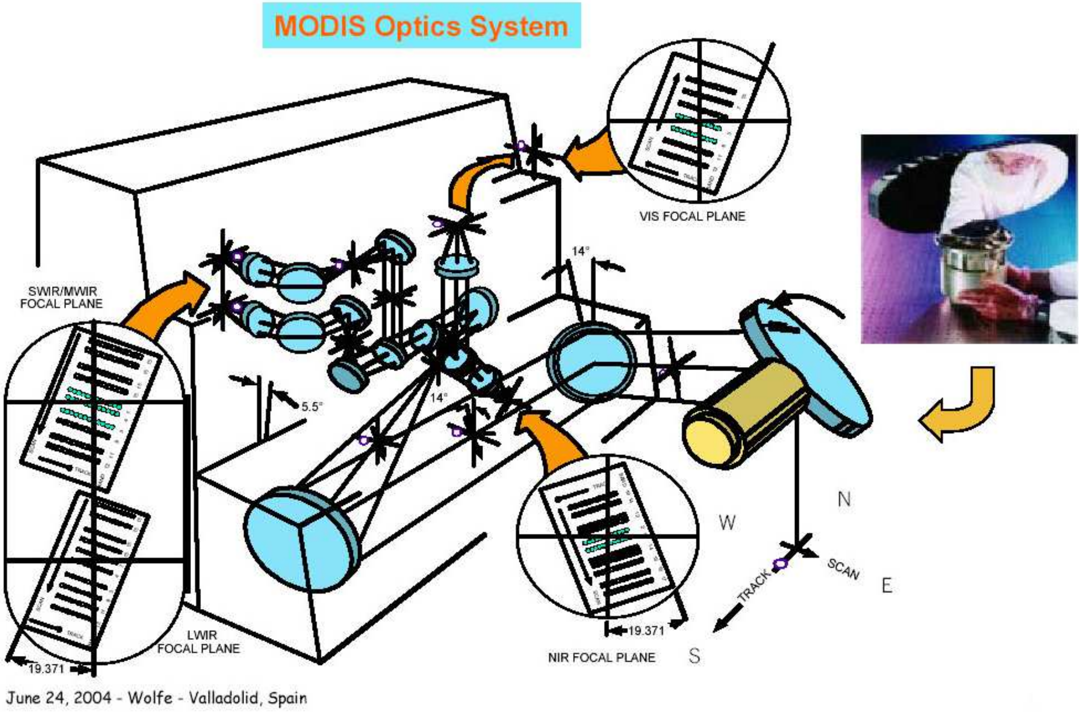

FIGURE 2.2:MODIS OPTIC SYSTEM

A MODIS scan line is supplied 2330 km long and 10 km wide at NADIR. It results orthogonal to the satellite track direction.

2.1.3 MODIS PRODUCTS

The data directly coming from the acquisition are primary products. From them can be generated the so called 'derived data', containing information at different spatial and temporal resolutions designed to meet the needs of the research community.

Primary products are made of the raw images of radiance/reflectance and geolocation files, which contain geometric references every kilometre.

Derived products are essentially divided in three classes:

- Atmosphere data products, data concerned with atmosphere events: cloudiness, precipitation and aerosols;
- Land data products, regarding phenomena like energy balance of earth, coverage of the soil, land thermal properties;
- Ocean data products, containing information about temperature, ocean primary productivity, presence of phytoplankton.

The name of all products in detail are listed in table 2.3

TABLE 2.3 MODIS DATA PRODUCTS

MODIS DATA PRODUCTS	
PRIMARY DATA	
MOD01 Level-1A Radiance Counts MOD02 Level-1B Calibrated Relocated Radiances MOD03 Relocation Data Set	
DERIVED PRODUCTS	
ATMOSPHERE DATA PRODUCTS MOD04 Aerosol Product MOD05 Total Precipitable Water MOD06 Cloud Product MOD07 Atmospheric profiles MOD08 Gridded Atmospheric Products (Level-3) MOD35 Cloud Mask	OCEAN DATA PRODUCTS MOD18 Normalized Water-leaving Radiance MOD19 Pigment Concentration MOD20 Chlorophyll Fluorescence MOD21 Chlorophyll_a Pigment Concentration MOD22 Photosynthetically Active Radiation (PAR) MOD23 Suspended-Solids Conc, Ocean Water
LAND DATA PRODUCTS MOD09 Atmospherically-corrected Surface Reflectance MOD10 Snow Cover MOD11 Land Surface Temperature & Emissivity MOD12 Land Cover/Land Cover Change MOD13 Vegetation Indices MOD14 Thermal Anomalies, Fires & Biomass Burning MOD15 Leaf Area Index & FPAR MOD16 Surface Resistance & Evapotranspiration MOD17 Vegetation Production, Net Primary Productivity MOD44 Vegetation Cover Conversion	MOD24 Organic Matter Concentration MOD25 Coccolith Concentration MOD26 Ocean Water Attenuation Coefficient MOD27 Ocean Primary Productivity MOD28 Sea Surface Temperature MOD29 Sea Ice Cover MOD31 Phycoerythrin Concentration MOD32 Processing Framework & Match-up Database MOD36 Total Absorption Coefficient MOD37 Ocean Aerosol Properties MOD39 Clear Water Epsilon MOD43 Albedo 16-day L3

2.1.4 MODIS DATA LEVELS

MODAPS (MODIS Data Processing System) provides MODIS data in a variety of levels. Primary data are stored in the first three levels, while derived product in the remaining ones (see figure 2.3).

Levels of storage of primary data:

- Level 0: Reconstructed unprocessed instrument/payload data at full resolution; any and all communications artefacts (e.g., synchronization frames, communications headers) removed.

- Level 1A Reconstructed unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended, but not applied, to the Level 0 data.
- Level 1B: Level 1A data that have been processed to sensor units (not all instruments have a Level 1B equivalent).

Levels for derived data:

- Level 2: Derived geographic variables at the same resolution and location as the Level 1 source data.
- Level 3: Variables mapped on uniform space-time grid scales, usually with some completeness and consistency.
- Level 4: Model output or results from analyses of lower level data (e.g., variables derived from multiple measurements).

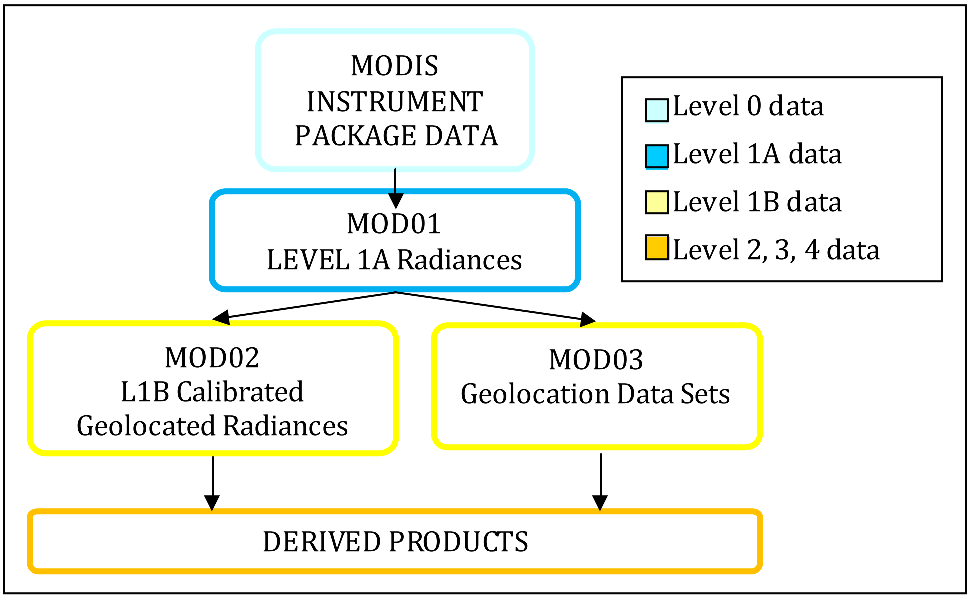

FIGURE 2.3: MODIS DATA HIERARCHY

Level 0 data is straight from the instrument, raw radiance values. This form is not friendly to most image processing systems and requires custom code to transform it into something more end user friendly. This product is not made available to the general public. Level 1 products are the lowest level of product available to the end user. These products are in 5 minutes chunks of sensor acquisition and are still in swath format.

The *Level 1A* code organizes a 2-hour Level 0 file into a set of scenes, each containing approximately 5 minutes of MODIS data.

Level 1A data sets contain counts for 36 MODIS channels, along with raw instrument engineering and spacecraft ancillary data. The Level-1A data are used as input for geolocation, calibration, and processing. Quality indicators are added to the data to indicate missing or bad pixels and instrument modes.

The *Level 1B* software, starting from Level 1A data, implements calibration algorithms to correct raw, Earth view-sector data for all known instrumental effects and to transform the corrected data into calibrated products, organizes the calibrated science data into a scene, performs quality assurance tests on the raw and calibrated data.

The Level 1B products (MOD02) contain calibrated and geolocated radiances, in W/ (m2 μm sr) and reflectances for 36 bands generated from MODIS Level 1A sensor counts (MOD01). Geolocation information stored in level 1B is derived from the MOD03 product. The MOD03 (MODIS Geolocation product), also derived from Level 1A data, contains geodetic coordinates, ground elevation, and solar and satellite zenith, and azimuth angle for each MODIS 1-km sample the file. This geolocation information is determined using the spacecraft attitude and orbit, instrument telemetry, and a digital elevation model.

The next step in the processing chain is *Level 2*. MODIS Level 2 products are corrected for atmospheric contamination (aerosols, and geo angles).

Level 3 and 4 products are end-user value added products, nearly all of which were derived from composites of some kind.

Available derived products can be stored in two different data formats: swath or grid. Level 2 data can be available in both the formats, while level 3 and level 4 data are only in grid format.

The *Swath format* is not a projection: it is the view as the sensor “sees” the Earth. One swath scene is the result of a 5-minute scan.

Since one scan of the MODIS mirror takes 1.4771 seconds, there are typically 203 full scans within a 5-minute product file and occasionally 204 full scans. The use of an occasional 204-scan file keeps the number of 5-minute granules or swath scenes per day at 288, assuming no data gaps (see figure 2.4). Though this product is available to the end user, special tools are still required to manipulate these data into a map projection and to correct for the MODIS panoramic distortion.

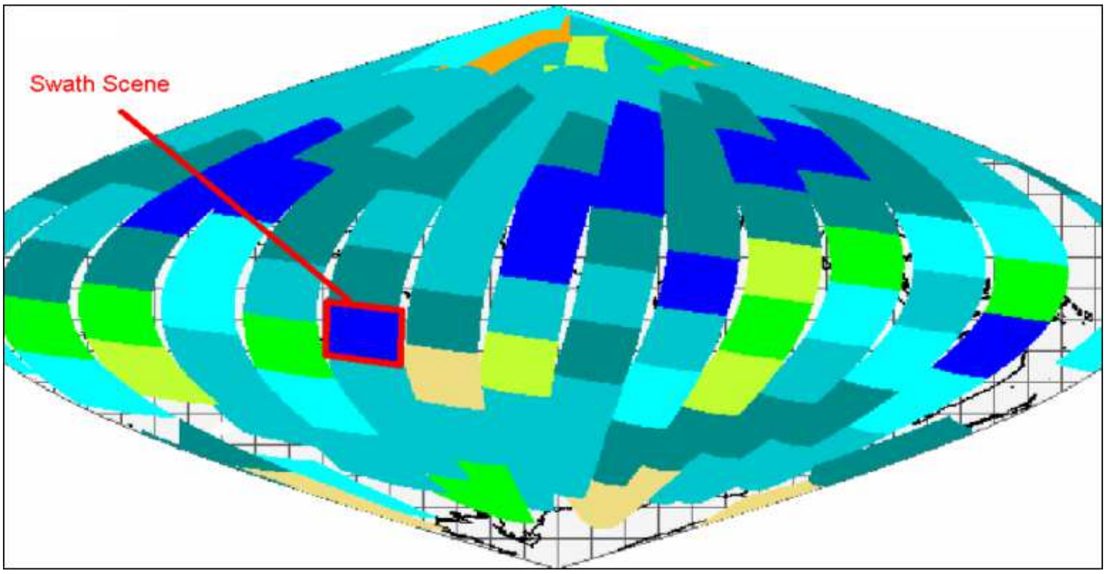

FIGURE 2.4: SWATH SCENES: A SWATH SCENE IS A 5 MINUTES ACQUISITION IMAGE

MODIS products can be provided in two kinds of grid: the sinusoidal and the Climate Modeling grid (CMG).

The sinusoidal projection is a pseudo-cylindrical equal area projection, with no distortion on the central meridian. Products mapped into the Sinusoidal projection are then gridded into 10-degree units called tiles (see figure 2.5). These files are geometrically corrected and are easily read by image processing software and manipulated by an end user. Products in the Climate Modeling Grid (CMG) provide information on the whole earth surface with a spatial resolution of 0,05 degrees , which correspond to about 5,6 km (see figure 2.6).

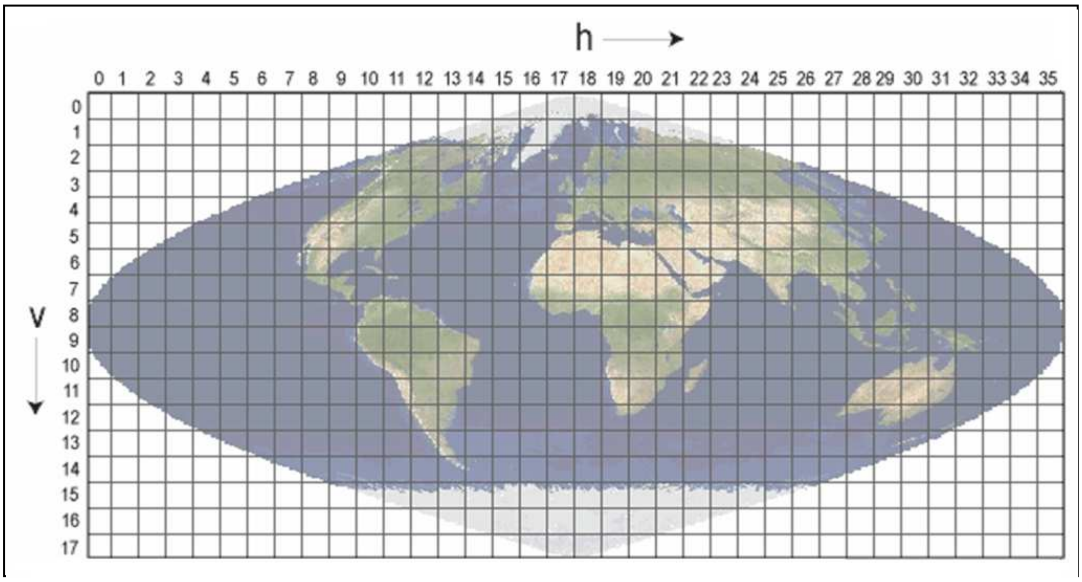

FIGURE 2.5: MODIS SINUSOIDAL GRID. TILES ARE 10 DEGREE BY 10 DEGREE AT THE EQUATOR. THE TILE COORDINATE SYSTEM STARTS AT (0, 0) (HORIZONTAL TILE NUMBER, VERTICAL TILE NUMBER) IN THE UPPER LEFT CORNER AND PROCEEDS RIGHT (HORIZONTAL) AND DOWNWARD (VERTICAL). THE TILE IN THE BOTTOM RIGHT CORNER IS (35, 17).

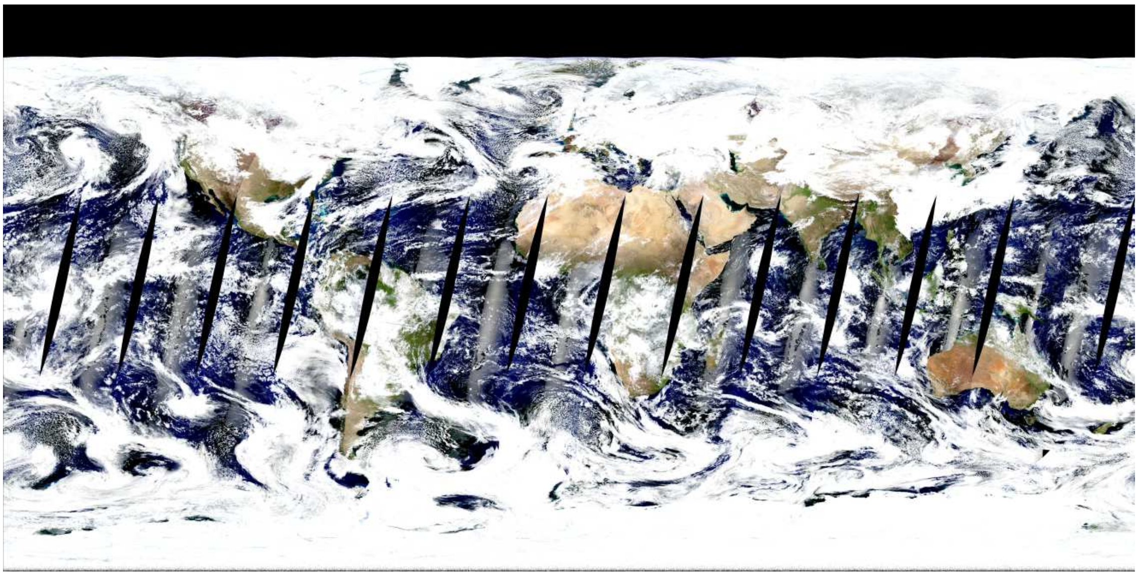

FIGURE 2.6 MODIS CLIMATIC GRID

For purposes of data reduction and removal (or minimization) of bad data it is useful to take the best observations over a series of days and produce one output. This process is called compositing and has been done for years with a variety of different satellite data products. For the MODIS instrument the repeat cycle of nadir overpasses is 16 days. This means that every 16 days the instrument will be travelling on nearly the exact same path. For this reason standard composite periods for MODIS are multiples of 8 days, exactly the mid-point of the repeat cycle. Thus we see time steps in composited products of 8, 16 and 32 days of data.

2.1.5 PRODUCT FORMAT

MODIS products are generally archived in Hierarchical Data Format - Earth Observing System (HDF-EOS) format files. HDF, developed by the National Center for Supercomputing Applications (NCSA), is the standard archive format for EOS Data Information System (EOSDIS) products.

HDF files are self-describing. The term "self-description" means that, for each HDF data structure in a file, there is comprehensive information about the data and its location in the file. This information is often referred to as metadata. So the product files contain global attributes (metadata) and scientific data sets (SDSs) (data arrays) with local attributes. Scientific Data Sets are digital arrays, which can contain radiance/reflectance values, geolocation information (latitude/longitude) and many other data depending on the product.

Unique in HDF-EOS data files is the use of HDF features to create point, swath, and grid structures to support geolocation of data. These structures (Vgroups and Vdata) provide geolocation relationships between data in a SDS and geographic coordinates (latitude and longitude or map projections) to support mapping the data. Attributes (metadata), global and local, provide various information about the data (see figure 2.7).

HDF was designed to address many requirements for storing scientific data, including:

- support for the types of data and metadata commonly used by scientists;
- efficient storage of and access to large data sets (Scientific Data Sets or SDS);
- platform independence; extensibility for future enhancements and compatibility with other standard formats.

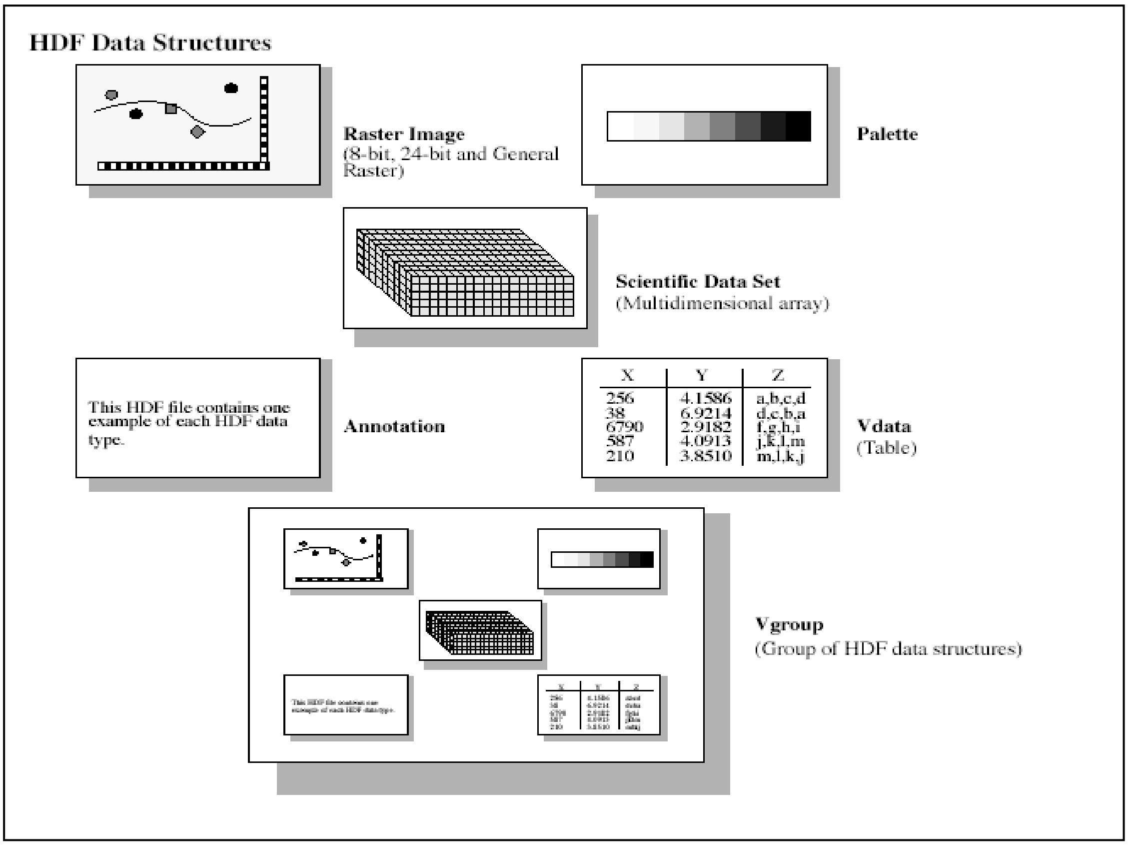

FIGURE 2.7: HDF DATA STRUCTURE (SEE HDF USER GUIDE)

In the work of thesis another kind of MODIS dataset has been used, the MODIS Rapid Response System. This 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, with a spatial resolution up to 250m, and georeferenced in a geographic reference system referred to WGS84 datum. Furthermore a simplified correction from the effects of the atmosphere (not considering the aerosol component of the atmosphere) is implemented on them before their distribution. They are distributed in a JPEG or GEOTIFF format, further characteristics of these data are provided in chapter 4.

2.1.6 PRODUCT AVAILABILITY

All MODIS products are available, completely free of charge, in the web. The most complete archive is WIST (Warehouse Inventory Search Tool) portal, a web-based search-and-order tool, provides a way for users to search for Earth Science data from multiple participating archives (https://WIST.echo.nasa.gov/api/). Then for each type of product different sites can be available, for example the GSFC DAAC (Distributed Active Archive Centers) (http://daac.gsfc.nasa.gov/) for Atmosphere products, the land processes DAAC (http://lpdaac.usgs.gov /) for land product, the National Snow and Ice Data Center (NSIDC) DAAC (http://nsidc.org/daac/) for cryosphere, climate and snow-ice products.

2.2 THE NOAA SYSTEM

The "Advanced Very High Resolution Radiometer (AVHRR)", is a sensor operating onboard of the National Oceanic and Atmospheric Administration (NOAA) POES series (Polar-Orbiting Operational Environmental Satellites). Its aim is to provide radiance data for investigation of clouds, land-water boundaries, snow and ice extent, ice or snow melt inception, day and night cloud distribution, temperatures of radiating surfaces, and sea surface temperature. The AVHRR data collection effort also provides opportunities for studying and monitoring vegetation conditions in ecosystems, including forests, tundra, and grasslands with applications that include agricultural assessment, land cover mapping, production of large-area image maps (e.g., country maps, continental maps, world maps), and evaluation of regional and continental snow cover.

2.2.1 THE NOAA-POES SERIES MISSIONS

The NOAA-POES series are regarded as the backbone of the US meteorological program. The current POES series satellites are named simply NOAA-9 through NOAA-17 in order of launch. The program has evolved over several years starting in 1960 with the first satellite TIROS. The philosophy of NOAA is to maintain at least two operational satellites in complementary orbit. The POES satellites are supposed to operate till 2010 (see table 2.4).

All NOAA-POES satellites have a circular, sun-synchronous polar orbit with a nominal flight height of 833 km. The even numbered satellites cross the equator at local solar times of approximately 7:30 and 19:30, while the odd-numbered satellites cross the equator at local solar times of approximately 2:30 and 14:30.

TABLE 2.4 NOAA POES LAUNCH DATES

Name after launch	Date of launch	AVHRR version	Mode of operation
TIROS-N	October, 1978		
NOAA-6	June 27, 1979		end of mission: November 11, 1986
NOAA-7	June 23, 1981	2	end of mission: June 7, 1986
NOAA-8	March 28, 1983	1	end of mission: October 31, 1985
NOAA-9	December 12, 1984	2	end of mission: November 5, 1994
NOAA-10	September 17, 1986	1	operational
NOAA-11	September 24, 1988	2	end of mission: September 13, 1994
NOAA-12	May 15, 1991	1	in operation; no generation of value-added products by DFD
NOAA-13	August, 1993		failure after launch
NOAA-14	December 30, 1994	2	degraded; generation of value-added products was stopped on October 3, 2001 by DFD
NOAA-15	May 13, 1998	3	in operation; no generation of value-added products by DFD
NOAA-16	September 21, 2000	3	in operation; no generation of value-added products by DFD
NOAA-17	June 2002	3	in operation; no generation of value-added products by DFD
NOAA-18	May 20, 2005	3	in operation; generation of value-added products (NDVI, LST and SST) by DFD ongoing

2.2.2 THE AVHRR INSTRUMENT

The AVHRR instrument is a cross-track scanning system which features four (AVHRR/1), five (AVHRR/2) or six (AVHRR/3) spectral channels. The spectral bandwidth (μm) and the "Instantaneous Field of View (IFoV)" of the AVHRR are listed in the table 2.5

The IFoV of each channel is approximately 1.4 milliradians leading to a resolution at the satellite subpoint of 1.1 km for a nominal altitude of 833 km. The scanning rate of the AVHRR is 360 scans per minute.

The analog data output from the sensors is digitized on board the satellite with 10 bit resolution at a rate of 39,936 samples per second per channel. Each sample step corresponds to an angle of scanner rotation of 0.95 milliradians. At this sampling rate, there are 1.362 samples per IFoV. A total of 2048 samples will be obtained per channel per Earth scan, which will span an angle of 55.4 degrees from the nadir (subpoint view).

TABLE 2.5 SPECTRAL CHARACTERISTICS OF AVHRR SENSORS

Channel	TIROS-N	NOAA-6,8,10	NOAA-7,9,11,12,14	NOAA-15,16,17,18	IFoV (mrad)
1	0.550- 0.90 µm	0.580- 0.68 µm	0.580- 0.68 µm	0.58 - 0.68 µm	1.39
2	0.725- 1.10 µm	0.725- 1.10 µm	0.725- 1.10 µm	0.725-1.0 µm	1.41
3A				1.58-1.64 µm	1.3
3B	3.550- 3.93 µm	3.550- 3.93 µm	3.550- 3.93 µm	3.55-3.93 µm	1.51
4	10.500-11.50 µm	10.500-11.50 µm	10.300-11.30 µm	10.3-11.3 µm	1.41
5	Ch4 rep.	Ch4 rep.	11.500-12.50 µm	11.5-12.5 µm	1.30

The IR channels are calibrated in-flight using a view of a stable blackbody and space as references. No in-flight calibration of the visible channels is performed (although the space view is available as one reference point). The calibration of the visible channels 1 and 2 was done using pre-launch calibration coefficients for NOAA-10 and -12 and time-adjusted values for the NOAA-7, NOAA-9 and NOAA-11 AVHRR. For the NOAA-14 AVHRR, pre-flight values were used between January 19, 1995 and November 21, 1995. On July 31, 1995, NOAA NESDIS(National Environmental Satellite, Data, and Information Service) published new calibration post-launch coefficients which were applied for all products between November 22, 1995 and November 11, 1996.(NOAA Polar Orbiter Data User's Guide).

Since November 12, 1996, the calibration coefficients have been updated once a month by NOAA NESDIS and these updates are implemented at German Remote Sensing Data Center (DFD)'s production chain after availability.

Currently, the AVHRR's onboard NOAA-12, NOAA-15, NOAA-17 and NOAA-18 are in operational mode. AVHRR on board of NOAA-16 severely degraded due to scan motor anomaly. NOAA-13 didn't become operational due to a power failure after launch.

2.2.3 AVHRR DATA DISTRIBUTION

NOAA environmental data are distributed by different providers, such as the electronic library CLASS (Comprehensive Large Array-data Stewardship System). CLASS is NOAA's premier on-line facility for the distribution of NOAA and US Department of Defense (DoD) Polar-orbiting Operational Environmental Satellite (POES) data, NOAA's Geostationary Operational Environmental Satellite (GOES) data, and derived data.

There are three data types produced from the POES AVHRR:

- The Global Area Coverage (GAC) data set is reduced resolution image data that is processed onboard the satellite taking only one line out of every three and averaging every four of five adjacent samples along the scan line. The spatial resolution of this product is 4.4-km pixel, which allows daily global coverage to be systematically stored and played back to NOAA ground stations;
- the Local Area Coverage (LAC) data set is recorded onboard at original resolution (1.1 km) for part of an orbit and later transmitted to earth
- the High Resolution Picture Transmission (HRPT) are full-resolution (1-km) data received directly in real-time by ground stations.

Furthermore NOAA presently produces a number of operational products from AVHRR imagery. Calculated Normalized Difference Vegetation Index (NDVI), Sea Surface Temperature, atmospheric aerosols, and sea ice cover data are available globally, while snow cover is mapped for the northern hemisphere.

2.3 SPOT VEGETATION

The SPOT program consists of a series of optical remote sensing satellites with the primary mission of obtaining Earth imagery for land use, agriculture, forestry, geology, cartography, regional planning, water resources and GIS applications (see table 2.6).

TABLE 2.6 SPOT MISSIONS

Satellite name	On board instrument	Date of launch	End of mission
SPOT 1	HRV, HRVIR	22 February 1986	28 November 2003
SPOT 2	HRV, HRVIR	22 January 1990	still operational
SPOT 3	HRV, HRVIR	26 September 1993	incident occurred on 14 November 1996. After 3 years in orbit the satellite has stopped functioning
SPOT 4	HRV, HRVIR, VEGETATION	24 Mar 1998	still operational
SPOT 5	HRS VEGETATION-2	4 May 2002	still operational

Currently, three SPOT satellites (SPOT 2, 4, 5) are operational. Up SPOT4, SPOT satellites have two viewing instruments (HRV, HRVIR); when operating simultaneously in the vertical viewing configuration, they are capable of imaging a 117 km wide strip on the Earth's surface. The instruments are fitted with programmable Strip Selection Mirrors giving access, off track, to areas within a 950 km wide corridor. This specific oblique viewing capability greatly increases the frequency at which the satellite can revisit a particular site (4 to 11 times within the 26 day cycle, according to latitude).

Consequently stereopairs, used for relief perception and elevation plotting (Digital Elevation Modelling) are formed from two SPOT images acquired at different viewing angles on successive satellite passes.

A single SPOT scene covers a geographical area of 60 x 60 km. Two alternative modes of imaging are possible:

- Panchromatic, black and white, with a ground resolution of 10 m;
- Multispectral, with 20 m ground resolution acquired simultaneously in 3 bands: green, red and near infrared (HRV) and medium infrared (HRVIR).

SPOT5 satellite is composed of two new HRG viewing instruments derived from SPOT4 HRVIR instruments, which have a better resolution: 2.5 to 5 meters in panchromatic mode and 10 meters in multispectral mode. SPOT5 has a new HRS instrument operating in panchromatic mode that takes images in front and behind the satellite at the same time which allows stereoscopy. SPOT 5 is also carrying the recurrent VEGETATION instrument. The VEGETATION Programme is conceived to allow daily monitoring of terrestrial vegetation cover through remote sensing, at regional to global scales. The instrument and associated ground services for data archival, processing and distribution are operational since April 1998. The first VEGETATION instrument is part of the SPOT 4 satellite and a second payload, VEGETATION 2, is now operational onboard SPOT 5.

The overall objectives of the "VEGETATION" system are to provide accurate measurements of basic characteristics of vegetation canopies on an operational basis, either for scientific studies involving both regional and global scales experiments over long time periods (for example development of models of the biosphere dynamics interacting with climate models), or for systems designed to monitor important vegetation resources, like crops, pastures and forests.

The "VEGETATION" system, consisting of a satellite-borne sensor and of its associated ground segment, provides long term basic measurements adapted to biosphere studies. Opportunities for scale integration are provided by the combination with the main SPOT instruments (HRVIR) which allow high spatial resolution for detailed modeling activities or multilevel sampling procedures. Availability of data to different types of users is facilitated through the centralisation of reception and archiving global data sets.

In this work of thesis free data derived from the VEGETATION sensor are analyzed.

2.3.1 THE SPOT VEGETATION INSTRUMENT

The VEGETATION instrument is an imaging system in 4 spectral bands: blue (0.43-0.47 microns), red (0.61-0.68 microns), near infrared (0.78-0.89 microns), and SW infrared (1.58-1.75 microns) (see table 2.7).

The red and near infrared are particularly well adapted to describe the vegetation photosynthesis activity, while the SW infrared is a good detector for the ground and vegetation humidity. The blue is designed in this case to make atmospheric corrections.

VEGETATION uses telecentric optics giving a quasi constant spatial resolution through the field of view, particularly wide (2 200 km on the ground) : this resolution is 1.15 km at nadir, and still 1.7 km on the sides of the field of view (101°).

The VEGETATION cameras cover a very wide field of view of 101° producing a swath width of 2250 km. the nominal resolution for optimizing the instrument mission is defined by pixels of 1.165 x 1.165 km instead of the ten metres for the HRVIRS.

TABLE 2.7 SPOT VEGETATION SENSORS SPECTRAL AND SPATIAL CHARACTERISTICS

VEGETATION features				
	Wavelenght range [µm]		Pixel resolution [m]	Swath [km]
B0	0.430	0.470	1165	2250
B1	0.610	0.680	1165	2250
B2	0.780	0.890	1165	2250
SWIR	1.580	1.750	1165	2250

For each of the main missions, some specific parameters are important and have to be derived from remote sensing data. To keep the measurements as robust as possible, only wide spectral band measurements (50 nm) are considered and the objectives are to characterize the main features of plant canopies: absorption by chlorophyll, water contents and structural properties. The best and minimal set of spectral bands known to fulfill this need is composed of:

- a red band centred on the absorption peak of the chlorophyll (0.665 µm),
- a near infrared band corresponding to the maximum vegetation spectral reflectance and principally related to the structural properties of the canopies and to percentage of soil covered by vegetation,
- a short wave infrared band centred around 1.65 µm where reflectance is related to water content of the canopy components and to its structure.

Considerations for atmospheric effects characterization or correction are added: among different possibilities that are under validation, both the use of adapted vegetation indices computed from red and near infrared reflectance and direct or indirect use of additional spectral bands in the blue region are retained. An additional band is provided in order to compute or characterise the atmospheric state (aerosols):

- a blue band (between 0.45 and 0.50 µm) where ground reflectance of vegetation cover is minimal and atmospheric aerosol diffusion effects are maximal.

The influence of atmospheric water vapor, which is most important in a wide near infrared band, is severely decreased by limiting the upper portion of the near infrared band to avoid the 0.935 µm water vapor absorption band.

The spectral response is as “similar” as possible to the high resolution instrument bands, at least for spectral bands that participate to the same mission: the red, near infrared and short wave infrared bands.

2.3.2 SPOT VEGETATION DATA DISTRIBUTION

All vegetation data are processed and distributed by the image processing and archiving centre (VITO), which is located in Belgium. Data older than three months are freely distributed to the scientific community (see http://www.vgt.vito.be/)

CHAPTER 3: OBJECTIVES OF THE THESIS AND METHODOLOGY

The work of thesis has been developed in the field of remote sensing to support the activities of humanitarian agencies for disaster management sector. In particular the work has been developed in cooperation with ITHACA (Information Technology for Humanitarian Assistance Cooperation and Action), a project developed by WFP in partnership with the Politecnico di Torino and the Higher Institute on Innovation Territorial Systems (SiTI). This joint project is establishing historical data patterns, developing forecasting models and applying satellite imagery to rapidly determine and monitor where hurricanes, floods, earthquake or famine may occur. This type of tailored information, if delivered in a timely fashion, can be an invaluable asset in planning and prioritizing locations for needs assessments, not to mention providing huge cost and time savings.

The project described in this thesis provides a transversal support to different Ithaca activities, which, in this chapter are described before the discussion of the objectives of the study and the methodologies.

3.1 WORK CONTEST: ITHACA

In its first years of life, ITHACA carried out many activities, which mainly deal with thematic maps production and related products and services supply, necessary to correctly plan and manage natural disaster mitigation operations. In particular, these products are used in a fruitful way for Early Impact and Early Warning Activities. Early Impact activities are devoted to the identification of effects caused by natural and manmade disaster events. The results of these analysis are suitable for the correct planning and support of necessary relief operations. On the other hand, the main aim of Early Warning activities and analyses is to foresee catastrophic events and their effects on population and territory.

At the moment, in Ithaca, several activities are related to WFP's request, in order to:

- develop tools to improve the collection and the distribution of emergency preparedness and response products generated in the different Ithaca products (Web Application for Geo-Data Sharing);
- develop technologies and systems for Early Warning and Early Impact analyses for emergencies arising from natural hazards (Early warning for floods and Detection and Monitoring of Drought Events);
- support Information Management in the realm of emergency preparedness and response;
- develop procedures and services to provide decision-makers engaged in humanitarian aid activities with real-time information about roads practicability (Snow Cover Service) and data generated by the Early Warning and/or the Early Impact developed systems ;
- support the implementation of the WFP Spatial Data Infrastructure (UN Spatial Data Infrastructure SDI);
- support the use of new methodologies by developing suitable tools;
- enable co-operation with other research centres and private entities.

Afterwards, in order to constitute a service support team for the WFP, ITHACA:

- supports the maintenance of web based tools and systems;
- provides remote sensing analysis;
- provides GIS support pre- and post- natural disaster impact;
- manages and maintains global spatial data, making them accessible.

3.1.1 UN SPATIAL DATA INFRASTRUCTURE

Accurate, easily accessible geographic information is crucial to good decision-making in humanitarian operations. The main aim of the Spatial Data Infrastructure (SDI) project is the development and implementation of a global database and of the rules to discover, use and exchange geospatial data for WFP and other actors in the humanitarian sector. The SDI, currently in production, is an efficient tool for storing, querying and manipulating large amounts of global geographic information and spatial data (see Figure 3.1). It is suitable for analysis, visualization purposes, and for sharing it between a large and differentiated community.

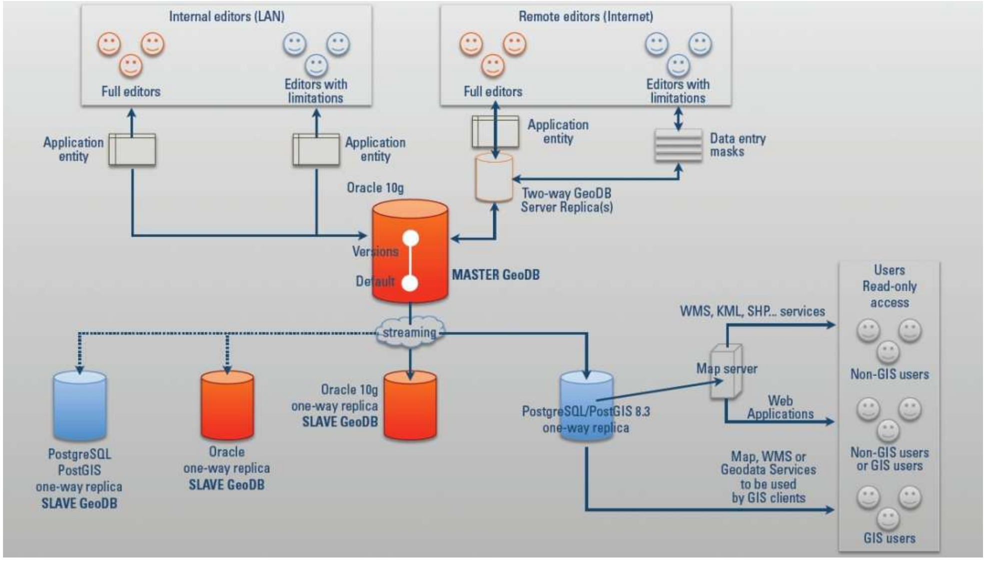

FIGURE 3.1. ITHACA SDI ARCHITECTURE

The project development process has three components:

- Data infrastructure: the complete SDI platform includes catalogues, search engines, data processing tools, map generating tools, user identification and security measures. The database structure has been chosen according to interoperability standards. The data model is conceived according to specific needs and demands of the user groups within the WFP organization. The set of geographic and alphanumerical data is structured in a geo-database, which is consistent, updated, standardized and easily accessible.
- Data management: WFP SDI integrates and maintains data coming from multiple sources, providing each user with the latest datasets, thus avoiding duplications and errors. Roles and privileges are defined and maintained, to control and guide the maintenance process;
- Services: data are accessible through user-friendly web-based platforms, and they will be integrated into the automated dynamic web-based maps for operational support (Logistics Capacity Assessments, Contingency Plans, etc.). The web platforms enable complete software customization based on user demands.

SDI will be available to WFP GIS practitioners in-house and, possibly, to partners. It also provides a significant source of information for the Early Impact Analysis service, which is a parallel project undertaken by ITHACA.

The European Union is also working toward the implementation of a European SDI. The INSPIRE directive entered into force in 2007 with the objective of sharing experiences and exchanging standards. Contacts with INSPIRE Thematic Working Group on Transport Networks (TWG TN) have been activated.

3.1.2 EARLY WARNING FOR FLOODS

The aim of the project is the development of a simplified and efficient Early Warning System for flood events, at worldwide extent. This system is able to give an alert in advance about the occurrence of floods around the world monitoring the heavy rainfalls in near real time. It can be used by WFP (World Food Programme) or other humanitarian assistance organizations to evaluate the events and to understand the potentially floodable areas in places where their assistance is needed.

The expected outputs consist of thematic maps that show the areas in which heavy rainfalls are occurring. River basins are shown in different colors indicating the level of criticality, according to the severity of rainfalls (see Figure 3.2). The maps are delivered through a dynamic website powered by a geonetwork node.

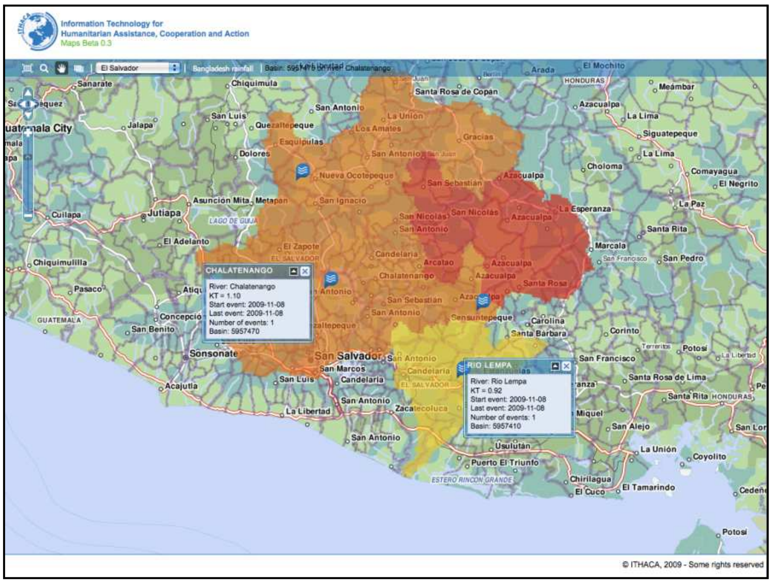

FIGURE 3.2 EXAMPLE OF THE OUTPUT MAP: THE ALERTED BASINS ARE SHOWN WITH 3 DIFFERENT POSSIBLE COLOURS, ACCORDING TO THE SEVERITY OF RAINS.

The method is based on precipitation analysis and it uses rainfall data from satellite at worldwide extent. This system works on river basins scale with global coverage by using satellite rainfall data from Tropical Rainfall Measuring Mission (TRMM) Multisatellite

Precipitation Analysis (TMPA). Oracle 10g © has been used to load and elaborate the whole dataset of rainfall that was structured in a data warehouse architecture.

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 (DDF) curves. The final output of this procedure is the creation of a complete historical database of heavy rainfalls with global coverage and single river basin resolution.

The same analysis is performed using near real-time data to monitor current rainfall conditions and to make alerts in near real time: it can be considered the proper "Early Warning System". The real time rainfall data (the 3-hourly real time 3B42RT product) are compared with the DDF curves, previously determined using historical data.

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.

Once they are triggered, alerts are automatically mapped and the overall situation can be visualized on the web application. The web interface, which was developed by extending OpenLayers functionalities, aims to offer an easy access to the data and also to map and alert the basins that are potentially under emergency. A tool to zoom at basin level and that can activate informative layers extracted from ITHACA GSDI (Global Spatial Data Infrastructure) is offered to the users.

Currently, the research team is working on testing the method on various scenarios, paying attention to the effective consequences of foregoing flood.

3.1.3 EARLY IMPACT ANALYSIS

The main aim of the Early Impact activities is to rapidly produce geo-referenced information on the impact of disasters, especially data on affected areas and population. The rapid mapping activities aimed at supporting the first stage of disaster management are generally based on satellite remote sensing data. Different types of satellite data can be used, mainly according to the type of disaster and the approximate extent of the affected areas:

- Low/medium resolution multispectral optical imagery (e.g.: MODIS, ALOS AVNIR, DMC, Landsat),
- High resolution optical data (e.g.: SPOT, Formosat, Ikonos, WorldView-1 and 2; Quickbird, GeoEye),
- Medium resolution radar data e.g.: Envisat, Radarsat, ALOS Palsar),

- High resolution radar data (e.g.: Cosmo-SkyMed, TerraSAR-X).

The aforementioned data belongs to two main families: public-domain data (generally accessible through the web (e.g.: MODIS, Landsat); or commercial imagery that should be purchased through reseller companies (e.g.: Eurimage, SPOTimage, DigitalGlobe). Sometimes it is possible to freely access satellite data on the basis of scientific agreements between Space Agencies and Principal Investigators in charge of a specific research (e.g.: ESA Category-1 users).

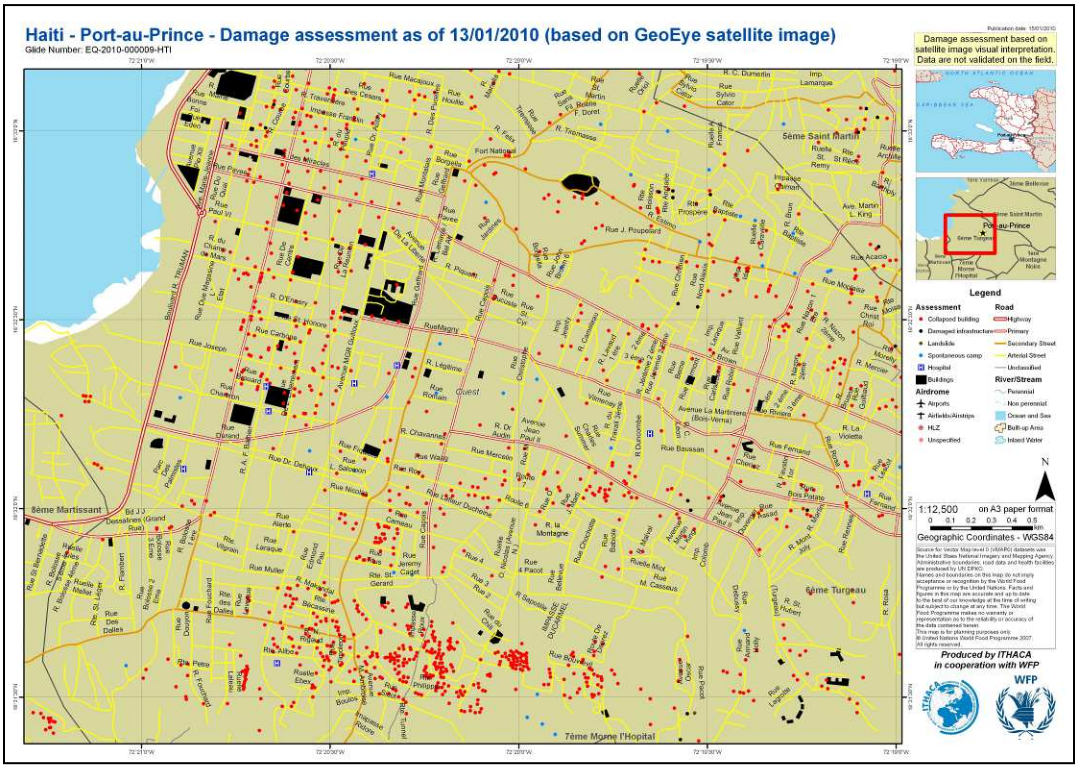

FIGURE 3.3 EXAMPLE OF CARTOGRAPHIC REPRESENTATION OF POST-EARTHQUACKE DAMAGE ASSESSMENT

Very high resolution optical data are generally analysed to identify damages to buildings and infrastructures (road accessibility, collapsed bridges, etc...), landslides or temporary shelters. Unfortunately this kind of data is affected by the cloud coverage, which is often persistent during flood or cyclone events. The all-weather capability of the radar technology and the possibility to acquire data also during night time are crucial advantages of a radar based approach for the monitoring of the latter events.

The number and distribution of potentially affected populations is another type of important information requested by managers responsible for the distribution of humanitarian help. A rapid estimate of this information can be obtained by overlaying potentially affected areas with population distribution data.

The main outputs of an early-impact analysis are normally in the form of cartographic products covering the area of interest (see an example in Figure 3.3). Data should be represented following cartographic rules, both in relation to output map scale and according to representation rules.

The dissemination process has the same importance as the analysis and map production phases. Final products should reach decision makers at the right time and in the right place. Integration of map products with adequate metadata in a standard format is of great help during discovery and exploitation phases.

3.1.4 UNMANNED AERIAL VEHICLES

The aim of this project is the development and construction of remotely controlled mini aircrafts (UAVs or "drones") for the acquisition of visual data necessary to plan emergency and relief food aid interventions. Each drone is equipped with photographic and multispectral sensors suitable for digital photogrammetric shootings. The aerial platform is capable of autonomous flight except for take-off and landing.

The UAVs are easily transportable and user-friendly. They produce images (see Figure 3.4) which will allow rapid updates of existing maps, and can produce thematic maps that highlight damages to buildings, infrastructure and flooded areas. In addition a micro-camera can acquire sequences to document the investigated area that can be downlinked in real-time.

The UAV is developed based on a Patent Pending model realized by the Aerospace Department of the Politecnico di Torino. It can be either electrically-powered or equipped with an internal combustion engine. The fixed-wing configuration gives the vehicle a better capability of withstanding adverse weather conditions, and guarantees a superior flight performance (in terms of endurance and range). Besides carrying out photogrammetric flights with perfectly respected flight plans, the presence of the autopilot and GPS allows the UAV to carry out missions at half of the optimum range limits, also beyond visual range of the Ground Control Station.

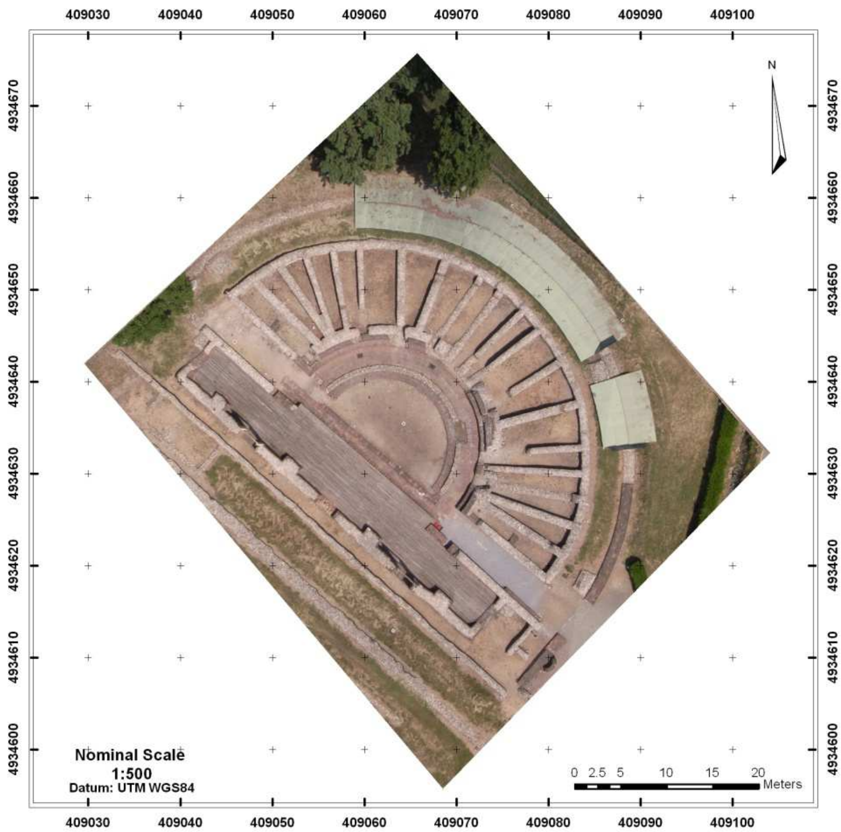

FIGURE 3.4. EXAMPLE OF MOSAICKED ORTHOIMAGES ACQUIRED BY THE UAV

Currently, the research team is working on:

- The assessment of the autopilot software parameters to improve the performance of autonomous photogrammetric flights;
- Testing other types of sensors (IR, radar, etc) whose weight and dimensions are sufficiently miniaturized to be compatible with the characteristics of the UAV;
- Developing software tools for the automatic processing and mosaicking of the acquired images

3.1.5 SNOW COVERAGE MONITORING

The aim of this project is to provide daily updated information on snow covered areas as support to WFP Logistics. The information provided enables the Road Accessibility Analysis which is required to ensure route optimization in humanitarian aids delivery. This is achieved through the implementation of a cost-free automated process based on MODIS (Moderate Resolution Imaging Spectroradiometer) satellite data.

The expected output consists of a daily updated web-GIS (see Figure 3.5) for the monitoring of particular areas of interest and a service on request for other areas (the expected map scale is 1:1000000). The web-GIS displays a synthesis of MODIS Snow covered satellite data. The service on request can provide snow vector data in near in real-time. The files refer to snow coverage information generally updated 24 hours before the request. A metadata file containing the reference time-period of the information is provided too.

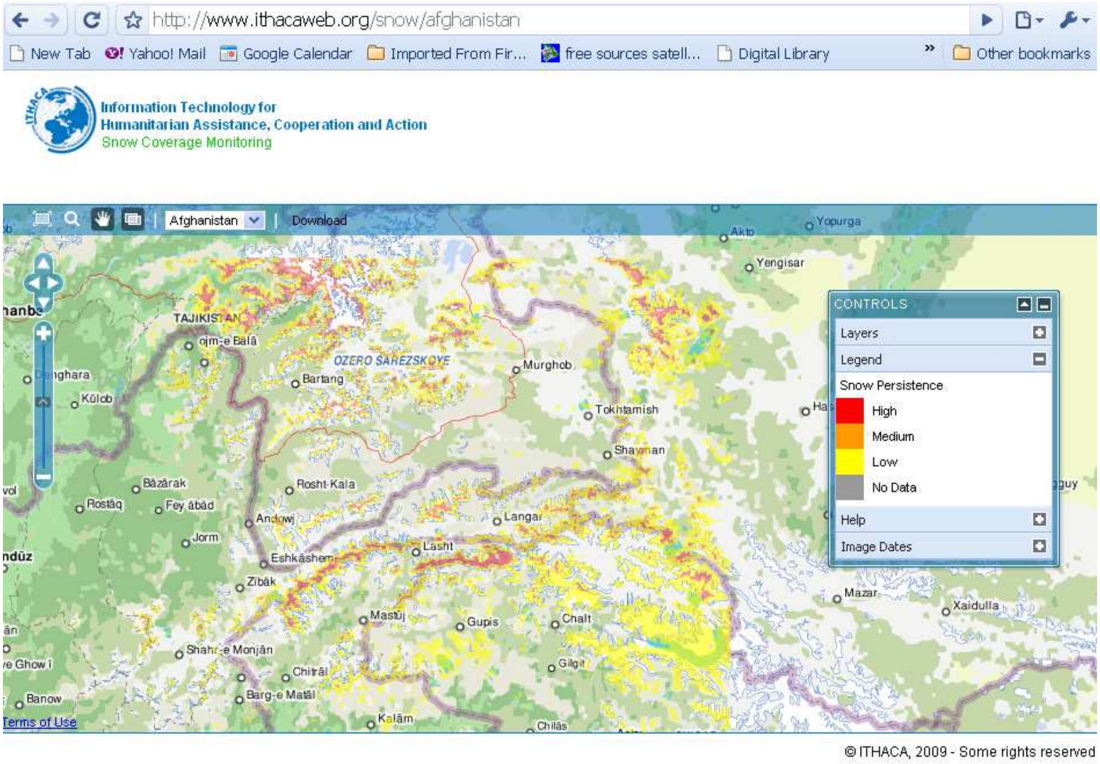

FIGURE 3.5. ITHACA SNOW COVER WEB GIS

The method consists mainly of MODIS data processing. This kind of satellite data is free of charge and produces a daily overview of the earth surface with the purpose of extracting snow coverage data. In order to produce the data, it is necessary to implement a mechanism which automatically performs the following operations:

- Images selection from archive, depending on required date and area of interest;
- Geometric correction, geo-referencing and resizing;
- Elaboration of the synthesis (monitoring service on the areas of interest);
- Elaboration of snow coverage vector information;
- Web-GIS updating.

3.1.6 WEB APPLICATION FOR GEO-DATA SHARING

The main aim of the project is to distribute and share georeferenced information for both early warning and early impact activities, by means of web-GIS application, based both on commercial or Open Source (OS) platform. Web based applications for georeferenced data distribution are highly customized and implemented using different software architectures, granting data reliability and integrity.

The project is aimed to develop new architectures for Web-GIS creation in response to emergencies. Each application needs a server to run on, a software architecture and customization scripts. Ithaca develops applications that thanks to the Java and Python technology can be hosted on any operating system.

The web GIS applications (see an example in Figure 3.6) can be accessed by means of a web browser without any additional plug-in. Published data may be stored on the same server the application runs on or may be gathered from the net using OGC standards such as WMS, WFS and WCS. The web service can also provide some analysis tool for querying, editing and downloading data through the web-GIS map.

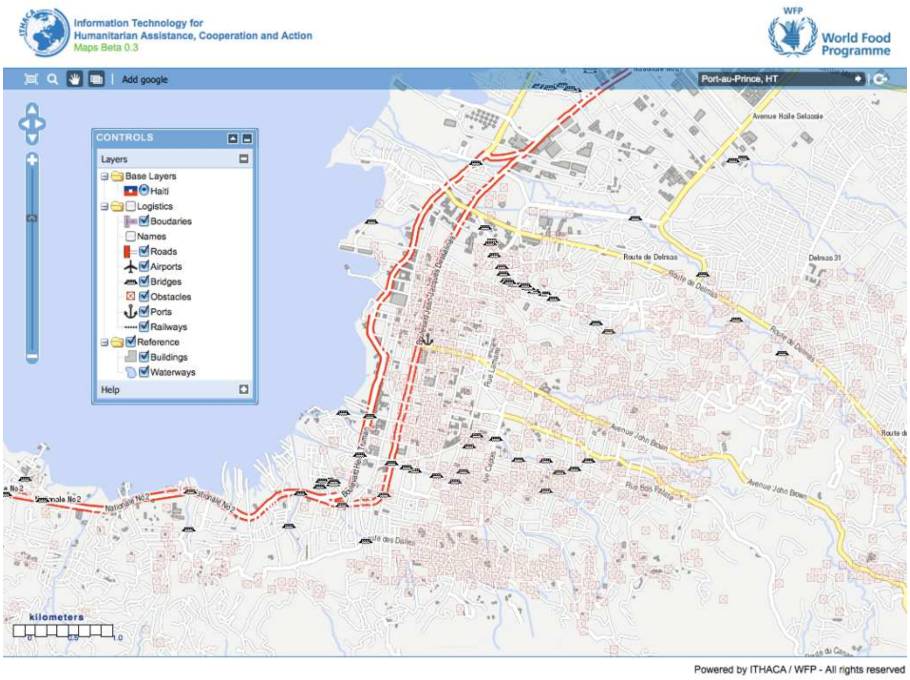

FIGURE 3.6. HAITI EMERGENCY WEB APPLICATION

The research team is working on a system powered by the Python programming language and the Django web framework while Geoserver is used as map engine with javascript libraries such as OpenLayers and GeoExt for the client side.

Working solutions are available on-line for the diffusion of Snowcover project data outputs and for the Flood Early Warning System. Implemented features are: Data navigation (Continuous pan and zoom), Layer transparency control, Layer visualization control, Data integration (different sources, including Google Maps), Data analysis (raster and vector data analysis), GeoRSS integration and geographic search.

New feature under development and testing is the real time data analysis through GRASS GIS.

3.1.7 DETECTION AND MONITORING OF DROUGHT EVENTS

The aim of this project is the development of global drought detection and monitoring system as a support to WFP activities. Drought is a water-related natural hazard with some peculiar characteristics: it has a slow onset, can affect extensive regions, and last even for many years, with serious impacts on population, first of all reducing food production.

The method is based on the analysis of a series of drought-related variables and indices, such as NDVI and SPI, obtained mostly from satellite data, in order to define thresholds and triggers suitable for early warnings. Land cover, land use, soil moisture, soil type and other relevant information may be integrated in the system to improve its effectiveness.

The NDVI (Normalized Difference Vegetation Index) is a satellite-based vegetation index, and its monitoring over time allows detecting water stress vegetation conditions. The spatial analysis of monthly historical time series of NOAA AVHRR NDVI data (1982-2007) allowed to identify long-term vegetation dynamics and to produce maps about the areas that were subject to increase or reduction in vegetation greenness. Moreover, the near real-time monitoring of NDVI data allows to detect deviations from identified trends, showing anomalous vegetation conditions, suitable for drought detection purposes (see an example of Vegetation Condition Map in Figure 3.7).

The SPI (Standardized Precipitation Index) is a meteorological drought index, which requires only precipitation input data. The SPI from one side gives a numerical value which offers quantitative information related to the deviation from normal conditions, which can be interpreted as the intensity of a drought spell in case of negative values; from the other side allows considering for every month different time scales, related to different drought conditions.

A Web-GIS application will be implemented, devoted to the rapid distribution of the results (maps and graphs) obtained during the different monitoring phases. Moreover, this service will allow users to search, explore and analyze all base data archived for the project, in order to visualize maps and graphs for customized reports generation.

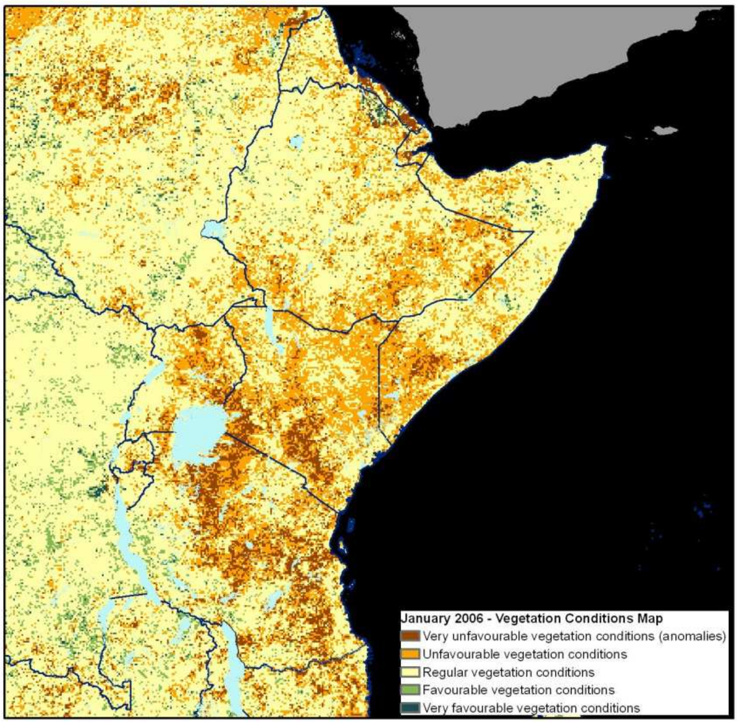

FIGURE 3.7. JANUARY 2006, HORN OF AFRICA, VEGETATION CONDITIONS MAP: SPATIAL DISTRIBUTION OF THE DETECTED NDVI ANOMALIES AND DEVIATIONS

3.2 OBJECTIVES OF THE THESIS

The work of thesis aimed at producing base input data for WFP officers and Ithaca activities by developing automatic satellite data-processing applications. In particular, the attention was focused on flood events and the effects that such events could have on the ground. Floods are one of the most widespread dangerous natural phenomena on Earth, which have large destructive force. During floods authorities must have timely and overall information about flooded territories in order to provide a prompt intervention on the area affected by the event. In developing countries this mission is particularly difficult because of lack of hydrological stations on the rivers. In addition for large areas, such as major river valleys, time and funds available are often limited. Moreover, humanitarian agencies, whenever possible, have to be prepared to the event also in the pre-event phase, in order to forecast possible floods and their impact on the environment and on population.

The applications developed during the work of thesis, in collaboration with Ithaca, aimed at providing information both in the pre and in post event phases. In particular, automatic procedures have been developed for processing huge amounts of data in order to provide an historical archive of flooded areas and phenological maps for the study area. This archive, integrated in a suitable data sharing system prepared by Ithaca, will mainly allow the WFP

Local Offices to analyze the gravity of past flood events, in terms of both the extension of flooded areas and the effects of these events brought on vegetation.

Furthermore, the produced archive will provide also a base data repository proper for several Ithaca early warning and early impact activities. In that case, the developed procedures can also be used in order to regularly update the archive.

In the work of thesis, applications based on free satellite data have been developed and implemented in order to provide base data for low cost services for regional analysis performed by humanitarian organizations.

In particular, automated applications have been implemented in order to:

- perform automated classifications of water bodies and use them in flood hazard management activities;
- use open source software for the production of phenological parameters maps suitable for the monitoring of vegetation status.

The first application is a proposal of a solution of an open issue in the remote sensing community as the automation of the classification of flooded areas from satellite optical data. As a matter of fact, classification of water bodies in images coming from optical sensors, the presence of cloud coverage prevent from the acquisition of the earth surface, and the shadows due to clouds can be wrongly classified as water bodies because of the spectral response very similar to the water one. Self-developed algorithm proposed in this work allows classifying the data of reflectivity and to temporally composite them in order to minimize cloud coverage effects and, therefore, identify flooded areas for each considered event. The phases of analysis, development and test of the procedure have been accomplished for sites with different morphological characteristics (Pakistan, Mozambique and Bangladesh) and using different input data; obtained results are mentioned in the next chapters. Different algorithms are applied for different uses: real time ad historical archive.

In food security, the monitoring of the vegetation status among the time is an important mean to predict famines and plan suitable interventions. The second application described in this work involves the integration of existing open source software in an automated procedure, which, starting from remote sensed data, allows the identification of several derived parameters, suitable for the definition of the phenological cycles, for the monitoring of the vegetation status with different spatial resolutions.

3.3 THE TEST AREA: BANGLADESH

In order to perform a first test of the methodology, a particular prone to natural disaster country where WFP operates and with an high population density has been chosen: the Bangladesh. Bangladesh is one of the most densely populated countries in the world, with 159 million people in middle 2009 living on 148 thousand square km [http://en.wikipedia.org/wiki/Bangladesh]. Poverty affects almost 50% of the population and 30 million people can be considered ultra poor (Bangladesh -European Community, Country Strategy Paper for the period 2007-2013).

FIGURE 3.8 BANGLADESH STATE

Most of the country lies within the broad alluvial delta formed by the confluence of the rivers Meghna, Ganges and Brahmaputra. The climate is tropical monsoonal with 80 percent of its heavy rainfall occurring between late May and mid-October.

The agriculture of Bangladesh is relatively productive with rice dominating crop production, grown in three distinct seasons (two rainfed and one irrigated).

On the basis of different climate and soil conditions, four types of rice have been identified in the course of time; three are grown in rainfed conditions (Aus and Aman) and one under irrigation (Boro). Wheat is cultivated during the winter under irrigation though the mild winter temperatures restrict yields. Other cereals, such as maize, are grown on very small areas (FAO Special Report, 1998).

Cereal production normally falls slightly short of demand: only during favorable years free from floods and droughts, self-sufficiency has nearly been reached.
Natural calamities, such as floods, tropical cyclones, tornadoes, and tidal bores occur almost every year, combined with the effects of deforestation, soil degradation and erosion.

3.4 METHODOLOGY

In order to provide instruments for the monitoring of the territory in case of flood events, two applications have been developed. The first application is suitable for the study of the extent of flooded areas, the second one for the analysis of the vegetation. For both the cases it has been decided to focus on particular steps in order to proceed in a structured way during the activities development. Therefore the necessary adopted steps are:

1. need assessment phase of the future users of the system;
2. literary review that allows the author to get inputs on what has already been done and to avoid the replication of existing systems;
3. selection of the input dataset suitable for the requirement;
4. choice and eventual improvement of the most suitable existing procedures for data processing;
5. test and validation of the procedures;
6. implementation of the procedures.

The first two phases are essential for the definition of the final procedures for data processing and classification to adopt in the applications developed. Details on the implemented procedures and applications (points 3 and 4) are provided in the next two chapter, in which the analysis of flooded areas and vegetation are treated separately, in order to better show the different peculiarities of each component, while an application of the integration of the data obtained by the use of the described procedures is provided in chapter 6.

In the following sections specific features of other steps are presented (points 1,2,5,6)

3.4.1 NEED ASSESSMENT

The need assessment is the first, necessary step for implementing successful procedures and applications. A need assessment is a systematic look at how the activities are currently performed by the future users and which are the spatial data needed to correctly carry out their activities.

During the need assessment phase, particular attention has to be paid to the definition of the objectives of the procedures to be developed, the Ithaca and WFP activities that must be supported and the specific output product and their format in order to provide effective information.

As a matter of fact, not always the final user has a clear idea of what is the most suitable information for his needs and which could be the best way to represent it. Thus, in this phase it is necessary to interact with the final user in order to understand his necessities, focusing on the following factors:

- Purpose of the planned application and definition of the content and format of the output product, according to the user capabilities. Actually, often the final users in WFP Local Offices make not part of the scientific community and it is necessary to understand which kind of software and space on disk they can dispose and which type of elaboration they need to perform on data in order to choose the most suitable format for the applications to deliver them. Vector and georeferenced raster (GEOTIFF or JPEG) are the most suitable format for users that only need to integrate the information provided in a GIS tool, with their own information. Other times they can need a standalone system on which they can perform some simple operations, in this case web GIS tools are the best option.
- Scale of the analysis. Definition of the elements of interest on the ground and the extension of the area of interest is essential for the choice of the base satellite datasets, which have to present a suitable spatial resolution to the scale of investigation. Low resolution data can be used for regional analysis, while high resolution datasets for a more detailed investigation.
- Updating frequency and urgency for the delivery of the desired products. In order to choose the most suitable data to use for the application to be developed, it is necessary to understand if the generated products needed to be update in real time and to define the range of years interested by the monitoring.

3.4.2 LITERARY REVIEW

Many applications have been developed by different centers of research in the field of flood management and vegetation monitoring. In order to avoid the replication of already existing procedures and tools, and to concentrate on integration and improvement of already existing ones, it was essential to perform an analysis of the status in order to define a range of methodologies among which to choose for the implementation of the procedure.

The different experiences found in literature have been classified according to different features:

- Spatial coverage: as reported in literature spectral characteristics of the elements at the ground can change according to the morphologic and land cover features of the region object of the study;
- Used dataset: it is necessary to consider the coverage of the satellite datasets used in the works found in literature and their time availability (for example, in case of a real time application, almost immediate delivery of the data is needed);
- Cost of the data: this is not a negligible factor, as the monitoring of environmental variables can request the use of several satellite data. Thus, in a contest of support of humanitarian agencies for disaster management in developing countries, procedures that use free or low cost data have been preferred;
- Scale of the analysis: not always analysis performed at a big scale can be generalized at a global scale: in this case it is necessary to understand the characteristics of the elements object of the analysis and decide if the procedure can be applied to a different scale;
- Possibility of automation of the procedures: it is important to choose procedures which do not need the intervention of an operator or that almost limit it just at the minimum possible;
- Processing time: this is an important issue, because, for example, for early impact analysis it is necessary to produce maps in a very short time (24/48 hours after the event), and, in this case, a long time of elaboration can slow down the optimized procedure for map production.

3.4.3 TEST AND VALIDATION OF THE PROCEDURES

Some existing procedures were selected for data classification and analysis, and then it was necessary to test them on the specific area of interest defined. The test was performed using a subset of historical data, in order to have an order of magnitude of the achievable accuracy.

A further issue specific of this phase was the choice of the reference data to use for the validation: these data couldn't be obtained easily and it was also difficult to obtain feedback from the field by the humanitarian agencies. In order to validate the products (such as flooded areas maps) obtained using the selected procedures, information deriving from satellite imagery have been mainly used in the work of thesis.

3.4.4 IMPLEMENTATION OF THE PROCEDURES

Since a huge amount of satellite data had to be processed in order to provide an historical archive of flooded areas and of vegetation conditions, the methodology for the treatment of the data needed to be based on several procedures that can work in an automatic way.

Therefore, all the operations which normally require the intervention of the operator need to work automatically: that is for image downloading, restoration, enhancement, classification, transformation, updating, and publication operations. This is the basic idea adopted in order to develop the procedures objective of the thesis.

The main software used for the analysis of the data is the RSI (Research Systems Inc) ENVI® (the Environment for Visualizing Images), which is an image processing system able to provide comprehensive data visualization and analysis tools for images of any size and any type.

ENVI is an application developed in the Interactive Data Language (IDL®), a programming language which allows an interaction with high-level data through simple instructions. Many ENVI processing routines are available in appropriate libraries, permitting to combine, in a self-developed routine, all of the functionalities provided for the ENVI program. For these reasons IDL was chosen as the main language for the developing of the automatic procedure.

As mentioned in chapter one, in the restoration phase are included geometric correction, calibration and mosaicking of satellite imagery.

For the applications described in this thesis satellite radiometrically calibrated products have been used. For some products, the only restoration operation needed was the resizing on the area of interest, while for other ones it has been necessary also to mosaic, resize and reproject them in the desired reference system.

In the process of image transformation are included the operations of time series analysis and the extraction of features.

In order perform restoration and resize operations on some particular datasets, tools that can be called by a main program in a batch mode have been used. Batch processing is the execution of a series of programs ("jobs") on a computer without manual intervention.

The main free tools used for the elaboration are:

- MODIS MRT tool
- VGT extractor
- TIMESAT and TIMESATimage

They will be described in detail in the next chapters.

CHAPTER 4: PROCEDURE FOR THE AUTOMATIC CLASSIFICATION OF WATER BODIES

Automated classification of water bodies is still an open issue in the community of remote sensing for emergency response. Many organizations and centers of research all over the world provide different solutions to this topic, implementing different indices for rapid mapping of flooded areas. Therefore, in order to avoid the replication of already existing system and tools, after the need assessment phase, a first literary review of the available methodologies for the detection of water bodies has been performed. In order to create an archive of flooded areas, with the elaboration of huge quantity of data, only techniques that allow detecting the flooded area with a simple processing phase have been considered (mainly histogram thresholding and indices). Some modifications to the existing classification techniques have been implemented, in order to adapt the methodology to the constitution of an historical archive of water bodies.

4.1 NEED ASSESSMENT

As mentioned in the previous chapter, Ithaca is involved both in the immediate post-event phase for the production of rapid mapping on the extension of flooded areas and in the pre-event early warning phase with the development of an early warning system based on satellite rainfall data.

The first request that came from Ithaca team was the development of an automated procedure for the detection of water bodies at a regional scale applicable in different developing countries which can be helped by the intervention of WFP food aid, like Mozambique, Niger in Africa, to Pakistan, Afghanistan, Bangladesh, Myanmar, Philippines in Asia or Central America and many others.

In particular, two kinds of applications were agreed in order to improve and optimize Ithaca activities in pre and post event phases: the first one is the processing of a huge quantity of satellite data for the production of an historical archive of water bodies which can provide useful scenarios both for Ithaca early warning system and also for WFP users, while the second one is a tool for the automation of the development of a water bodies layer for rapid mapping in the post event phase.

For the development of the historical archive, in order to allow different levels of analysis both for the Ithaca team and for the final users of WFP, it was decided to distribute data mainly in two levels: a first level of raster data for internal analysis and a web GIS application that can easily enable members of local offices of humanitarian agencies to get the information they need.

The raster data will be used by Ithaca to perform historical analysis on the past flood events, by a study of the seasonality of water bodies (variation of the extension of water bodies in the time) and the extraction of reference water, information necessary to avoid misinterpretation of recurrent flood events or irrigation for flood events.

For the real time application is instead expected the capacity to detect flooded areas and at the same time to indicate in which areas it is not possible to get information on the ground because of cloud coverage.

4.2 LITERATURE REVIEW

Remote sensing technology, especially space technology, provides an economically feasible alternative means of supplementing traditional hydrologic data sources. Modern possibilities of space observation allow getting the image of different spatial resolution of large surface area with different periodicity that cannot be provided by hydrometeorological network with its discrete supervisions in time and space. Therefore the effective monitoring and prediction of floods and their consequences becomes impossible without assimilation of information of Earth observation from space.

Actually, both active (radar) and passive (optical) sensor are commonly used to get information on water bodies. 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 behavior 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. Since it is not available a complete archive of radar image which can allow historical analysis on developing countries, for this project it was decided to focus on optical imagery which have a more complete archive and more competitive costs. Methods applied on optical imagery 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.

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. Since the institution of an historical archive requires the processing of several satellite data, it was decided to investigate on classification procedures that require short processing time. Thus the literary review was focused on simple classification techniques, mainly based on indexes derived from differential band ratios or histogram thresholding. The advantage of the indices is that they make threshold values more independent from image acquisition parameters. In particular, according to literature two indices are commonly used to identify and classify flooded areas: the Normalized Difference Vegetation Index (NDVI) and the Normalized Differential Water Index (NDWI), .accompanied by histogram threshold and band ratios techniques (HU Zhuowei, August 28-29, 2007).

The NDVI is an index generally used for vegetation analysis (see chapter 5), but it can find application also for the detection of water bodies. Its definition is:

$$NDVI = (\rho_{NIR} - \rho_{RED})/(\rho_{RED} + \rho_{NIR})$$

where ρ_{RED} and ρ_{NIR} stand for the spectral reflectance measurements acquired in the red and near-infrared regions, respectively.

Since water absorbs energy in the IR band, NDVI present low values in correspondence of flooded areas but also in correspondence of bare soils, which have similar characteristics both in the visible than in the IR band. For this reason, in order to extract water bodies, NDVI values are often combined with IR histogram thresholding, which allows to mask bare sole areas.

The NDWI is instead generally used for describing temporal and spatial dynamics of surface moisture, but application have been found which use this index combined to IR thresholding. Unfortunately a unique definition of the NDWI 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 diffused 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:

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

where ρ_{GREEN} and ρ_{NIR} stand for the spectral reflectance measurements acquired in the green and near-infrared regions, respectively.

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:

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

where ρ_{GREEN} and ρ_{SWIR} stand for the spectral reflectance measurements acquired in the green and Short-wavelength infrared regions, respectively.

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

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

where ρ_{BLUE} and ρ_{NIR} stand for the spectral reflectance measurements acquired in the blue and near-infrared regions, respectively.

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:

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

where ρ_{RED} and ρ_{SWIR} stand for the spectral reflectance measurements acquired in the red and Short-wavelength infrared regions, respectively.

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

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

where ρ_{NIR} and ρ_{SWIR} stand for the spectral reflectance measurements acquired in the near-infrared and Short-wavelength infrared regions, respectively.

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). For this reason it has been necessary to perform test on the different versions of NDWI in order to choose the best definition for the classification.

In literature, the NDVI and NDWI indices are applied to different dataset at different spatial and time resolution, such as Landsat, AVHRR, MODIS, very high resolution data. In order to perform an analysis at a regional scale, it was chosen to use as base data MODIS products. In particular, for the historical analysis MODIS surface reflectance products have been chosen, because they present a 10-years archive and, with their 250 maximum spatial resolution, allow a discrete regional historical analysis on water bodies. Since MODIS surface reflectance data are available only four days after the acquisition, for the real time application, the MODIS rapid response system datasets were chosen, since they are generally available few hours after the acquisition.

The issue due to cloud coverage, which is the main cause of missing information on the ground, is generally solved by the use of compositing techniques. By a literary review, it resulted that the simplest and more easily implementable techniques perform a temporal composite of data, by taking into account only the best reflectance data according to different factors (for example the zenith angle of satellite or presence of cloud or missing values). After the creation of an image with all the best values in the reference period, the image is classified and the water bodies are extracted.

Thus, the methodology for water bodies' classification will include two main features: extraction of water bodies by the use of suitable indices and thresholds and a compositing part for reducing the effects of cloud cover.

In the next paragraphs we describe the MODIS products used for the analysis, and the experimental procedures for the choice of the best methodologies of classification in the area of interest (Bangladesh, see chapter 6).

4.2.1 SELECTED DATASET

As previously mentioned two products have been chosen for their spatial and time characteristics: the MODIS surface reflectance product and the MODIS rapid response system subsets.

4.2.1.1 THE MODIS SURFACE REFLECTANCE PRODUCT

MODIS Daily Surface Reflectance products, are product derived from the raw data acquired by Terra and Aqua satellites, and contain the surface spectral reflectance as it would be measured at ground level, in the absence of atmospheric scattering or absorption. On the raw product are implemented the correction for gaseous scattering and absorption, for aerosol effects, for the presence of cirrus, for BRDF coupling and for surface adjacency effects (see Vermote, 1999 for further details). In order to operate these corrections other MODIS products are used, such as atmosphere and geolocation ones (see figure 4.1 for more details). The product is provided at different spectral, spatial and time compositing characteristics (see table 4.1).

In order to perform analysis at a regional scale the product with the highest spatial resolution have been considered: the one at 250 m spatial resolution (MOD09GQ and MYD9GQ), which provides a daily estimate of the surface spectral reflectance for two bands, the red one (620-670 nm) and the IR one (841-876 nm), and the one at 500 m resolution (MOD09GA and MYD09GA), which provides 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 250 m product is the most suitable for the scale of investigation, but not for its spectral characteristics, since it provides only two bands, which only allow histogram thresholding techniques and the calculation of the NDVI index. 500 m product, which a bigger spectral resolution, allows to calculate both the NDVI and NDWI found in literature, of course with a loss of spatial resolution. A validation phase, different for each area of interest will allow choosing which characteristic between spatial and spectral resolution will provide the best results.

TABLE 4.1 MODIS SURFACE REFLECTANCES PRODUCTS

product name	satellite	bands	format	Spatial resolution	Temporal frequency
MOD09A1	Terra	Surface Reflectance Bands 1–7	Tile	500m	8 Day
MOD09CMG	Terra	Surface Reflectance Bands 1–7	CMG	5600m	Daily
MOD09GA	Terra	Surface Reflectance Bands 1–7	Tile	500/1000m	Daily
MOD09GQ	Terra	Surface Reflectance Bands 1–2	Tile	250m	Daily
MOD09Q1	Terra	Surface Reflectance Bands 1–2	Tile	250m	8 Day
MYD09A1	Aqua	Surface Reflectance Bands 1–7	Tile	500m	8 Day
MYD09CMG	Aqua	Surface Reflectance Bands 1–7	CMG	5600m	Daily
MYD09GA	Aqua	Surface Reflectance Bands 1–7	Tile	500/1000m	Daily
MYD09GQ	Aqua	Surface Reflectance Bands 1–2	Tile	250m	Daily
MYD09Q1	Aqua	Surface Reflectance Bands 1–2	Tile	250m	8 Day

The products are provided in a grid format, into the SIN grid (Sinusoidal grid, see chapter 2). Therefore different regions of the earth are represented by different squares of the grid (tile) of the dimension of 10°x10°. In order to use a more standard reference system, (generally Geographic WGS84 system), the product will have to be reprojected.

The product can be reprojected in different way: by using apposite routines in ENVI or by a specific tool provided by NASA to reproject the data in the desired reference system, the MODIS MRT Tool. From tests on different tiles, the MODIS MRT Tool resulted to be the quickest to process the data, so it has been used for the integration in the procedure.

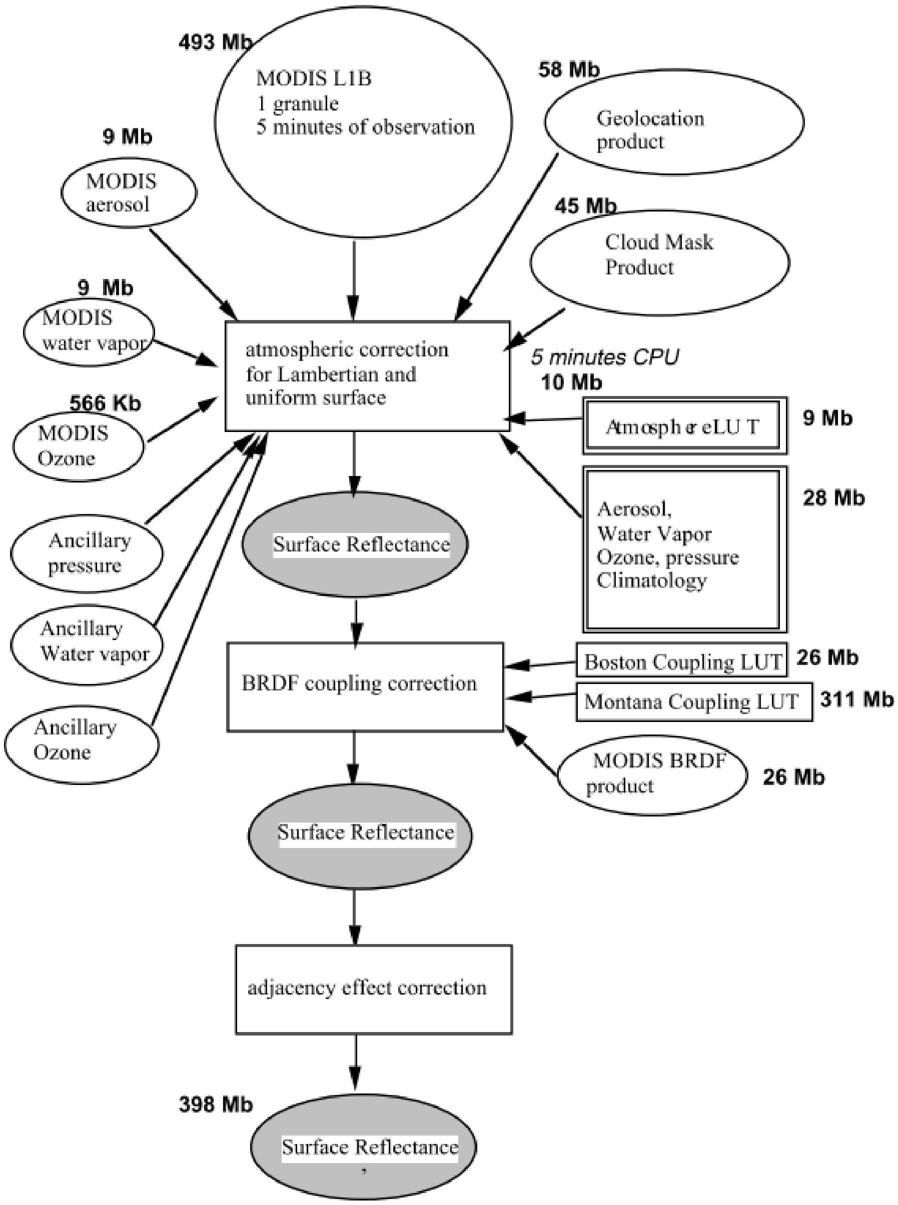

FIGURE 4.1. VERSION 1 ATMOSPHERIC CORRECTION PROCESSING FLOW CHART

MODIS MRT TOOL

The MODIS Reprojection Tool (MRT) is a tool developed by The Land Processes Distributed Active Archive Center (LP DAAC), which was established as part of NASA's Earth Observing System (EOS) Data and Information System (EOSDIS) initiative to process, archive, and distribute land-related data collected by EOS sensors. The tool has been developed to support MODIS data which are projected to the Integerized Sinusoidal (ISIN) mapping grid or in the Sinusoidal (SIN) projection and stored in HDF-EOS (Hierarchical Data Format – Earth Observing System). The MRT supports all MODIS versions and projections.

The MODIS Reprojection Tool is software designed to help individuals work with MODIS data by reprojecting MODIS images (Level-2G, Level-3, and Level-4 land data products) into more standard map projections. If desired, the user may reproject only selected portions of the image (spatial subsetting) and only selected image bands (spectral subsetting). The software outputs MODIS data in file formats that are supported by existing software packages (raw binary and GEOTIFF) as well as HDF-EOS. The MODIS Reprojection Tool runs on several platforms, including Sun Solaris workstations, SGI IRIX workstations, Linux, 64-bit Linux, Microsoft Windows, and Mac OS X.

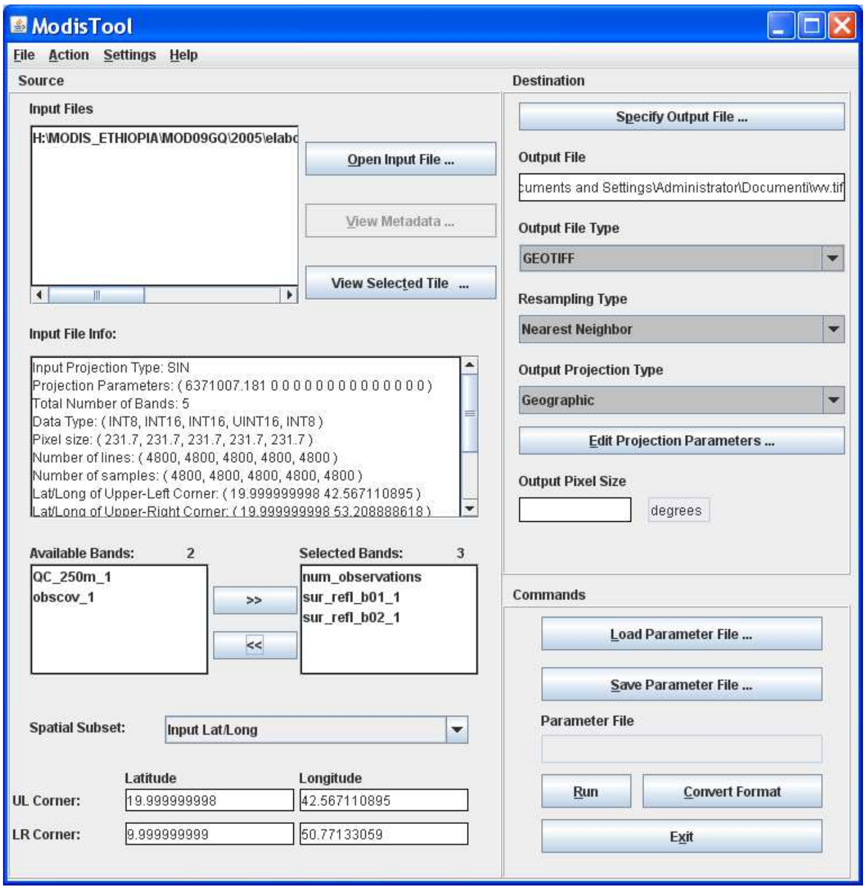

FIGURE 4.2. MODIS REPROJECTION TOOL GUI

The heart of the MODIS Reprojection Tool is the resampler and MRTmosaic, executable programs that may be run either from the command-line or from the MRT Graphical User Interface (GUI, see figure 4.2). The command line option has been used for the elaboration of the data for the elaboration of the archive, since it allows background processing of large MODIS data files.

4.2.1.2 MODIS RAPID RESPONSE SYSTEM SUBSETS

The MODIS Rapid Response System was developed to provide daily satellite images of the Earth's landmasses in near real time. True-color, photo-like imagery and false-color imagery are available within a few hours of being collected, making the system a valuable resource for the emergency response community. The system is freely available to everyone, scientists, operational users, educators, and the general public. The most part of the globe is covered by these datasets, but still some regions are not covered by this dataset and in this case it is necessary to elaborate the raw data by routines provided by the distributors of the MODIS rapid response system subsets (see figure 4.3).

The Rapid Response system receives MODIS Level-0 data from a special near real-time data feed. These data are processed to Level1B using a specific code from the Ocean Color/SeaDAS group. The calibration look-up-tables for the Level0 to Level1B processing are updated regularly when they are received from the MODIS Calibration and Support Team. The geolocation for the Terra satellite uses the same entrained data as used by the standard processing. For Aqua, Rapid Response uses predicted satellite ephemeris rather than the definitive ephemeris. Normally, the differences in location are relatively small, under 400m and often under 100m.

Rapid Response true and false color images are created using the “corrected reflectance” algorithm developed by Jacques Descloitres at NASA/GSFC. The purpose of this algorithm is to provide natural looking images by removing gross atmospheric effects, such as Rayleigh scattering, from MODIS visible bands 1-7. In particular, the algorithm produces quasi-true-color images from MODIS bands 1, 3, and 4. In contrast, the MODIS land surface reflectance product (MOD09) is a more complete atmospheric correction algorithm that includes aerosol correction, and is designed to derive land surface properties. In clear atmospheric conditions the corrected reflectance product is very similar to the MOD09 product, but they depart from each other in presence of aerosols, which are present in the corrected reflectance product, but not in the land surface reflectance product.

After the correction, the data are converted, with a loss of information, into to RGB brightness values. A non-linear enhancement is used to increase the brightness of darker portions of the scene, such as land or water.

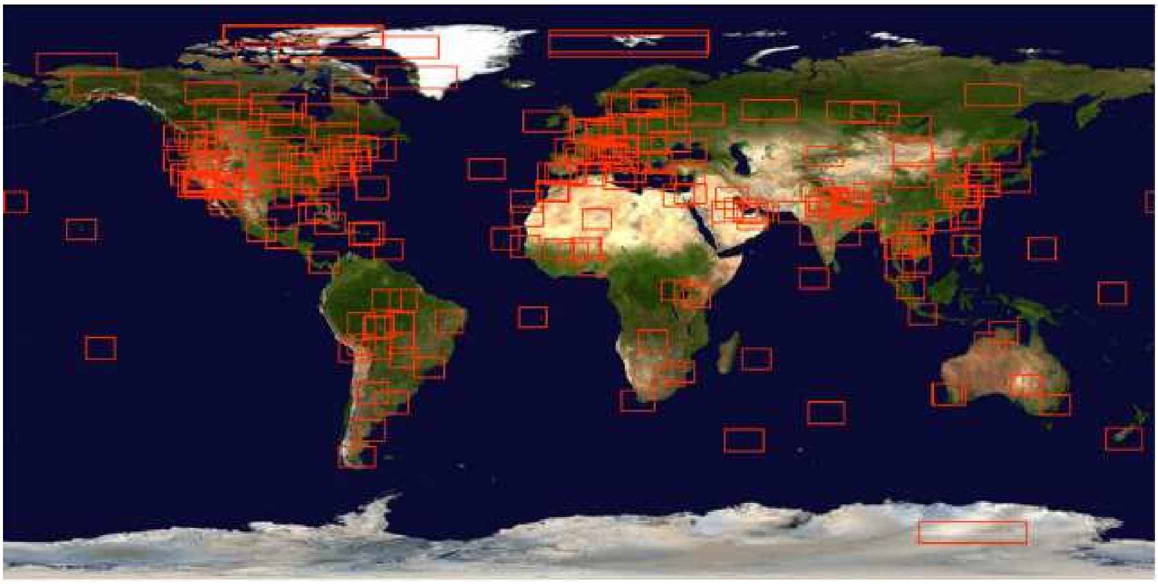

FIGURE 4.3 - THE MODIS RAPID RESPONSE SYSTEM AVAILABLE SUBSETS.

Even if the lack of aerosol correction badly affects the classification process, the near real-time image availability is fundamental to built-up an operational emergency management process.

Relatively to spectral characteristics, the following additive syntheses, suitable for the extraction of clouds and snow masks and for the classification of water bodies, are available:

- 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).

4.3 SELECTION OF SUITABLE DATA PROCESSING PROCEDURES

After the selection of the base dataset, it has been decided to approach to the choice of the methodology for water bodies' classification firstly deciding the index/indexes to use, then the compositing algorithm and later testing the methodology on available data (see figure 4.4). The first test on data is performed on the version of the chosen dataset, which should allow temporal consistence and less variation due to atmospheric condition, the MODIS surface reflectance product, which is already corrected from the effects of the atmosphere. Later the methodology is tested on the data of the MODIS rapid response system in order also to verify the applicability of this data not corrected by aerosols effects and corrupted by the compression process.

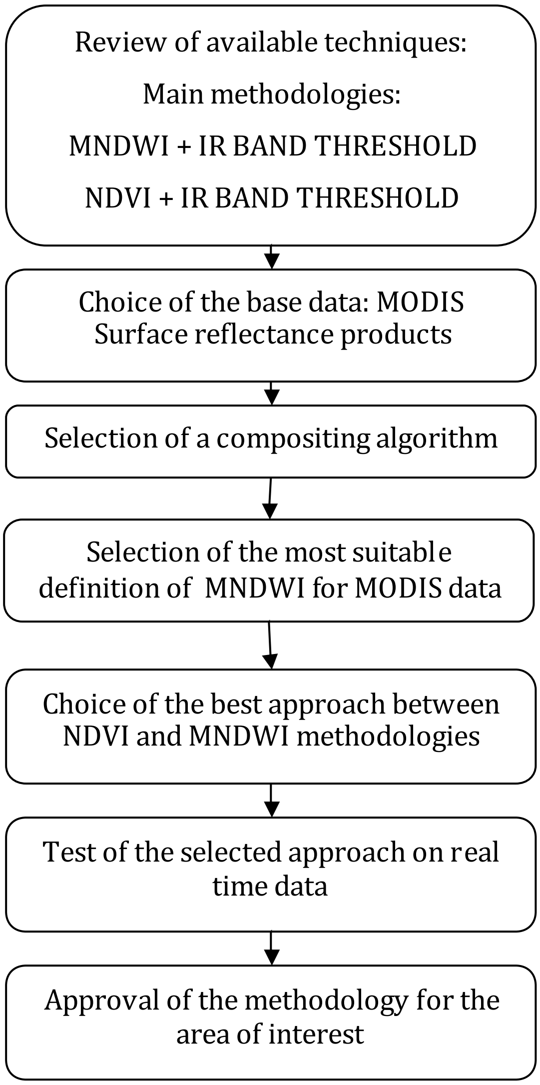

FIGURE 4.4. TESTING PHASE FLOW CHART

4.3.1METHODOLOGIES SELECTED FROM LITERARY REVIEW

From literary review, two methodologies for water bodies' extraction emerged. The first approach includes:

- MNDWI index calculation, using MODIS spectral bands 4 and 6, which are available only for the product at a spatial resolution of 500 m, the product MOD09GA (see table 4.1);
- Masking of areas covered by clouds or snow, through thresholding procedures in the visible bands.

The second approach includes:

- NDVI index calculation, using MODIS spectral bands 2 and 1 of product MOD09GQ, which only has the two bands of interest and 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.

4.3.2 SELECTION OF THE COMPOSITING ALGORITHM

In order to detect water bodies, it was necessary to choose the compositing algorithm, which allows to fill the daily missing information. A part from cloud coverage also the presence of cloud shadows has been considered. As previously mentioned, dark cloud shadows, since they have spectral characteristics quite similar to those of water bodies, can be interpreted by an automated algorithm as water areas (Luo et al., 2008). For this reason it has been decided to adopt a compositing technique different from the ones encountered in literature. This approach first aims at the classification of the available daily images and later composite them. The steps performed by this composition algorithm are:

- automated detection of water bodies from daily satellite imagery
- combination information on water bodies coming from the 2 daily images of MODIS Terra and Aqua satellite;
- choice of the time compositing period, according to final users needs and cloud coverage;
- time compositing , application of a threshold to the number of times in which a pixel is classified as water in order to assign it to the class water bodies;
- extraction of water bodies.

In this way water bodies are first detected on every available image with the technique chosen for the area of interest. In the phase of combination of Terra and Aqua data, each pixel of the image which results to be covered by water in almost one of the two daily available images is considered as water body. The time compositing is performed by counting the number of days a pixel results to be covered by water in the compositing period (see figure 4.5).

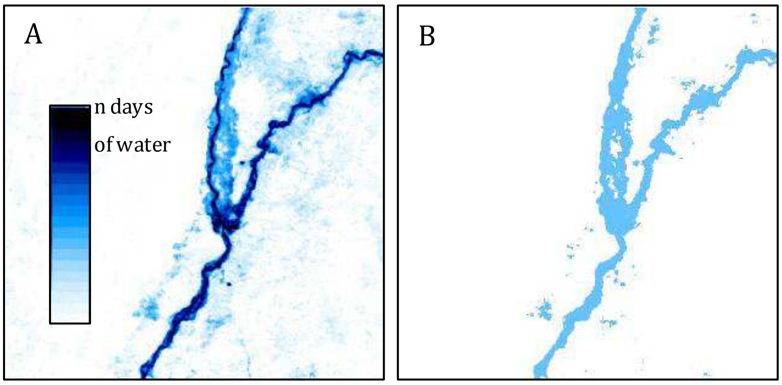

FIGURE 4.5. – TIME COMPOSITION (A) AND RESULT OF CLASSIFICATION ERRORS MASKING, USING THE TEMPORAL THRESHOLD APPROACH (B).

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 cloud shadows 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 over 10) will be considered as areas affected by water. The same approach may be used for the cloud coverage.

This methodology presents the advantage that it can detect pixels that are for more than an established threshold of days covered by water and so the more persistent water and allows to perform a correction on cloud shadow. If associated to the information on how many days a pixel result to be covered by cloud also quality information on the classified pixel can be provided.

4.4 TEST AND VALIDATION OF THE PROCEDURES

Since developing countries all over the world can present different morphologic, land cover and cultivation characteristics and the of NDVI and NDWI indices can be different according to the region of earth occupied by a country, in order to detect water bodies, it can be necessary to adopt different indices and reflectance threshold values for each country. After the choice of the best NDWI definition, two validation steps have been performed: the first for the choice of the best index to extract water on the area of interest from the dataset of surface reflectance data, the second to verify the suitability of real time data for the application, since it is provided in a format (GEOTIFF or JPEG) in which the quality of the image has not been corrected from all the effects of the atmosphere and furthermore has been deteriorated due to the transformation to RGB Brightness values.

4.4.1 CHOICE OF THE BEST DEFINITION OF NDWI

Since from literature emerged the necessity to choose a suitable MNDWI definition for the identification of water bodies, different classifications of images belonging to regions with different geographic features, such as Pakistan, Mozambique have been tested by the author by photointerpretation. The classifications have been performed by the use of several MODIS band combinations for the MNDWI 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, to be effective, 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, and for this reason it was decided to adopt the afore mentioned MNDWI index:

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

where ρ_{GREEN} and ρ_{NIR} stand for the spectral reflectance measurements acquired in the green and near-infrared regions, respectively.

4.4.2 CHOICE OF THE BEST METHODOLOGY OF WATER CLASSIFICATION BASED ON SURFACE REFLECTANCE DATA

In order to choose the best index to classify MODIS surface reflectance products in the area of interest (Bangladesh), the water bodies' information produced by the two procedures based on NDVI or MNDWI have been compared to a reference image of the area of interest.

The first point to solve before the implementation of the procedure was the choice of compositing period. During a flood event, as already mentioned, cloud cover is an important issue to solve. Time compositing allows, in presence of areas that are not covered by clouds for the whole compositing period, to get information on a pixel almost for a percentage of the days belonging to the compositing period. Furthermore flooded areas generally are not persistent for several days on the ground since they are absorbed by the soil and, if the compositing period is too large, they can be mistaken for errors due to cloud shadows (if their permanence on the soil is less than the percentage of days established to distinguish water pixels from cloud shadows ones). Finally, it has been chosen a temporal composite of 10 days, since from a screening on the real time products on the area of Bangladesh, this period resulted to be a good compromise between numbers of days in which a pixel is hopefully not covered by clouds and the relevance of the number of days in which a pixel is covered by water during a flood respect to the compositing period.

In order to obtain a temporal composite, MODIS Terra and Aqua Surface Reflectance data belonging to the 10-days period between 21st May and 1st June were collected. The daily images were classified according the parameters shown in table 4.2.

For each day of the composition time, the water bodies regions were calculated as the union of the pixels classified as water in the Aqua and Terra image. After this operation it was counted how many times each pixel of the area resulted to be classified as a water body. Every pixel that resulted to be classified as a water body for a number of days greater or equal than the established day threshold was finally detected as a water body.

TABLE 4.2 INPUT PARAMETERS FOR THE ESTIMATION OF THE ACCURACY OF THE CLASSIFICATION METHODOLOGIES BASED ON DIFFERENT INDICES

CLASSIFICATION METHODOLOGY	MNDWI + THRESHOLD	NDVI + THRESHOLD
DATASET USED	MOD09GA-MYD09GA	MOD09GQ-MYD09GQ
SPATIAL RESOLUTION	500 m	250 m
COMPOSITING PERIOD	10 days: from 21st May to and 1st June 2009	10 days: from 21st May to and 1st June 2009
INPUT PARAMETERS	MNDWI >0 Reflectance in B2 < 0.15	NDVI < 0.1 Reflectance in B2 < 0.15
DAYS THRESHOLD TO CLASSIFY A PIXEL AS COVERED BY WATER	2 days over 10	2 days over 10

Since no ground information was available, radar imagery, whose efficacy in detecting water is not affected by cloud coverage and which has a higher spatial resolution, has been considered as a reliable reference image for the determination of reference water bodies. During the AILA cyclone Cosmo-Skymed constellation acquired radar images, belonging to the period between the 30th may and the 1st June 2009 and with a spatial resolution of 30 m. Thus the water extension obtained by the classification of a radar image, performed by the application of a threshold value, was considered as water effectively present on the soil.

In order to compare the two methodologies for water bodies' classification the reference radar image and the classification at 500 m were resampled to 250 m.

Once the data to compare had the same spatial resolution, for the two result of classification were calculated three indices of accuracy (see table 4.3):

$$Averall\ accuracy = \frac{\text{number of correct classifications}}{\text{total number of pixels in the image}}$$

$$User\ accuracy_{water} = \frac{\text{number of correct classifications}}{\text{total number of pixel classified as water}}$$

$$Producer\ accuracy_{water} = \frac{\text{number of ground points correctly classified}}{\text{number of points belonging to the category water in the ground}}$$

The overall accuracy provides the extent to which classification process correctly identifies selected classes. The number of pixels that both in the classification image and in the truth image are classified in the same way is assumed to be the number of correct classifications.

User accuracy is a measure of the reliability of a classification. It is a statistic that can inform the user of the map about the percentage of a class which corresponds to the ground-truth class and furthermore provides a quantification of the commission error.

Producer's accuracy is a measure of how much of the land in each category was classified correctly and provides a quantification of the omission error of the proposed classification.

TABLE 4.3 CLASSIFICATION ACCURACIES OF MODIS SURFACE REFLECTANCE DATA ACCORDING TO THE USED METHODOLOGIES

CLASSIFICATION METHODOLOGY	OVERALL ACCURACY	USER ACCURACY (class water)	PRODUCER ACCURACY (class water)
MNDWI + IR THRESHOLD	92,81%	68,85%	71,83%
NDVI + IR THRESHOLD	94,89%	79,85%	76,12%

As shown in table4.3, 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 (MOD09GQ and MYD09GQ), i.e. the one based on thresholding techniques using both NDVI index and infra-red bands.

4.4.3 TEST OF THE SUITABILITY OF REAL TIME DATA FOR THE APPLICATION

In order to test the chosen classification methodology on MODIS rapid response system products, a dataset of 20 images (10 for Aqua and 10 for Terra, belonging to the same composition time used in the first phase, from 21st May to and 1st June 2009) were collected and classified with the methodology NDVI + IR threshold.

Since in these images are not provided reflectance values, but enhanced brightness values, before applying the thresholds and calculate the NDVI index, the inverse transformation of the enhancement process has been performed.

The enhancement table used for the rapid response system is the one reported in table 4.4. In order to obtain the original brightness values, a non linear transformation has been applied and afterwards data were brought to the reflectance ranges 0-1.

TABLE 4.4: NON-LINEAR BRIGHTNESS ENHANCEMENT TABLE

Input Brightness	Output Brightness
0	0
30	110
60	160
120	210
190	240
255	255

The same parameters and time compositing procedure used in test of surface reflectance data was used.

In order to obtain an independent index on the quality of this process of classification, the results were compared to the truth image obtained from the higher resolution radar image.

Furthermore, in order to quantify the loss of information of classification of the rapid response system data due to the lack of aerosol correction respect to the classification of the surface reflectance product, the results of the two classifications were compared. Using classification of the surface reflectance products as the ground truth, the accuracy indices, previously mentioned were calculated. The results of the two tests are shown in table 4.5.

TABLE 4.5 MODIS RAPID RESPONSE SYSTEM DATA CLASSIFICATION ACCURACY ACCORDING TO THE ASSUMED GROUND TRUTH

ASSUMED GROUND TRUTH	OVERALL ACCURACY	USER ACCURACY	PRODUCER ACCURACY
RADAR	95,00%	79,10%	80,50%
SURFACE REFLECTANCE PRODUCTS	96,40%	83,25%	86,40%

Since the loss of information due to the compression can be estimated around the 15% and the accuracy of the data results to be of the same order of the surface reflectance product, it was assumed that these data can be used for the real time application.

4.5 IMPLEMENTATION OF THE PROCEDURES

The first development of the procedure of classification and time compositing has been performed for the case study of Bangladesh (see chapter 6).

Two kinds of applications were required: one for the real time, the other one for the creation and analysis of the historical archive. Both the applications require the automatic implementation of the following steps:

- Download of satellite data
- Data pre-processing and water bodies classification
- Time compositing
- Extraction of flooded areas.

The main frame in which data are processed is IDL, but different applications are also called by the main procedure in order to perform all the operations needed.

4.5.1 DOWNLOAD OF THE DATA

The GNU WGET has been used for data downloading. It is a free utility for non-interactive download of files from the Web and supports HTTP, HTTPS, and FTP protocols, as well as retrieval through HTTP proxies.

The download of the data is performed in a different way according to the type of application (see figure 4.6).

In the case of the historical analysis, the MODIS data, stored in an ftp sites (ftp://e4ftl01u.ecs.nasa.gov/MOLT/MOD09GQ.005/, ftp://e4ftl01u.ecs.nasa.gov/MOLA/MYD09GQ.005/), are downloaded, if belonging to the area of interest, for the entire archive.

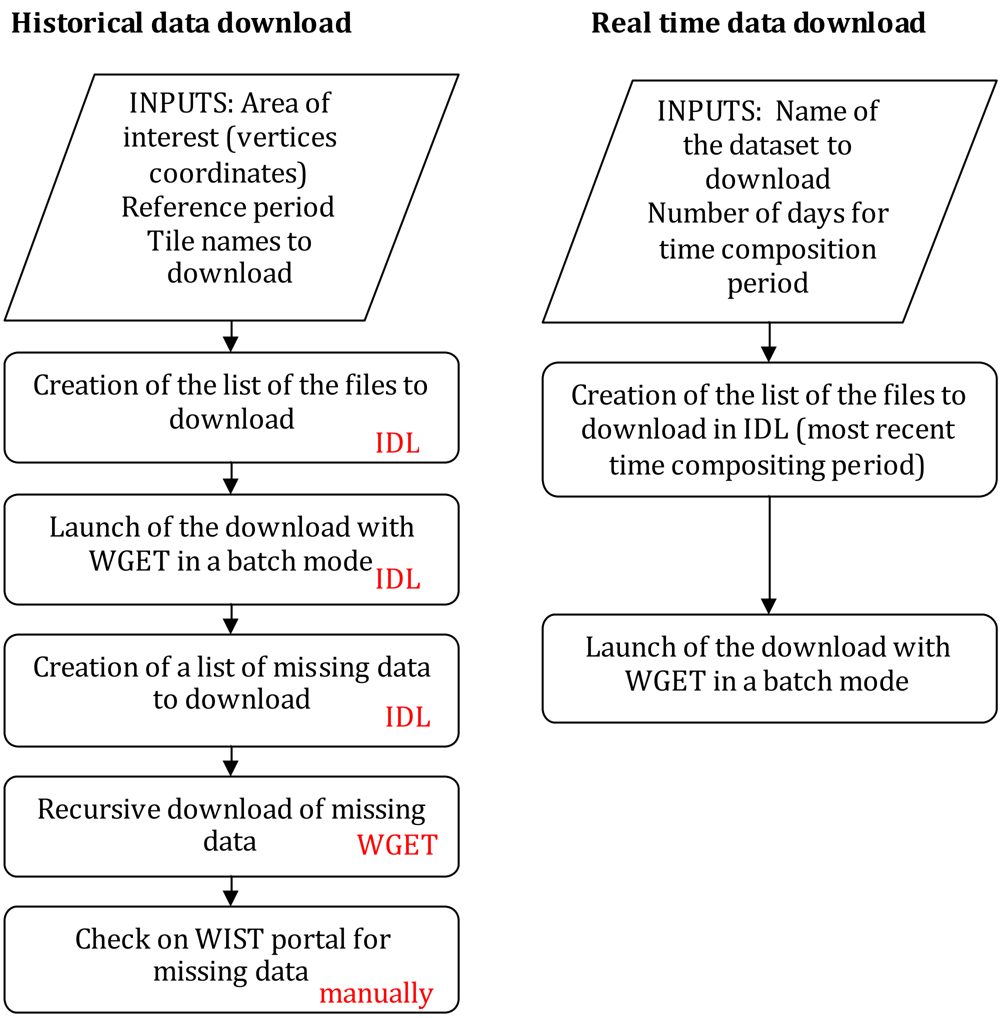

FIGURE 4.6. DATA DOWLOADING FLOW CHART FOR HISTORICAL ARCHIVE APPLICATION AND REAL TIME APPLICATION

In order to provide the right set of data, it is necessary to know the extent of the area of interest, with information of the coordinates of the upper left and lower right pixel of the area of interest.

The MOD09GQ and MYD09GQ data used for this application are distributed in tiles, so before starting the download is necessary to check the name of the tiles to download in the WIST portal, performing a test request with the coordinates of the vertices of the area of interest, for a short period of interest.

After information on the name of the tile and the kind of product to download is available, a list of files to download is generated by an IDL routine, which automatically generates the name of the subfolders where to look for the data and the name of the files to download for all the available archive of MODIS data.

The acquisition of the data is performed by a launch in a batch mode of WGET, which downloads all the files contained in the list prepared by IDL.

With such an amount of data, it is necessary to perform a check on the missing data: it is first performed automatically by an IDL routine which test for each day if all the tiles belonging to the area of interest have been downloaded and create a list of files that still need to be downloaded. After the implementation of this control for 10 times, a second manual check is performed. In this case it is checked if data not available in the ftp site are available in the WIST portal. After this operation, eventual data available only in the WIST portal are downloaded and a list of missing data is compiled.

In the case of the near real time application, the data are distributed in an http link, which varies according to the zone and day of interest (for example: http://rapidfire.sci.gsfc.nasa.gov/subsets/?subset=FAS_Bangladesh.2008141.aqua.721.250m.tif). Only the most recent data are considered according to the time compositing period more suitable for the area of interest. Also in this case the name of the link where to download the data is automatically produced by IDL and the data are downloaded with WGET launched in a batch mode.

4.5.2 DATA PRE-PROCESSING AND CLASSIFICATION

In this phase the data downloaded are projected to the required reference system and resized on the area of interest. in case one tile or real time subset is not sufficient to cover the area of interest, also a phase of data mosaicking is planned (the procedure used for the two applications are shown in figure 4.7).

For the processing of MODIS Surface reflectance data, the MRT Tool is called in a batch mode by an IDL routine to perform the three operations of mosaicking, reprojection, and resizing, while in the case of real time data ENVI and IDL routines are used for this purpose. In order to launch in a batch mode the MRT tool a parameter file is created by IDL before the launch of the tool.

After the resize on the area of interest, the two daily images (one coming from Terra satellite and the other from Aqua) are classified, using the NDVI and IR thresholds tested in the validation phase for detecting water bodies in the surface reflectance. For near real time products, the enhanced brightness values need to be transformed into reflectance ones before the classification process (see figure 4.8). In order to provide information on the cloud coverage and with this on the reliability of data, also the class cloud is included in the classification of the daily image. In this way, in the case of the historical analysis a first archive of couples of daily classified images (one for Terra and one for Aqua) containing information on water bodies and cloud extent is built for the 10 years of MODIS data.

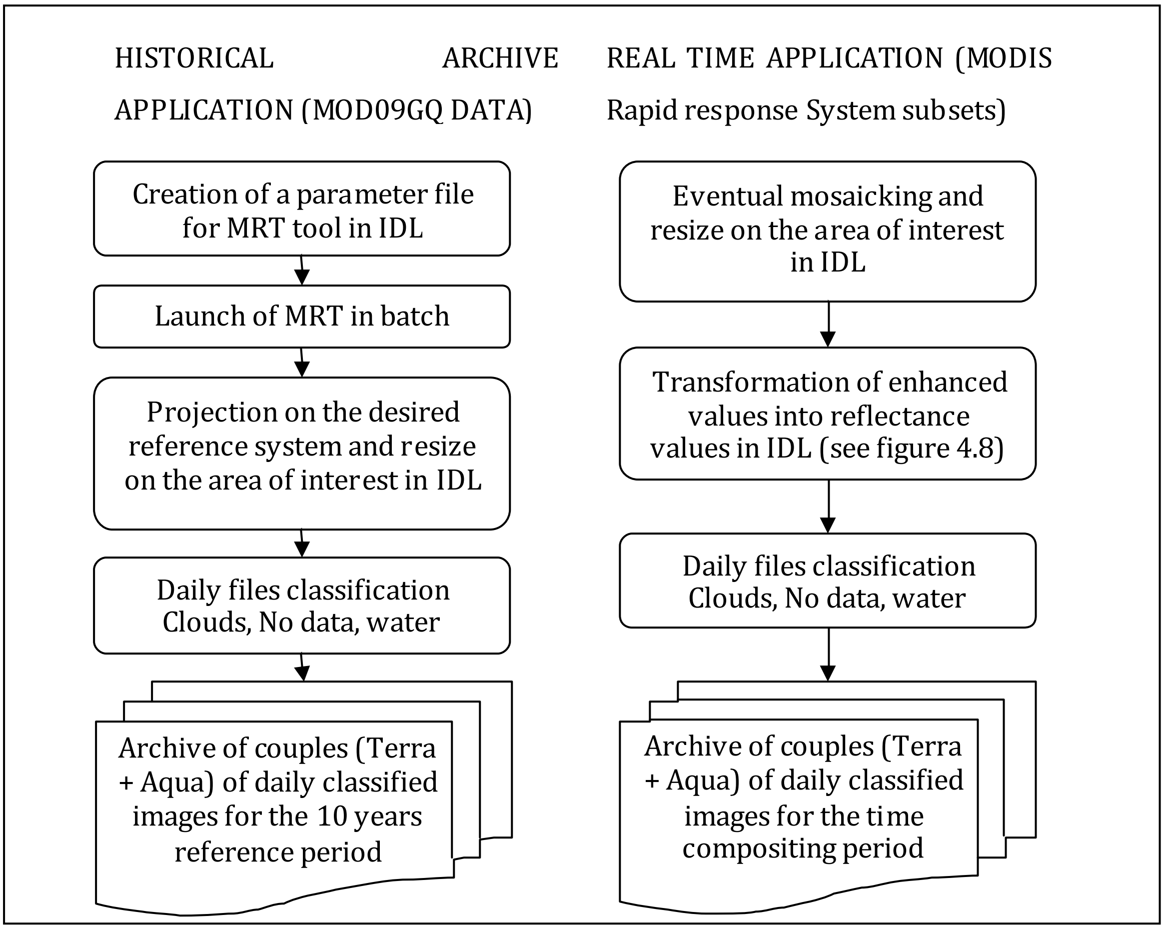

FIGURE 4.7. DATA PRE-PROCESSING AND CLASSIFICATION FLOW CHART ACCORDING TO THE TYPE OF APPLICATION

```
scale=[[0,30,60,120,190,255],[0,110,160,210,240,255]] ;first line: not enhanced values, second line: enhanced values
 for i=0, n_elements(scale)-2 do begin
   in=where (b2 ge scale[i,1] and b2 lt scale[i+1,1], count)
   if count gt 0 then b2[in]=float(scale[i,0])+float((b2[in]-scale[i,1]))*(scale[i,0]-scale[i+1,0])/(scale[i,1]-scale[i+1,1])
   in=where (b1 ge scale[i,1] and b1 lt scale[i+1,1], count)
   if count gt 0 then b1[in]=float(scale[i,0])+float((b1[in]-scale[i,1]))*(scale[i,0]-scale[i+1,0])/(scale[i,1]-scale[i+1,1])
   in=where (b7 ge scale[i,1] and b1 lt scale[i+1,1], count)
   if count gt 0 then b7[in]=float(scale[i,0])+float((b7[in]-scale[i,1]))*(scale[i,0]-scale[i+1,0])/(scale[i,1]-scale[i+1,1])
 endfor
;transformation of 8 bits values into reflectance values
b1=b1/255
b2=b2/255
b7=b7/255
```

FIGURE 4.8 IDL CODE FOR THE TRASFORMATION OF ENHANCED VALUES INTO REFLECTANCE VALUES

4.5.3 TIME COMPOSITING

Since daily images are affected by cloud, they are temporally composited according to the chosen time compositing period. An ad hoc routine has been developed in IDL in order to first combine the couples of daily classifications coming from Terra and Aqua and to count how many times in a compositing period a pixel results to be covered by water. The outputs of this procedure are three types of temporal composites containing information on how many times in a compositing period:

- a pixel results to be covered by water;
- a pixel results to be covered by clouds;
- there is no information on a pixel (to check the number of missing information);

In order to obtain water bodies classification, a synthesis relative to the compositing period the threshold of days, tested in the validation phase, is automatically computed by an IDL routine. Furthermore, if a pixel results for a percentage greater than the 80% of the compositing period covered by clouds or no data are available on this point, in the synthesis file this pixel is marked as covered by clouds or with no available information (this is a very rare case, in the Bangladesh case has happened only once for the 10 years archive).

The time compositing algorithm is the same both for real time and historical archive application, with the difference that for the real time application only one image representing the most updated time compositing period is produced.

4.5.4 EXTRACTION OF FLOODED AREAS

For the extraction of flooded areas it is necessary to compare the data synthesis to water body reference information. Since water bodies' extension can vary according to the seasons, it can be useful to exploit the archive of synthesis on water bodies.
An IDL routine has been implemented in order to perform statistics on the data, the main information extracted are:

- Number of times in the 10 years period that a pixels results to be covered by water in a synthesis period. This information allows the user to understand which are the periods when there is more water on the ground because of seasonality or irrigation.
- Number of times in the 10 years period that a pixels results to be covered by clouds in synthesis period. This information can be used as a reliability index to the first information provided.
- Number of synthesis periods in which a pixels result to be covered by water in the 10 years period. This information can give a general overview on how many times a pixel

can be covered by water in the historical period of reference: the thresholding of the histogram can allow to detect permanent water, seasonal or temporary water , and floodable areas, according to the percentage of occurrence a pixel covered by water in the 10 years period

- Number of synthesis files (10-days period) in which a pixels result to be covered by clouds in the 10 years period. This information can be used as a reliability index to the information on seasonal or permanent water provided by the analysis of the histogram in the previous point

These information need to be analyzed just manually only once in order to extract the reference water. Thus flooded areas are detected automatically by the comparison of the water bodies with the extracted water references both for historical dataset and real time data (see data flow in figure 4.9). The application of this information is shown in chapter 6.

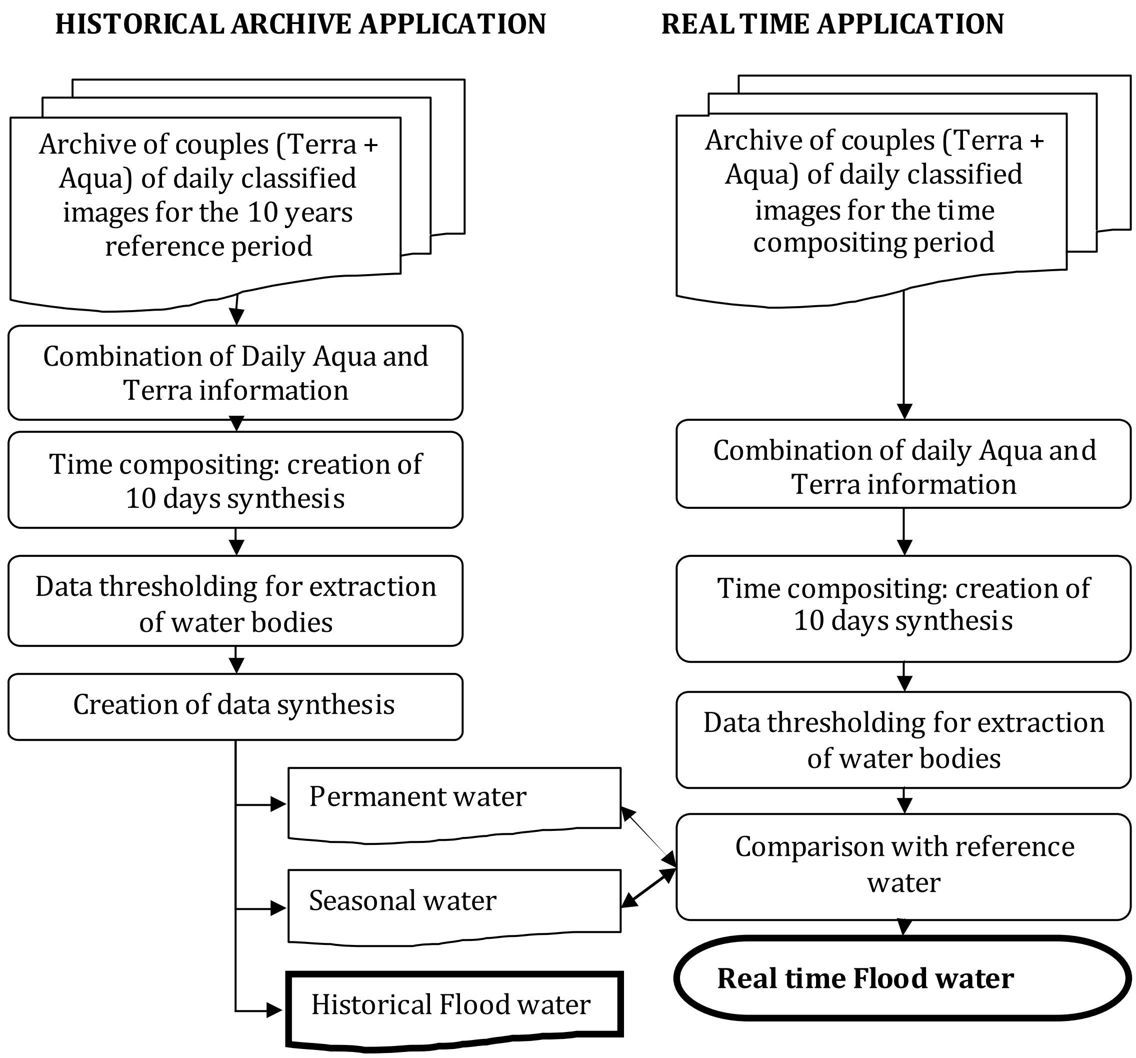

FIGURE 4.9 TIME COMPOSITING AND EXTRACTION OF FLOODED AREAS FLOW CHART FOR HISTORICAL AND NEAR REAL TIME APPLICATIONS

4.5.5 AUTOMATION AND PROCESSING TIMES CONSIDERATIONS

The developed procedures for the classification of water bodies allow a good degree of automation, since the intervention of the operator is required only for the historical analysis in a few phases concentrated at the beginning and at the end of processing of data:

- choice of the compositing period and indices thresholds to adopt in the validation phase;
- input of the area of interest (name of the tiles of MOD09 and MYD09 products and name of the subset of the MODIS Rapid Response System);
- manual check on missing data after the automated one;
- extraction of reference water according to water bodies seasonality.

For the development of the historical archive a dataset corresponding to 10 years of acquisition of MODIS data has been processed. Since Bangladesh surface is covered by 2 tiles of the MODIS sinusoidal grid the number of processed file can be so divided:

- About 7000 tiles of the product MOD09GQ (the MODIS terra product from 2000 to 2009, all the available archive);
- About 5800 tiles of the product MYD09GQ (the MODIS aqua product from 2002 to 2009, all the available archive);

The total space on disk occupied by base data was 750 Gb. Only the 1.3 % (164 files) of the total amount of data is not available after the manual check. Furthermore the missing data are concentrated in the first three years of MODIS mission, since the 85% of them belong to the period between February 2000 and December 2003. Hopefully the percentage of missing data will decrease in the next years.

The download and processing of the whole archive required around one month of downloading and elaborations on a single pc. That can be divided, according to the activities as reported in table 4.6:

TABLE 4.6 PROCESSING TIMES

	Percentage of total processing time
Downloading	45%
Reprojecting-Moisaicking-Resizing	36%
Classification	14%
Compositing	5%

Results shows that downloading and georeferencing phases are the most critical, while classification and temporal composite production require significant lower processing times. After the creation of the historical archive, no other heavy processing is required and therefore the update of the archive and the real time processing will require only few minutes, since data will be downloaded and processed once a day.

CHAPTER 5: PROCEDURES FOR THE AUTOMATION OF PHENOLOGICAL INFORMATION EXTRACTION

Phenology is the science dealing with the influence of climate on the recurrence of annual phenomena of animal and plant life, such as leafing and flowering of plants, maturation of agricultural crops, emergence of insects, and migration of birds. As a signal of the impact of environment on vegetation, phenology plays an important role in research of global carbon circulation, continental net primary production, biogeochemical circulation and so on. Changes in phenology further depict a canopies' integrated response to environmental change and provide valuable information for land degradation studies, integrated pest and invasive species management, drought monitoring, wildfire risk assessment, and agricultural production. Satellite-derived phenology allows monitoring of terrestrial vegetation on a global scale and provides an integrative view of the regional landscape-scale processes.

One of the main used indicators for the study of vegetation is the Normalized Difference Vegetation Index (NDVI), which uses the visible and near-infrared bands of the electromagnetic spectrum, in order to detect vegetation from the rest of land surface and to understand the state of health of vegetation. The NDVI is calculated from these individual measurements as follows:

$$NDVI = (\rho_{NIR} - \rho_{RED})/(\rho_{RED} + \rho_{NIR})$$

where ρ_{RED} and ρ_{NIR} stand for the spectral reflectance measurements acquired in the red and near-infrared regions, respectively. The values resulting from this calculation range from – 1 to +1, where negative values usually indicate a vegetation absence and approximately 0.80 generally indicates very high vegetative biomass. NDVI has found a wide application in vegetative studies as it has been used to estimate crop yields, pasture performance, and rangeland carrying capacities among others. It is often directly related to other ground parameters such as percent of ground cover, photosynthetic activity of the plant, surface water, leaf area index and the amount of biomass. Seasonal changes in Normalized Difference Vegetation Index (NDVI), a self-standardizing measure of land surface greenness, have proven useful in tracking land surface phenology. Time series of NDVI – often employed as per-pixel series of NDVI over time – are particularly useful for phenological research, since they provide a means to capture time-sensitive phenological shifts and vegetation developmental stages. NDVI increases and decreases with annual cycles of superficial vegetation green-up and senescence, and much of the research involving NDVI incorporates multi-temporal data sets for this reason.

In this chapter a procedure for the elaboration of the NDVI index for phenological studies in disaster management is shown.

5.1 NEED ASSESSMENT

In order to increase the possibilities of analysis of the WFP in order to preview and manage shortage in the harvest, it has been requested an instrument of monitoring of the vegetation status by the analysis of NDVI indices. Ithaca, as mentioned in chapter 3, has developed a tool to analyze automatically NDVI trends at a global scale in the last 25 years (see figure 5.1). In order to integrate this information with seasonality information, it was agreed to produce phenological information derived from NDVI indices, which will be validated in a second phase with the help of experts of the area. A range of possible datasets in raster format will be proposed and displayed in a web-GIS application.

In order to provide an extended archive with available NDVI datasets, the final per-pixel phenological information has to be created automatically and has to provide:

- Phenological parameters layers for each season of the past years (according to the historical archive of the considered satellite imagery). This information will allow the analysis of single years and the extraction of statistic information for long term behavior of the parameters in the area of interest.
- Phenological parameters statistics which can resume the long term behavior of phenological parameters in the different seasons of the year. This information will allow detecting eventual anomalies from the long term behavior of vegetation by a comparison with the historical analysis.

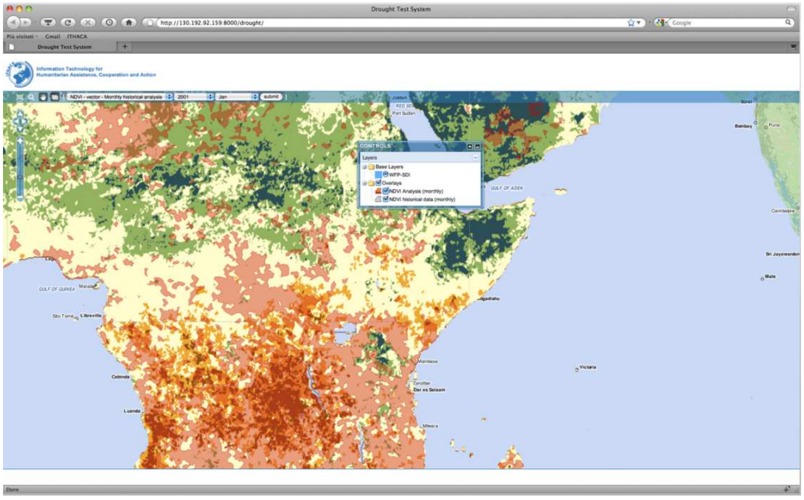

FIGURE 5.1 ITHACA DROUGHT MONITORING SYSTEM WEB GIS

5.2 LITERARY REVIEW

As a source for the synoptic, repeatable views necessary to monitoring environmental parameters, satellite remote sensing becomes increasingly popular to regional and global-level investigations of vegetation phenology. Multi-temporal data sets of the normalized difference vegetation index (NDVI) are particularly prevalent throughout the literature and continue to be the principal data source for numerous remote sensing-based phenological studies (Pettorelli et al. 2005). The main data used for vegetation analysis are the ones coming from the AVHRR, MODIS and SPOT sensors.

In spite of its prominent usage in phenological research, NDVI time series are prone to noise and error from a number of sources. The main perturbing factors include:

- Atmospheric effects: The actual composition of the atmosphere (in particular with respect to water vapor and aerosols) can significantly affect the measurements made in space. Hence, the effects of aerosols may be misinterpreted if they are not properly taken into account (as is the case when the NDVI is calculated directly on the basis of raw measurements).
- Clouds: Deep (optically thick) clouds may be quite noticeable in satellite imagery and yield characteristic NDVI values that ease their screening. However, thin clouds (such as the ubiquitous cirrus), or small clouds with typical linear dimensions smaller than the diameter of the area actually sampled by the sensors, can significantly contaminate the measurements. Similarly, cloud shadows in areas that appear clear can affect NDVI values and lead to misinterpretations. These considerations are minimized by forming composite images from daily or near-daily

images. Composite NDVI images have led to a large number of new vegetation applications where the NDVI or photosynthetic capacity varies over time.

- Soil effects: Soils tend to darken when wet, so that their reflectance is a direct function of water content. If the spectral response to moistening is not exactly the same in the two spectral bands, the NDVI of an area can appear to change as a result of soil moisture changes (precipitation or evaporation) and not because of vegetation changes.
- Anisotropic effects: All surfaces (whether natural or man-made) reflect light differently in different directions, and this form of anisotropy is generally spectrally dependent, even if the general tendency may be similar in these two spectral bands. As a result, the value of NDVI may depend on the particular anisotropy of the target and on the angular geometry of illumination and observation at the time of the measurements, and hence on the position of the target of interest within the swath of the instrument or the time of passage of the satellite over the site. This is particularly crucial in analyzing AVHRR data since the orbit of the NOAA platforms tended to drift in time. At the same time, the use of composite NDVI images minimizes these considerations and has led to global time series NDVI data sets spanning more than 25 years.
- Spectral effects: Since each sensor has its own characteristics and performances, in particular with respect to the position, width and shape of the spectral bands, a single formula like NDVI yields different results when applied to the measurements acquired by different instruments.

For these reasons, the NDVI should be used with great caution.

For the extraction of the phenological layers, datasets, whose algorithms attempt to account for these complicating factors through processing have been used. They are: MODIS, SPOT, and the new version of AVHRR data, the GIMMS dataset

MODIS, as shown in table 5.1, provide vegetation indices products at different spatial and temporal resolutions. All MODIS NDVI indices are calculated from the reflectance values contained in the daily surface reflectance product (MOD09 series), which are corrected for molecular scattering, ozone absorption, and aerosols (Vermote, et al., 2002). Furthermore the MODIS NDVI compositing methodology is based on data that is quality-filtered, sorted, and processed in a sequence of 6 steps:

- Collect all data from the sixteen days (multiple observations per day).
- Perform integrity checks and keep only non-water pixels.

- Evaluate the quality of the data and assign an internal quality measure, keeping only the highest quality group for compositing.
- Apply the weighted average scheme to reduce data to one observation per day, per orbit per pixel.
- Perform the constrained view angle – maximum value composite (CV-MVC) procedure and choose output observation. The 'two' highest NDVI observations are retained, and the observation with the smallest view angle is then selected.
- Assign quality assurance measures to the output VI product.

TABLE 5.1 MODIS VEGETATION INDICES PRODUCTS

Product name	satellite	format	Spatial resolution	Temporal frequency
MOD13A1	Terra	Tile	500m	16 Day
MOD13A2	Terra	Tile	1000m	16 Day
MOD13A3	Terra	Tile	1000m	Monthly
MOD13C1	Terra	CMG	5600m	16 Day
MOD13C2	Terra	CMG	5600m	Monthly
MOD13Q1	Terra	Tile	250m	16 Day

Different global NDVI datasets derived from AVHRR data have been produced (see table 5.2). In literature, are mainly used the following:

- Pathfinder Land (PAL)
- GVI (Global Vegetation Index)
- GIMMS (Global Inventory Monitoring and Modeling Studies)

TABLE 5.2 MAIN AVHRR VEGETATION INDICES PRODUCTS

Dataset	Temporal span	Temporal resolution	Missing data	Spatial resolution
PAL	July 1981-sept 2001	1 day, 10 days, monthly, seasonal	Sept 1994-January 1995	8 km
GVI	May 1982-present	Weekly, monthly, seasonal	Sept 1994-January 1995	16 km
GIMMS	July 1981-present	15-days	None	8 km

Among the three products the GIMMS dataset have been chosen, since it has the greatest spatial resolution and allows comparisons with the more recent datasets of MODIS and SPOT. In the GIMMS dataset the NOAA AVHRR GAC 1B data are processed (see chapter 2).

The NDVI indices derived from these reflectance products are corrected for different parameters that effects the AVHRR acquisitions. They are:

- residual sensor degradation and sensor inter-calibration differences;
- distortions caused by persistent cloud cover globally;
- solar zenith angle and viewing angle effects due to satellite drift;
- volcanic aerosols;
- missing data in the Northern Hemisphere during winter using interpolation due to high solar zenith angles;
- low signal to noise ratios due to sub-pixel cloud contamination and water vapor.

Furthermore, the GIMMS datasets so corrected are temporally composited by the Maximum Value Composite (MVC) technique, which retains as output value composite the maximum NDVI value. The compositing periods are defined according to the legal calendar: from 1st to 5th, from 16th to the end of each month.

The SPOT VEGETATION dataset is the result of a ten-day synthesis (S10) computed from all SPOT passes on each location acquired during 10 day periods. The periods are defined according to the legal calendar: from 1st to 10th, from 11th to 20th, from 21st to the end of each month. The time composition of SPOT measurement criteria are the followings:

- the measurement does not correspond to a blind or interpolated pixel;
- the measurement is not flagged as cloudy;
- the measurement corresponds to the highest value of NDVI.

Nevertheless it should be recognized that despite the smoothing effects of temporal compositing, residual atmospheric and bidirectional effects often still exist in an NDVI time series, particularly in regions of prevalent cloud cover or at higher latitudes where sun and sensor view angles can be large (Carreiras et al. 2003). For this reason, further noise reduction is often performed on composited NDVI time series.

Many are the techniques used for smoothing out noise in the NDVI time series: in (Jennifer N. Hird, 15 January 2009) there is an accurate work of comparisons of several selected methods (see table 5.3), in which is demonstrated the general superiority of Asymmetric Gaussian (AG) and Double Logistic (DL) function-fitting techniques over four alternative filters. AG and DL showed a balanced ability to reduce noise while maintaining the relevant NDVI signal integrity. These procedures are implemented in the software TIMESAT, a program available from the authors, Jönsson and Eklundh, which allows the correction from residual noises of the data and the extraction of phenological features.

TIMESAT has been used in different applications, such as for mapping environmental and phenological changes in Africa from 1982 till today (Eklundh and Olsson 2003, Hickler et al. 2005, Olsson et al. 2005, Seaquist et al. 2006, Heumann et al. 2007, Seaquist et al. 2009), for improving data in ecosystem classification (Tottrup et al. 2007), for use with MSG SEVIRI data (Stisen et al. 2007), for mapping high-latitude forest phenology (Beck et al. 2007), and to evaluate satellite and climate data-derived indices of fire risk in savanna ecosystems (Verbesselt et al. 2006).

In order to check the suitability of the automated extraction of phenology from satellite image, as a first approach it has been decided to make use of the TIMESAT analysis software, which allows, correcting the residual noise of the data and extracting from satellite NDVI data eleven phenological metrics that include dates of the growing season, such as start, peak, end, and duration, and corresponding vegetation greenness values, among others (Jönsson and Eklundh 2004, 2002).

TABLE 5.3 NOISE REDUCTION TECHNIQUES

Candidate technique	Description
Asymmetrical Gaussian function-fitting	Fits local, nonlinear functions at intervals around local maxima and minima, then merges these into a global function describing the full NDVI time series
Double logistic function-fitting	Uses a series of parameters (e.g. winter, maximum NDVI) to model NDVI time series with a double logistic function
Savitzky–Golay filter	Applies an iterative weighted moving average filter to NDVI time series, with weighting given as a polynomial of a particular degree
4253H, twice filter	Applies a series of running medians of varying window size and a weighted average filter, with re-roughing, to the NDVI time series
Mean-value iteration filter	Iteratively compares each date with the average of the dates before and after it, replacing the date with this average if the difference is above a certain threshold
ARMD3-ARMA5 filter	Applies an autoregressive running median filter with a window size of 3, followed by an autoregressive running mean filter with a window size of five

5.2.1 SELECTED DATASETS

In order to provide a range of datasets to validate for the use in phenological analysis in Bangladesh, products derived from AVHRR, SPOT and MODIS mission have been selected. Their names and characteristics are shown in table:

TABLE 5.4 SELECTED DATASETS

Dataset	Temporal span	Temporal resolution	Spatial resolution	Radiometric resolution	Format
GIMMS	July 1981-present	15-days	8 km	12 BIT	TIFF
MODIS MOD13CMG	February 2000-present	16-days	5,6 km	12 BIT	HDF
SPOT S10	January 1998-present	10-days	1 km	8 BIT	HDF
MODIS MOD13Q1	February 2000-present	16-days	250 m	12 BIT	HDF

With SPOT data a tool for the pre processing of SPOT VEGETATION data is also provided: the VGTExtract, which allows users to extract certain Regions Of Interest (ROIs) from SPOT Vegetation (VGT). These products are distributed by VITO (CTIV) and its partners under license from CNES (Centre National d'Études Spatiales) and SPOT IMAGE. VGT products are usually distributed in the form of zipped archives that consist of one or more images in HDF file format, a text file describing the product's technical properties, a copyright document and a Quick Look Tiff image. VGTExtract only extracts regions of interest from the HDF images. The other product files are unaffected. VGTExtract searches for VGT products in a given directory and its subdirectories. For each product found, VGTExtract can unpack (uncompress) the product and can perform the following actions on the resulting unzipped HDF images:

- Extract a given ROI ;
- Convert image format ;
- Change data type.

VGTExtract has 2 modes of operation:

- GUI or interactive mode (see figure 5.2);
- Batch or automatic mode.

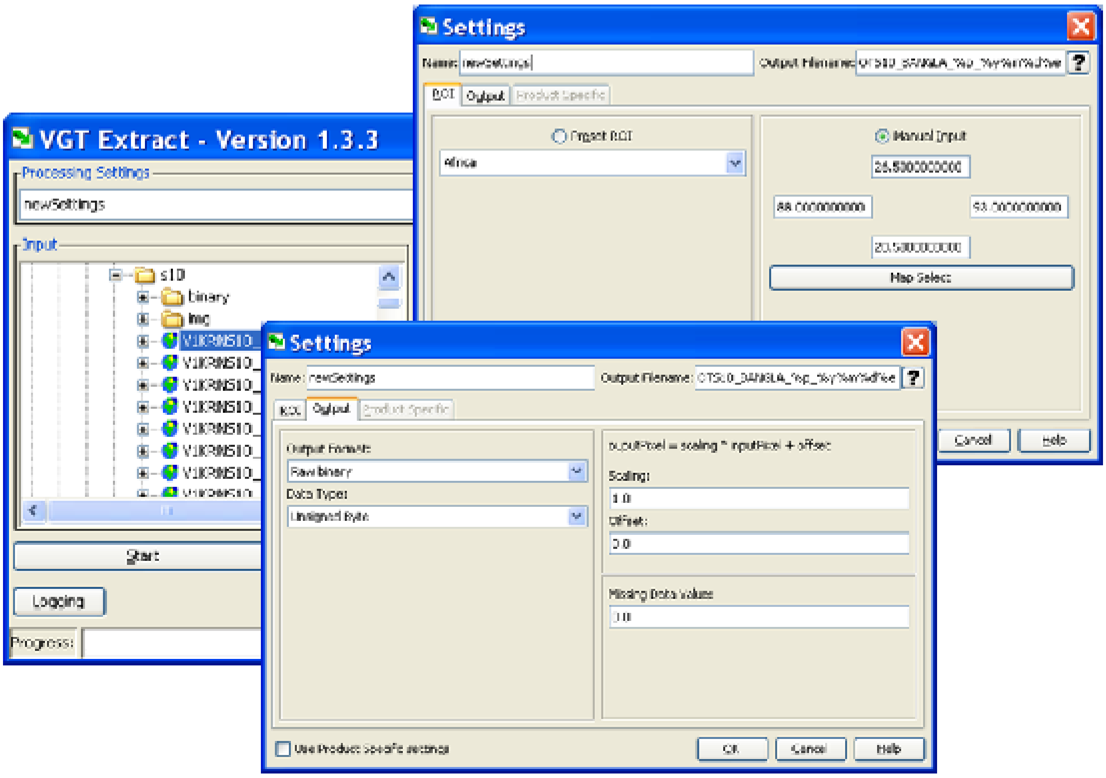

FIGURE 5.2 VGTEXTRACT GRAPHICAL INTERFACE

When run in batch mode, VGTExtract automatically processes a set of VGT products, using saved or pre-defined settings (ROI, type conversion, data type information needed for processing the VGT products). In GUI or interactive mode, the user has the ability to add new settings, change existing settings and save the settings for later use. In GUI mode, the user also can process some VGT products interactively for instance to test his new settings.

5.2.2 TIMESATIMAGE

The software package TIMESAT was developed for estimating growing seasons from satellite time-series, as well as for computing phenological metrics from the data (Jönsson and Eklundh 2002, 2003, 2004, Eklundh and Jönsson 2003). The main functions of TIMESAT are to iteratively fit smooth mathematical functions to time-series of noisy satellite data, and key phenological metrics are extracted for each image pixel.

TIMESAT was originally intended for handling noisy time-series of AVHRR NDVI data and to extract seasonality information from the data. The program now has the capability to handle different types of remotely sensed time-series, e.g. data from Terra/MODIS at different time resolutions.

The Output of the program is a set of files containing seasonality parameters (beginning of season, end of season, amplitude, integrated values, derivatives, see table for their significance and method of extraction), as well as fitted function files containing smooth renditions of the original data.

The function fitting is done in steps. In the first step the number of seasons and their approximate timing is defined. The second step filters the data or fits smooth functions to the data (Savitzky-Golay filter, or least-squares fitted asymmetric Gaussian or double logistic smooth functions). After the fitting has been achieved, the seasonality parameters are computed and written to output files.

TIMESAT runs from within Matlab, or under a Windows command prompt (Fortran version). Matlab is recommended by the authors for tuning the settings and for running small images or time-series, while for large images or long time-series the use of Fortran executables in a command window is recommended.

During the work of thesis, the version of the program (TIMESATimage) to process time series of NDVI binary images has been used. This program requires as input:

- **Name of the sensor image list;**
- **File type for image files**; file type of the sensor and mask data images. (to choose among 8-bit unsigned integer (byte), 16-bit signed integer, and 32-bit real)
- **Number of rows and columns** in the image tile;
- **Subset of the image to process** start row, stop row, start column, stop column;
- **Number of years and number of points per year** in the time-series;
- **Range for sensor data values**; the sensor data values are assumed to be in the range [ylow, yhigh]. Values outside this range are assigned weight zero. This setting can be used to capture obvious spikes in data.
- **Mask image list**: containing information on quality of every pixel of each image of the time series
- **Conversion of mask data to weights**; if mask data are available the user should define three ranges [a1, b1], [a2, b2], [a3, b3] and corresponding weights w1, w2, w3. If a mask data value is in range [ai, bi], i = 1, 2, 3, then the corresponding sensor data value is assigned weight wi. If a mask data value is outside the three ranges the sensor data value is assigned weight zero;
- **Cutoff for amplitude in sensor data**; only time-series with an amplitude, as obtained from the fit to the sine and cosine functions, higher than a certain cutoff are processed. This setting can be used to make sure that uninteresting time-series with low variation, e.g. time-series over deserts or water, are not processed.

- **Cutoff for spike**; spikes are detected by comparing each data value yi, i = 1, 2, . . . , N with median filtered values and with closest neighbors. If the distance is greater than S × ystd, where ystd is the standard deviation for the values in the time-series, we have a spike and the value yi is assigned weight 0.
- **Parameter for determining the number of annual seasons**; the number of annual seasons are determined from the fitted sine and cosine functions. The fitting procedure always gives primary maxima. In addition, secondary maxima may be found. If the amplitude of the secondary maxima exceeds a certain fraction of the amplitude of the primary maxima we have two annual seasons. Setting the fraction close to 0 causes the program to interpret a small depression in the main curve as a second annual season. Setting the fraction close to 1 forces the program to process data as if there were only one annual season. The latter setting should be used for areas where it is known that there is only one annual season.
- **Number of fitting steps**; the user should specify the number of fitting steps (1,2 or 3). If more than one step the fitted functions are forced to the upper envelope of the sensor data.
- **Strength of envelope adaptation**; a value between 1 and 10, where 10 strongly forces the fit to follow the upper envelope of the data.
- **Processing methods**; indicate what processing methods should be used. Savitzky-Golay filter, Asymmetric Gaussians, or Double Logistic functions can be used.
- **Window size**; window size for the Savitzky-Golay filter. A small window captures abrupt changes in the beginning and end of the growing seasons but yields comparatively little smoothing. A large window gives smooth curves at the cost of less accurate time information. Separate window sizes are set for each of the fitting steps.
- **Level that defines the season start and end;** the start and end of the season are defined as the times for which the fitted curves have, respectively, increased and decreased to a user defined level measured from the minimum level.
- **Job name**; give a string that will be used to identify the output data.

All these parameters can be input from a command line or saved in an ASCII file for a batch processing.

The outputs are binary files containing information on the phenological parameter described in figure 5.3

NDVI metric	Significance	Method of calculation
Timing of start of growing season (SOS)	Time at which seasonal photosynthetic activity begins	Composite period in which 10% of the NDVI range between spring minimum and seasonal maximum is reached
Timing of end of season growing (EOS)	Time at which seasonal photosynthetic activity ends	Composite period in which 90% of the NDVI range between autumnal minimum and seasonal maximum is reached
Length of growing season (LGS)	Time period of seasonal photosynthetic activity	Number of composite periods between, and including the SOS and EOS composite periods
Time for the mid of the season	Time at which seasonal photosynthetic activity reaches its maximum	Mean value of the times for which, respectively, the left edge has increased to the 80 % level and the right edge has decreased to the 80 % level.
Maximum NDVI	Maximum level of seasonal photosynthetic activity	Largest data value for the fitted function during the season, maximum NDVI reached during the growing season
Base level	Photosynthetic activity before the start of the season	The average of NDVI value corresponding to the start and end of the season
Seasonal amplitude	Amplitude of seasonal photosynthetic activity, measure of seasonality	Difference between the maximal value and the base level
Rate of increase at the beginning of the season	Rate at which levels of photosynthetic activity increase at the beginning of the season	Ratio between the values evaluated at the season start and at the left 80 % level divided by the corresponding time difference.
Rate of decrease at the end of the season	Rate at which levels of photosynthetic activity decrease at the end of the season	Ratio between the values evaluated at the season end and at the right 80 % level divided by the corresponding time difference.
Large seasonal integral;	Sum or accumulation of seasonal photosynthetic activity, related to biomass	Integral of the function describing the season from the season start to the season end
Small seasonal integral	Seasonal change in photosynthetic activity,	Integral of the difference between the function describing the season and the base level from season start to season end

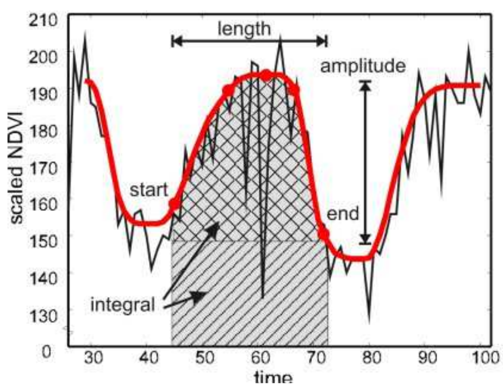

FIGURE 5.3 TIMESAT OUTPUT PARAMETERS

5.3 TESTS AND VALIDATION

In order to allow the validation of the produced dataset and choose the best options for future developments, time series of NDVI distributed at different spatial resolution have been elaborated with the TIMESATimage software. The TIMESATimage outputs produced are then transformed in raster format in order to carry out a validation with feedbacks from the field of analysis.

GIMMS data are characterized by the longest time series, but also by the worst spatial resolution. The output of the elaboration of these data have the purpose to give a synoptic view of the region of interest and allow to identify, with a processing time slightly reduced respect to the other products, homogeneous regions in terms of phenological behavior and to detect (thanks to the long time series) the variation of phenological parameters in the last 25 years.

MODIS data at 2 different spatial resolution have been selected: 250 m and 5,6 km. The first dataset will allow to detect spatial variation of phenological parameter at the biggest scale, and to perform comparisons with the outputs of SPOT data processing, while the one with the lowest spatial resolution will allow comparisons with GIMMS data.

Free SPOT vegetation data, even if they are the less suitable for automated applications, since they need to be downloaded manually and are available only three months after their acquisition, are used as a further element of comparison for historical analysis in a preliminary phase, in order to evaluate the loss of quality in the use of free data respect to data that would have a relevant cost, if distributed in near real time.

In order to choose suitable input parameters to TIMESAT, as suggested by the creator of the programs, some trials have been performed in the TIMESATgui application for some points belonging to the areas of interest. The input parameter that demonstrated to best fit the NDVI time series for pixels belonging to different areas of Bangladesh have been chosen. The main open issue is the choice of the parameter for determining the number of annual seasons, only one value cannot produce produce fitting results with the same quality in the different pixels of the area of interest. Thus the parameter for determining the number of annual seasons was set to the most neutral value, represented by the inter medium value of 0,5 in order to avoid to force the individuation of 1 or more season because.

Other parameters were sets according to the fitting quality that they assured to the time series in different pixels of the image. According to this principle, most of the input parameters were kept to the value advised by the authors, while the number of fitting steps was set to a value of 2, while the window size for the Savitzky Golay filter which resulted to better fit the data in most of the cases was the one (4, 5, 6.), the cutoff for amplitude was set to a value of 0,05 in order to avoid the processing of urban,desert or water areas.

As a further development for improving the fitting, a first step of validation will be performed with the help of WFP Local Offices, which, after the implementation of the developed procedure, will dispose of phenological data and statistic information on them in a suitable format for their interpretation.

After the first steps of validation, further improvements will be performed on the procedures in order to consider in order to assign different weight to the values of NDVI according to quality parameters provided with data.

5.4 IMPLEMENTED PROCEDURE

In order to provide a range of dataset that can be integrated in a web-GIS application, a series of procedures have been developed in order to create raster phenological datasets.

The main operations performed are:

- Download of the data, that has to be performed in an automatic way, in case of necessity of updating of the data;
- Data pre-processing, that is performed in a different way according to the considered datasets. It mainly envisages the processes of reprojection of the dataset into the desired reference system, the resize on the area of interest, the transformation of the input in a format suitable for TIMESAT;
- Launch of TIMESAT in a batch mode;
- Re organization of TIMESAT output phenology files in raster format, which can be easily consulted in a GIS;
- Calculation of statistics for phenological parameters for each season of the year, in order to detect anomalies in the vegetation cycle.

5.4.1 DATA DOWNLOADING

In the data downloading phase, it is necessary to collect all the data on the area of interest from data providers and to check for eventual missing data. The automation of this process is only possible with the use of MODIS and GIMMS datasets. Unfortunately it is not possible to automate this process for SPOT data, to which the access is not allowed in an automatic way. For the SPOT case it has been necessary to manually download the data and locate them in a local folder of the workstation where the system of analysis is located

For MODIS and GIMMS data, it is possible to download the data from an ftp folder by the launch in a batch mode of the WGET program. A scheme of the downloading process is shown in figure 5.4

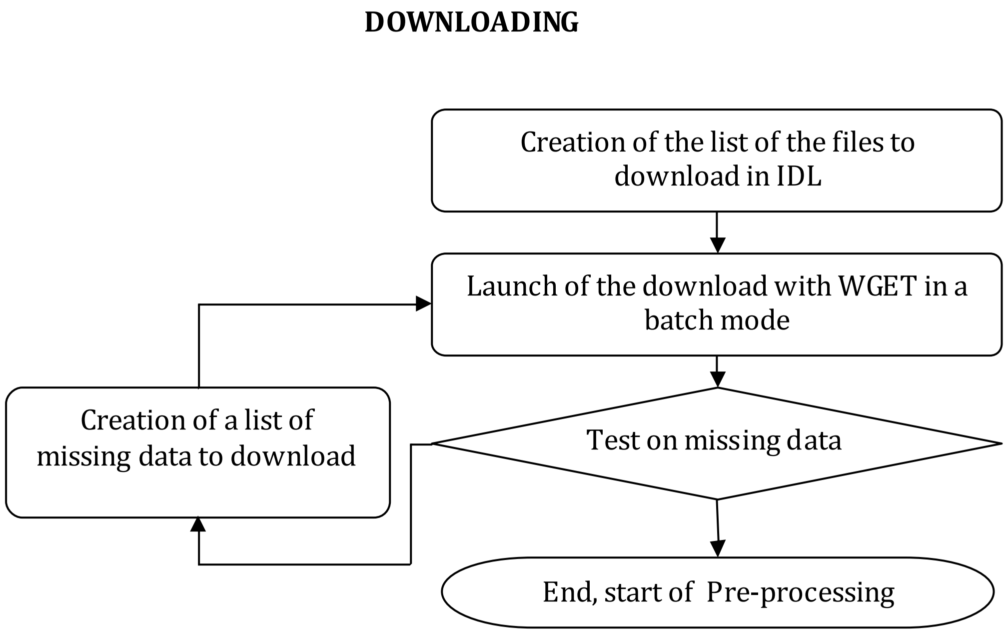

FIGURE 5.4 VEGETATION DATA DOWNLOADING FLOW CHART

5.4.2 DATA PRE-PROCESSING

In this phase data are prepared in order to create the input for the software TIMESATimage. TIMESATimage requires images in a binary format and an input parameter file containing the input parameters values.

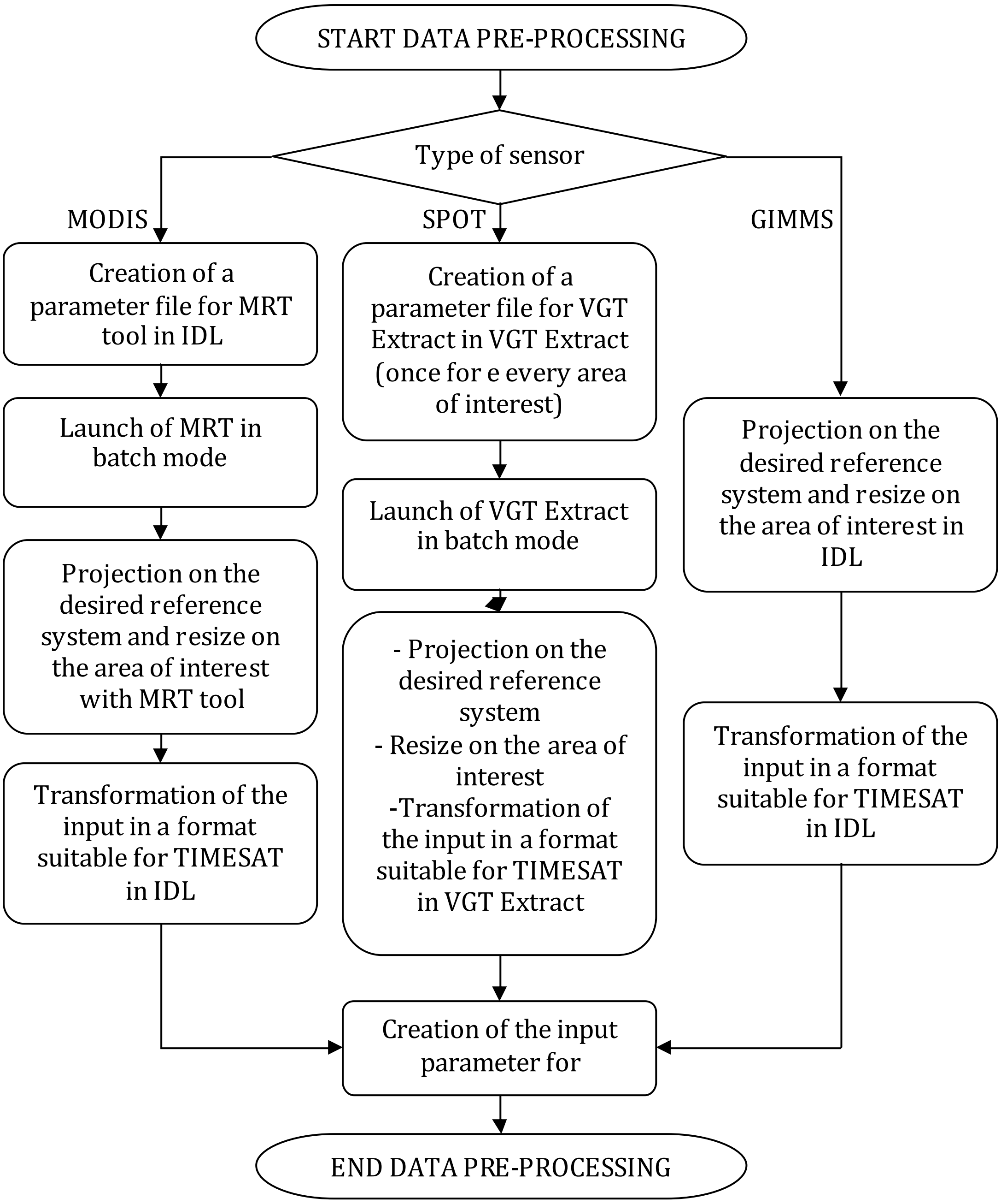

FIGURE 5.5 PREPARATION OF THE DATA FOR THE INPUT IN TIMESAT

Every dataset is provided in a different format and for this reason three different methodologies have been developed in order to prepare the data for the input in TIMESATimage (see figure 5.5 for a schematic view of the elaboration steps and the software used for each processing step). The programs MRT Tool and VGT are called in a batch mode by the IDL main program. The last step of the IDL main program is the preparation of the input files for TIMESAT.

5.4.3 TIMESAT ELABORATION AND CREATION OF RASTER OUTPUT IMAGES

In order to process data the program TIMESATimage has been used. This program elaborates images provided in a binary format and gives as output a binary file containing all the phenological information about the input dataset. The structure of the binary file is displayed in figure 5.6. For each processed pixel are provided the number of seasons and after that, in subsequent lines, the phenological information.

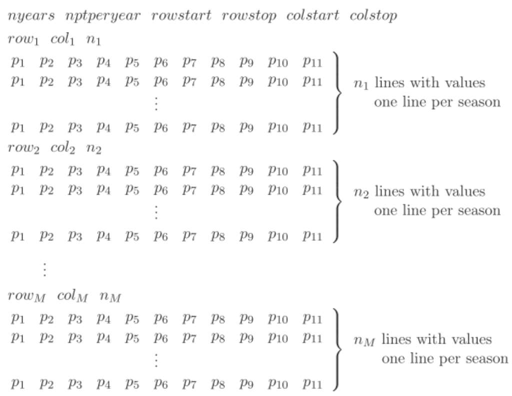

FIGURE 5.6: DATA STRUCTURE OF BINARY FILES CONTAINING SEASONALITY PARAMETERS EXTRACTED FROM THE FITTED FUNCTIONS. THE FIRST LINE OF THE FILE GIVES INFORMATION ABOUT THE NUMBER OF YEARS SPANNED BY THE TIME-SERIES, NYEAR, THE NUMBER OF DATA VALUES PER YEAR, NPTPERYEAR, AND THE SPATIAL EXTENT OF THE AREA. FOR EACH PIXEL (ROW, COL) IN THE AREA SEASONALITY PARAMETERS ARE GIVEN FOR N SEASONS.

The TIMESAT group provides a Fortran program (Phen2img) to extract the information from the output binary file. Phen2img, every time is launched, extracts information for only one phenological parameter and provides as outputs 3 files containing:

- the first season of the selected phenological parameter in the period of interest which is provided as input by the user
- the second season of the selected phenological parameter in the period of interest which is provided as input by the user
- a warning file indicating if some pixels have a number of seasons greater than 2 in the period of interest which is provided as input by the user.

Unfortunately this program cannot be run in a batch mode. Thus, in order to automate and facilitate the output of raster files for long time periods, a routine in IDL has been developed (see figure 5.7). In this routine, the code of the Fortran program has been modified in order to transform the output of TIMESATimage into raster files (one for each season of the period considered, with a number of bands equal to the number of phenological parameters) containing for every pixel the information about all the phenological parameters. In order to render more readable the phenological information, the obtained raster outputs are re-organized. Two types or re-organizations are performed:

- for years and seasons: for each year of the period of interest are created 11 files (one for each phenological parameter) containing the phenological information for each season in the year in exam. This information is useful for the analysis of a year of interest and as intermedium step for the calculation of the statistics.
- for parameter and seasons: : for parameter are created a number of files equal to the number of season, containing, for the season and parameter of interest, all the phenological information for the years of analysis. This information is useful in order to detect anomalies and trends in phenological parameters.

The output files are finally transformed in easily readable tiff files, in which the information is divided in classes (period of year for the start –end and mid of the season, or range of integrals for the small and large integral, rates or NDVI values).

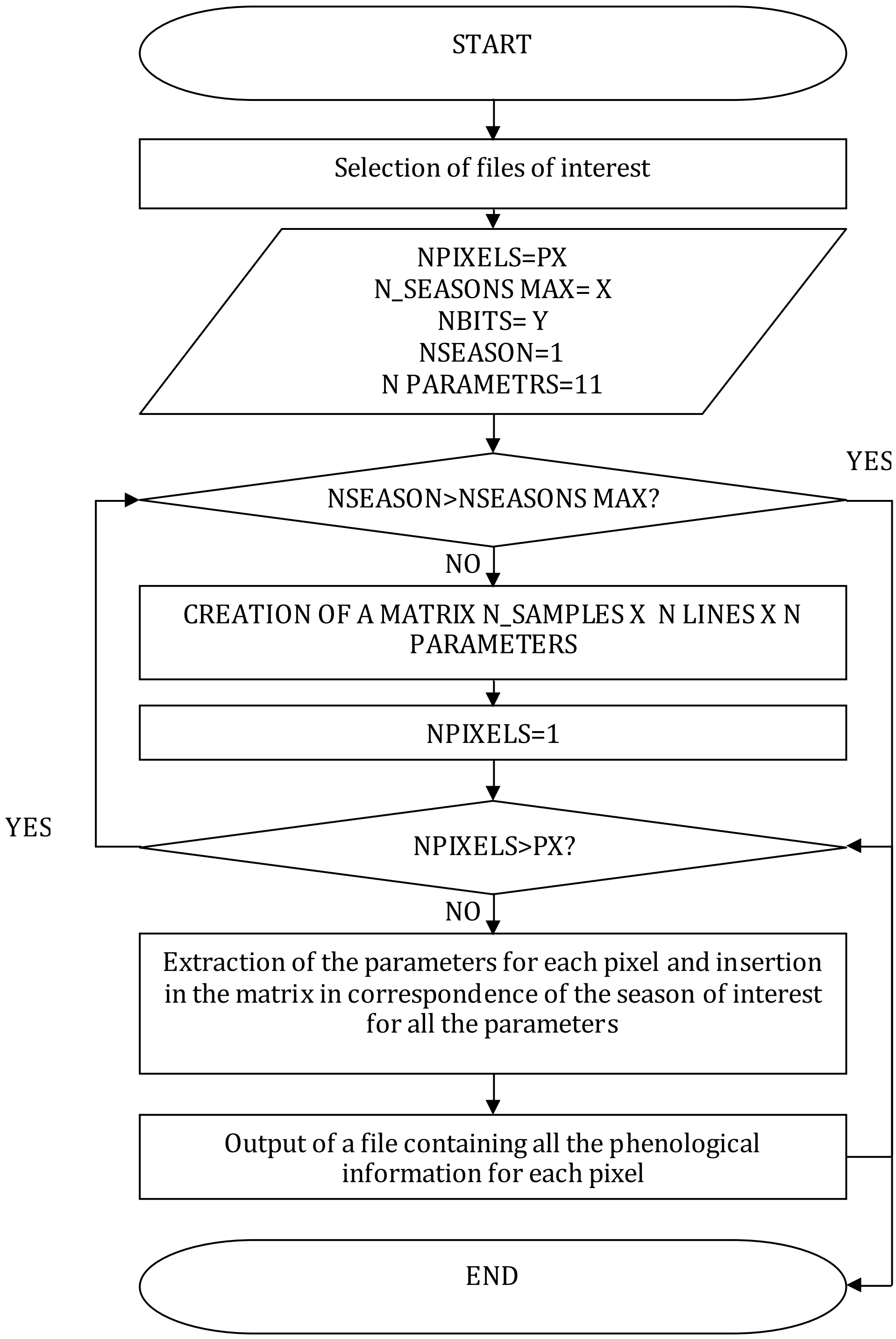

FIGURE 5.7 PROCESSING OF TIMESAT BINARY OUTPUT IN ORDER TO OBTAIN RASTER DATA

5.4.4 AUTOMATION AND PROCESSING TIMES CONSIDERATIONS

For the creation of phenological data the following datasets have been processed:

- 9 years of MODIS 250 m data (197 files, from 2000 to 2009, 2,3 GB)
- 12 years of SPOT data 1 km data (432 files, from 1998 to 2009, 160 MB)

- 9 years of MODIS 5,6 k m data (197 files, from 2000 to 2009, 7,2 MB)
- 25 years of GIMMS 8 km data (600 files, from 1982 to 2006, 5 MB)

The extraction of the phenological parameters required very short time of elaboration for the low resolution data (GIMMS and MODIS MOD13CMG), few hours for the product SPOT at 1 km, while almost a week for the downloading and processing of the product at the spatial resolution of 250 m. after the extraction of long term statistics the processing will require reduced elaboration times compared to the monitoring time scale and can allow to perform an near real time monitoring of vegetation.

CHAPTER 6: APPLICATIONS OF THE PRODUCED DATASET

As mentioned in the previous chapters, two applications for the automated processing of satellite data have been developed in order to create georeferenced thematic information, which can be used as base data for Ithaca activities and that can consulted by WFP during the different phases of risk management. In this chapter applications of the produced datasets are provided.

6.1 APPLICATIONS OF THE WATER BODIES EXTRACTION PROCEDURES

The procedure developed for the extraction of water bodies and the archive of the historical distribution of water bodies can have different applications, that can be divided between the ones for an analysis of the historical data and the ones for near real time data processing in case of flood events.

6.1.1 ANALYSIS OF HISTORICAL WATER BODIES ARCHIVE

In order to obtain information on flooded areas, water bodies' extensions during a flood event are generally compared to a layer of reference water. Since the available global hydrographic references are generally static layers, that do not take into account the variability of water bodies during the year, it has been decided to take advantage of the produced archive in order to detect the variability of water bodies during the year and consequently to create a reference water information for the different periods of the year.

The implementation of the automated procedure for water bodies' extraction on the 10 years of MODIS data brought to the definition of an archive containing different information of water bodies in the last 10 years. They are:

- daily classifications (3 classes: water, clouds and soil) both of Terra and Aqua satellite .
- for every year are available 36 10-days synthesis containing the number of times that a pixel results to be covered by water, clouds or no data in the 10 days taken into consideration. One pixel is defined as covered by water in one day if it results to be covered by water or in Terra or in Aqua image or in both of them. One pixel is defined as covered by clouds/no data if it results to be covered by clouds/no data both in Terra and in Aqua images.
- for every year are available 36 10-days classifications indicating if a pixel is covered by water, clouds or no data. This information is derived from the application of thresholds of days in which a pixel is classified as water, clouds or no data. In the case of water, a pixel is classified as water if for more than 2 days over 10 it resulted to be covered by water, it is classified as cloud/no data if for more than 9 days over 10 it results to be covered by clouds/no data.

Since these outputs are not sufficient for an operational instrument finalized to flood monitoring, statistical information derived from the historical archive was extracted:

a) **Number of times in which a pixel results to be classified as water in the 360 10-day periods** in the ten years reference period (see figure 6.1). The analysis of the histogram in of the image in figure 1 allowed to detect three classes of water cover, information useful to have a complete idea of the evolution of water bodies during the year in one single image (see figure 6.2 and 6.3):
   - Permanent water, defined as the areas that for more than the 70% of the 360 10-days periods in 10 years have been classified as water
   - Seasonal water, defined as the areas that for more than the 20% and for less than the 70% of the 360 10-days periods in 10 years have been classified as water, is the water that have a limited, but almost regular occurrence during the year
   - Historical flooded areas, defined as the areas that for more than the 1% and for less than the 20% of the 360 10-days periods in 10 years have been classified as water (1% is considered the residual error of the procedure), is the water due to flood events.

N° of 10-days periods in which a pixel is classified as water body in 10 years

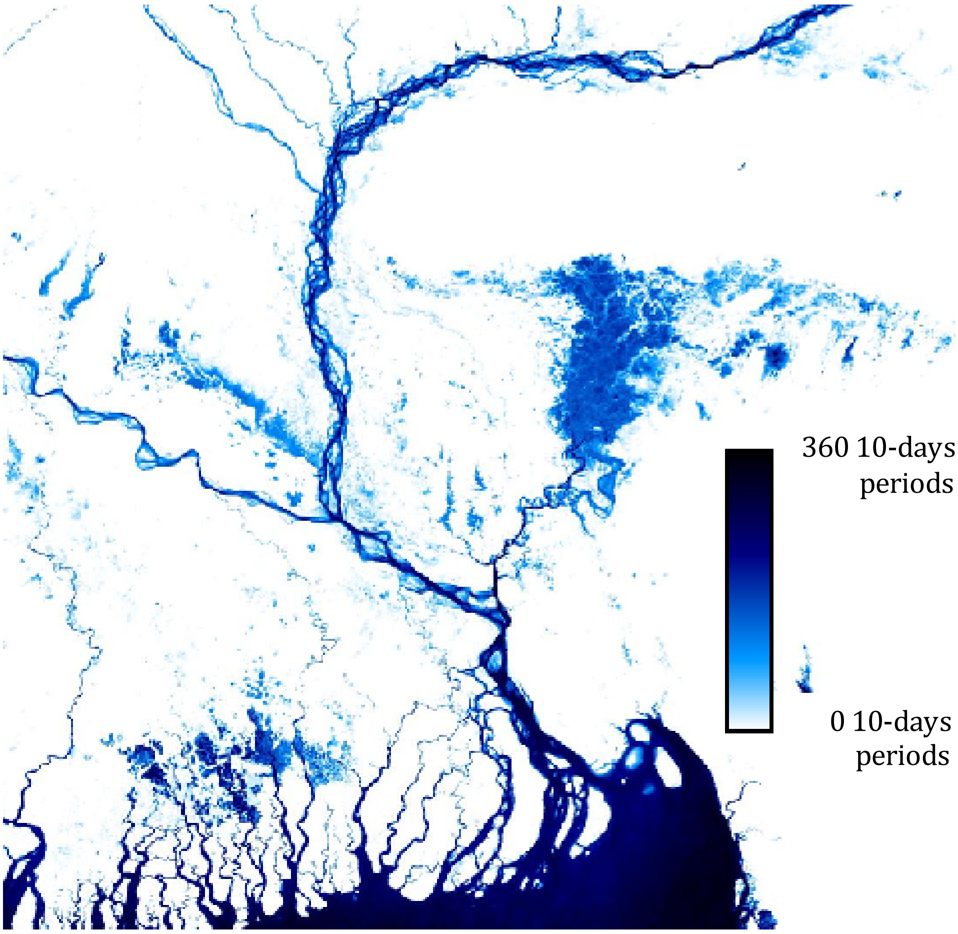

FIGURE 6.1. NUMBER OF TIMES IN WHICH A PIXEL RESULTS TO BE CLASSIFIED AS WATER IN THE 360 10-DAY PERIODS IN THE TEN YEARS REFERENCE PERIOD. IN DARK BLUE THE AREAS WHICH HAVE BEEN CLASSIFIED AS WATER IN MORE 10-DAYS PERIODS

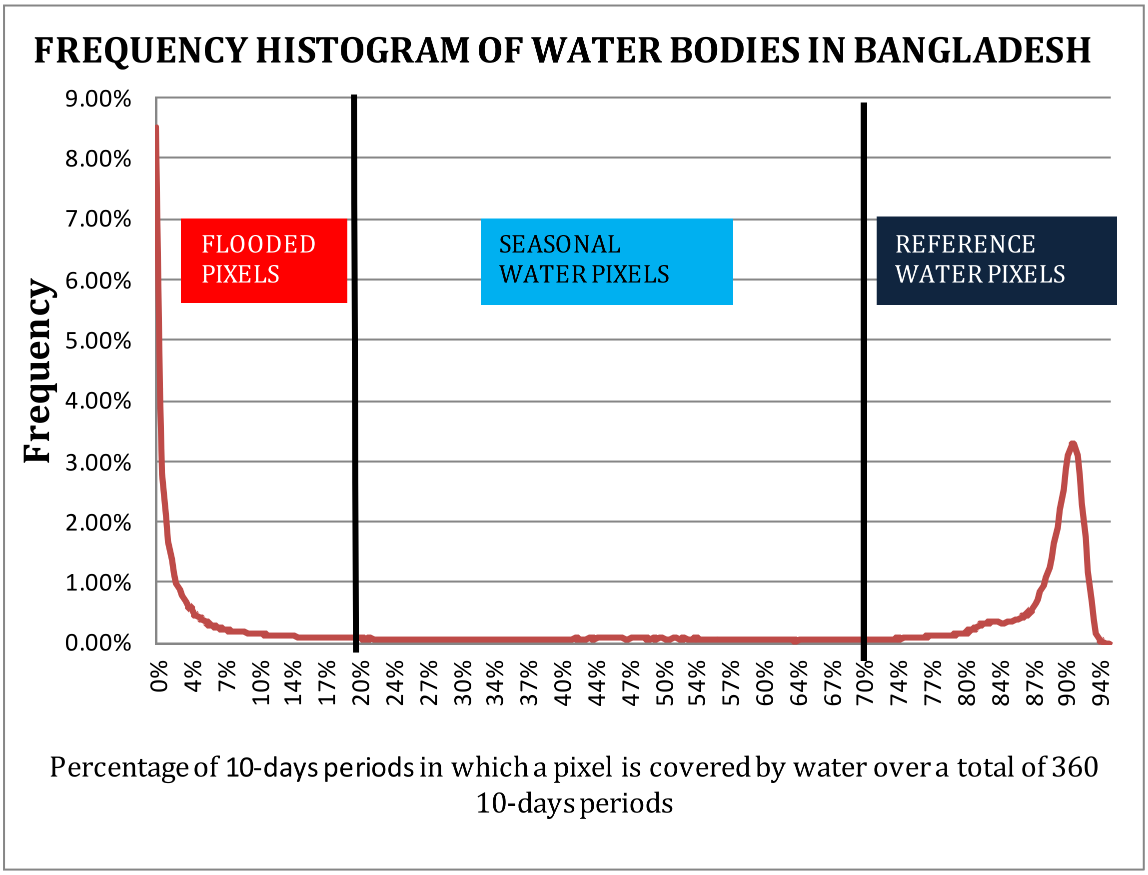

FIGURE 6.2. SLICING OF FREQUENCY HISTOGRAM OF FIGURE 6.1. THE HISTOGRAM IS DIVIDED IN 3 AREAS BY THE PERCENTAGE OF TIMES THAT A PIXEL RESULTS TO BE COVERED BY WATER IN A 10-DAYS PERIOD DURING THE 10 YEAR REFERENCE PERIOD.

Water bodies variability during the year

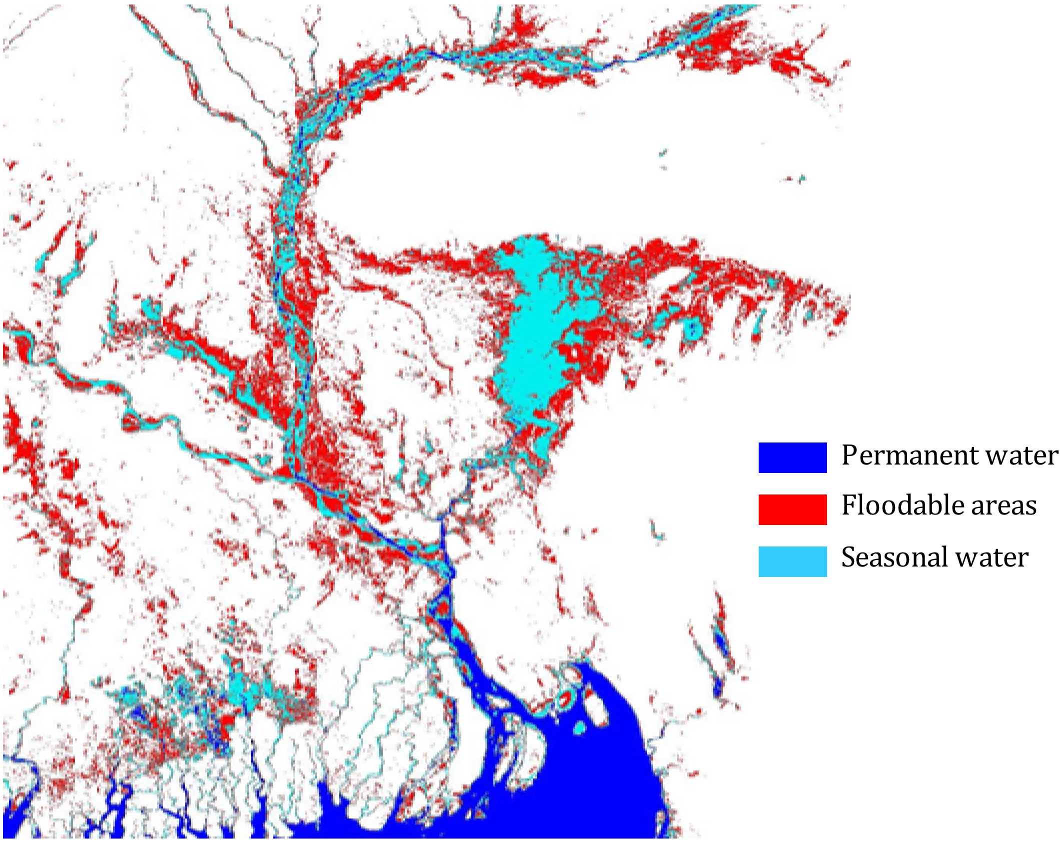

FIGURE 6.3. RESULT OF THE HISTOGRAM SLICING IN FIGURE 6.2 PERMANENT, FLOODABLE AND SEASON WATER ARE CLASSES DERIVED IN BASE OF THE PERCENTAGE OF DECADES IN WHICH A PIXEL RESULTS TO BE COVERED BY WATER IN 10 YEARS

b) **Number of times in which a pixel results to be classified as water during a 10-days period** in the ten years reference period (one image for each 10-days period of the year has been produced). Since a great variability in water bodies has been detected, this information has been used to extract '10-days period reference water': if a pixel results to be covered by water for more than 50% of the same 10-days period of the 10 years, it can be reasonable to define it as reference water, assuming as flooded areas all the pixels that do not result to be covered by water less than the 50% of the years. Therefore all the pixels that for more of the 50% of the 10 years resulted to be covered by water in a 10-days period were considered as reference water for the 10-days period of interest (see figure 6.4).

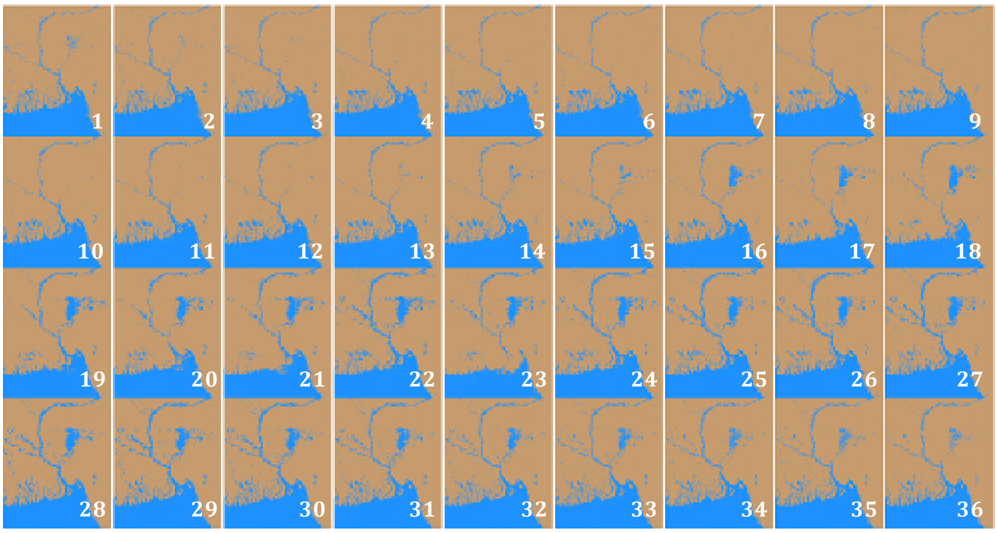

FIGURE 6.4. 10-DAYS PERIOD REFERENCE WATER (IN BLUE) EXTRACTED FROM THE 10 YEARS ARCHIVE OF MODIS CLASSIFIED DATA. THE NUMBER AT THE BOTTOM RIGHT CORNER OF THE IMAGE IS THE 10-DAYS PERIOD NUMBER STARTING FROM THE 1ST

c) **Number of times in which a pixel results to be classified as cloud during** a 10-days period in the ten years reference period (one image for each 10-days period of the year has been produced). This information will be used in the future applications in order to provide for the reference water extracted a sort of 'quality index', in order to be aware of which are the area most prone to cloud coverage in the different pixels of the image.

In order to create an archive of historical flooded areas it will be only necessary to compare the extension of water bodies during flood events and the produced layer of 10-days period reference water. Furthermore, the information of flooded areas can be used as input for the creation of maps indicating the most prone areas to flood in the last ten years to propose to WFP.

In next section the example for the case of the flood that occurred in Bangladesh in June-July 2004 is shown.

6.1.1.1 EXAMPLE OF EXTRACTION OF HISTORICAL FLOODED AREAS

In late-June 2004, heavy monsoon rains swelled the waters of the Meghna River, which reached its peak level in early-July. The Jamuna and Padma Rivers also burst their banks in early-July, due to heavy rains in the north of the country, causing flash floods in the north and the west-central districts. Approximately 38 percent of Bangladesh was inundated by the time the waters began to recede in late-August, including 800,000 hectares of agricultural land.

In order to detect the flooded areas during an historical flood event it only necessary to compare the 10-days classifications referring to 10-days period in which the flood occurred and to the corresponding 10-days period reference water images. The result of this operation for the case of the flood in 2004 is shown in figure 6.5, where the reference water is displayed in blue and the flooded areas, which are the additional water for the last 10-days period of July 2004 (10-days period number 21) respect to the reference, are displayed in red.

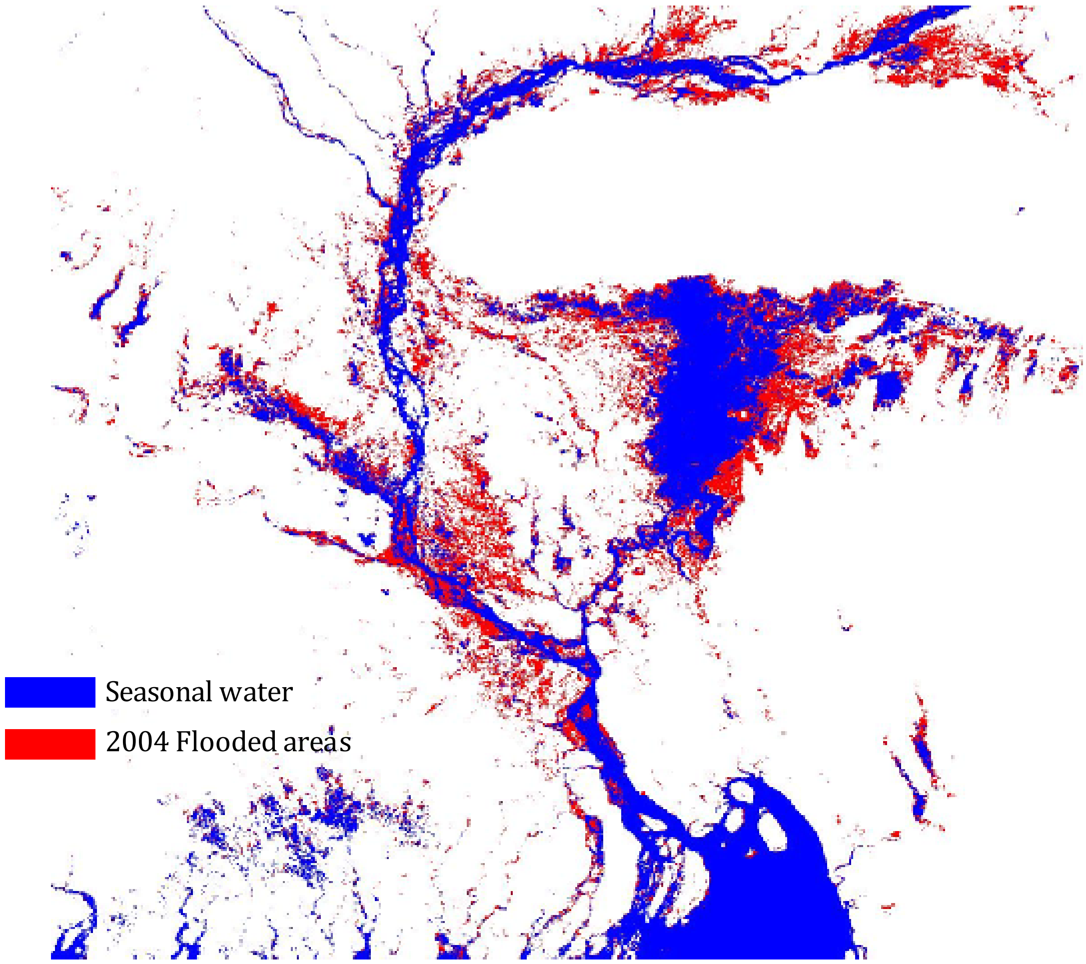

FIGURE 6.5. FLOODED AREAS REPRESENTATION IN OCCASION OF THE FLOOD OF JUNE-AUGUST 2004. THE IMAGE IS REFERRED TO THE SITUATION DURING THE LAST 10-DAYS PERIOD OF JULY.

6.1.2 NEAR REAL TIME DATA PROCESSING IN CASE OF FLOOD EVENTS

The same procedures used for the definition of flooded areas in the historical archive can also run in near real time, in order to provide a support for the rapid mapping activity during an emergency due to a flood event. In this case it will be necessary to create in near real time the layer containing the water bodies' extension during a flood event, using the available satellite imagery of the MODIS Rapid Response System belonging to the most recent 10 days and compare it with the 10-days period reference water. In this case the indication of the number of days in which a pixel is covered by clouds is fundamental, in order to be aware of the areas in which no information is available (see figure 6.6).

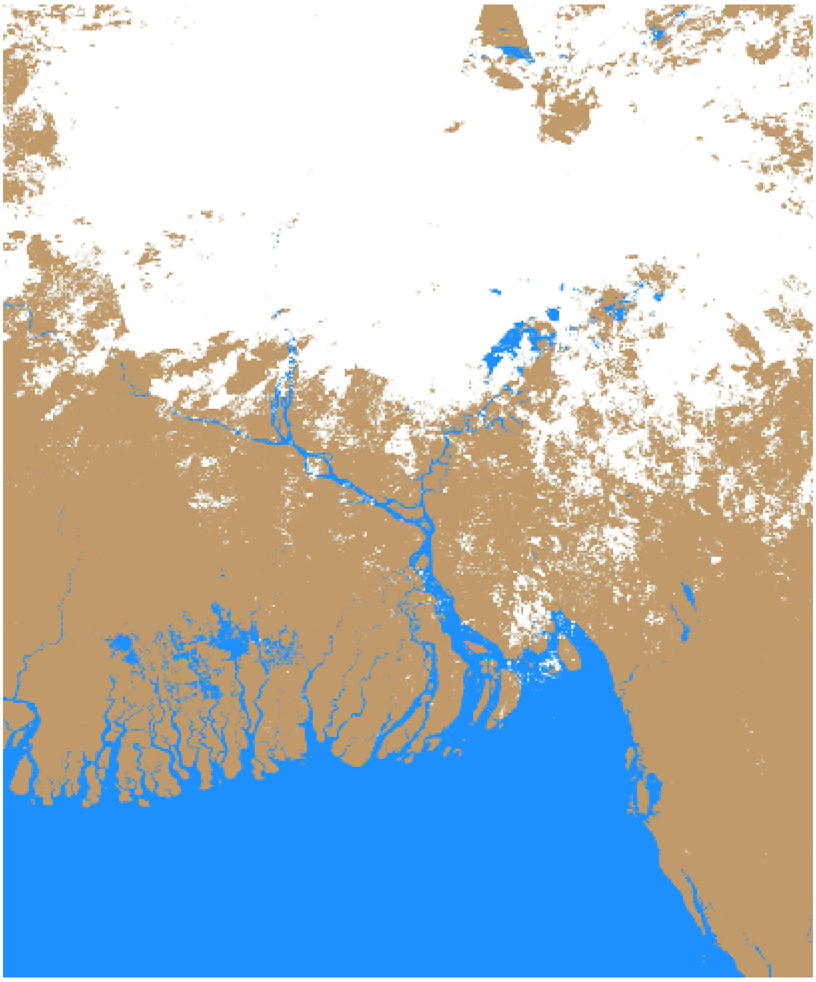

FIGURE 6.6. EXAMPLE OF IMAGE INDICATING THE AREAS FOR WHICH NO INFORMATION IS AVAILABLE IN THE 10-DAYS PERIOD OF INTEREST (IN WHITE). IN BLUE THE DETECTED WATER BODIES

The information derived from the classification of the near real time image can be inserted in a map with additional information on the administrative districts, roads, center of transportation, main cities in order to provide a first view of the situation during a flood few hours after the flood event has occurred, and of the involved population (see figure 6.7)

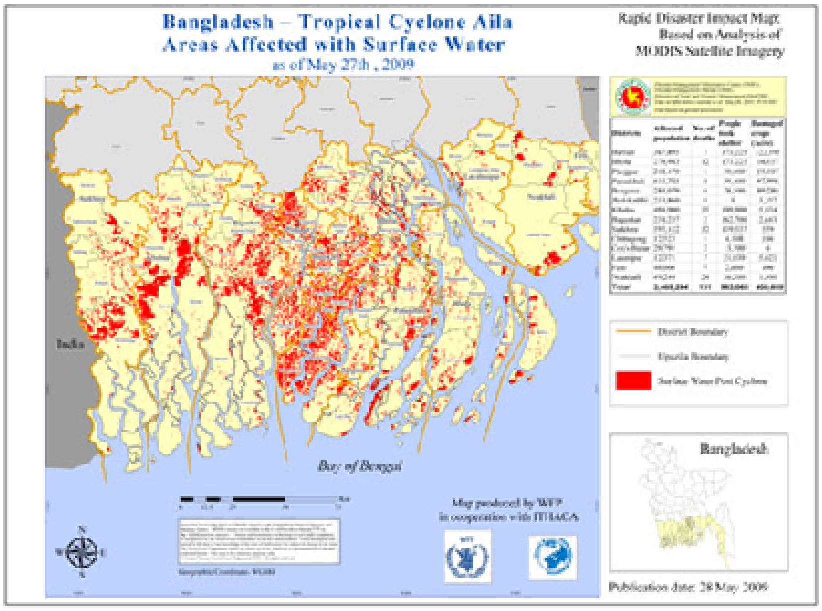

FIGURE 6.7. EXAMPLE OF RAPID MAPPING AFTER A FLOOD EVENT, INDICATING FLOODED AREAS AND INVOLVED POPULATION.

The water information derived from the real time application can also be integrated in monitoring applications requiring a coarser spatial resolution. An example is shown in figure 6.8, where the information on flooded areas has been converted into a flood index.

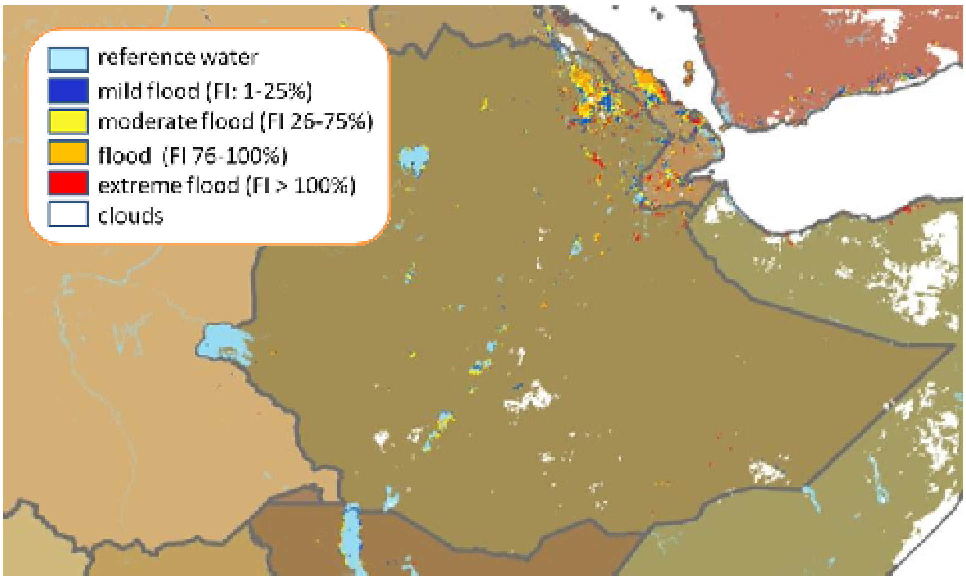

FIGURE 6.8. EXAMPLE OF A FLOOD INDEX FOR THE AREA OF ETHIOPIA, AT A 3KM SPATIAL RESOLUTION.

6.2 PHENOLOGICAL PARAMETERS

Phenological parameters extracted with the procedures described in chapter 5 could provide useful information on vegetation condition. The data can be used both for historical analysis and for near real time applications (updating the available information every 10 or 15 days, according to the base data compositing period). Since global land cover data do not provide information on the seasonality for the data of interest, it will be necessary a validation by asking for a feedback of the proposed datasets from WFP Local Offices. Waiting for these feedbacks, at the moment is only possible to illustrate which kind of thematic map can be from processed datasets. The first outputs of the procedure are raster images containing information on seasonal parameters. The statistical analysis allows extracting representative parameters values (for example mean, standard deviations) for each season of the year. Anomalies can be detected from this medium behavior, which can allow identifying areas that have been subjected to famines.

From a first analysis, data at the lowest spatial resolution (GIMMS at 8 km, MOD09CMG at 5.6 km and SPOT at 1 km) provide a general idea of the distribution of vegetation phenology parameters and will result useful for the analysis of larger areas respect to the one of Bangladesh, that can be easily covered by two tiles of MODIS products at the spatial resolution of 250 m.

An example of outputs is reported in figures 6.9-6.15

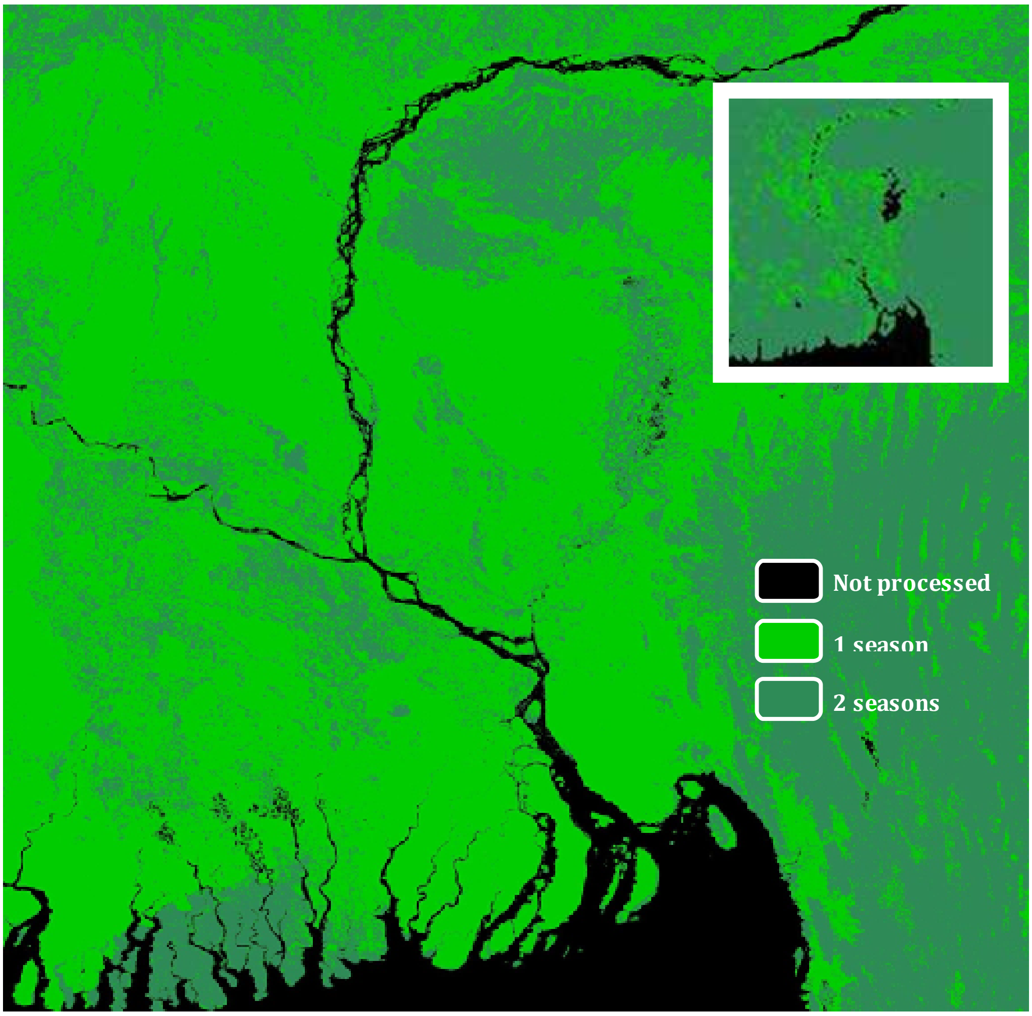

FIGURE 6.9. EXAMPLE OF RASTER DATA PROVIDING THE MEDIUM NUMBER OF SEASON PER YEAR. IT IS CALCULATED BY THE MEAN OF THE EXTRACTED NUMBER OF SEASON FOR EACH YEAR DERIVED FROM MODIS DATA AT 250 M. ON THE UPPER RIGHT CORNER OF THE IMAGE THE SAME LAYER, EXTRACTED FROM THE MODIS PRODUCT AT A SPATIAL RESOLUTION OF 5.6 KM, IS PROVIDED.

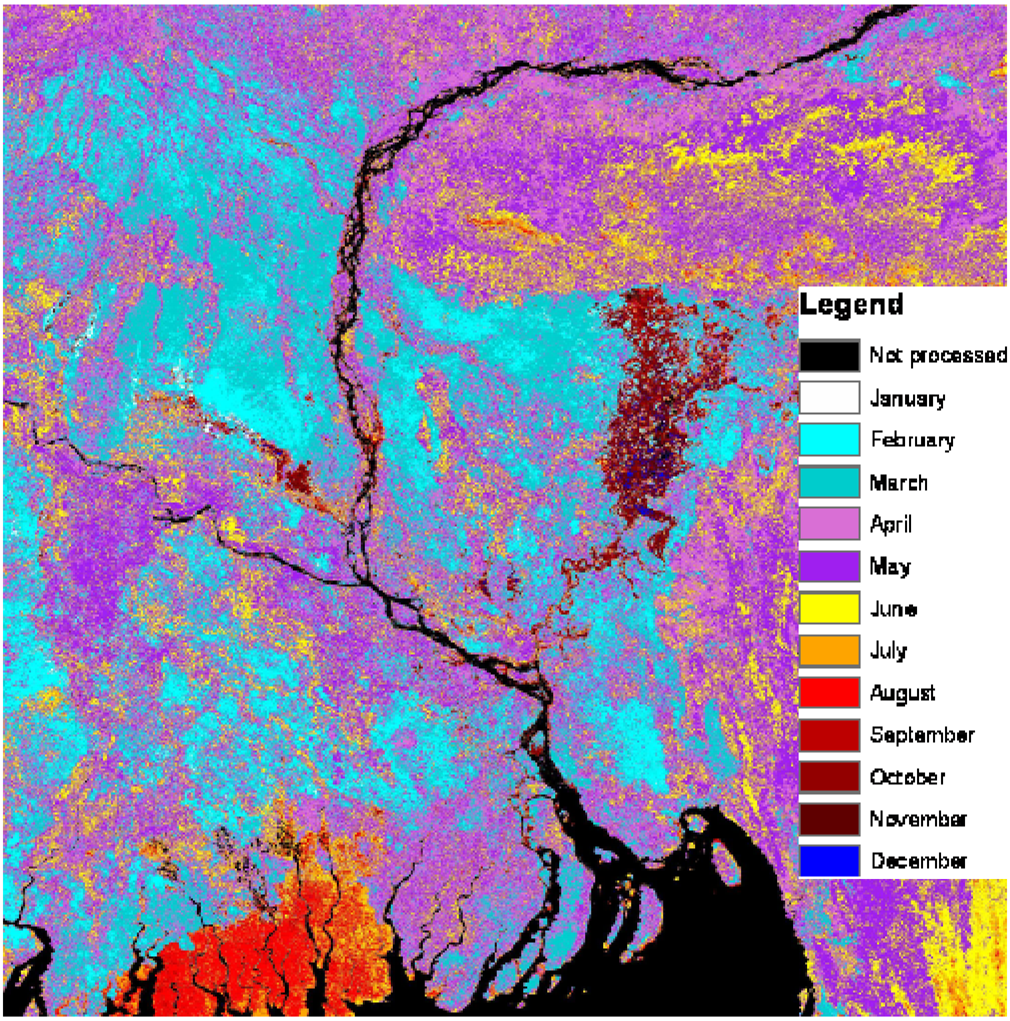

FIGURE 6.10. START OF THE FIRST SEASON OF THE YEAR MEDIUM VALUES FOR THE 10 YEARS COMPOSITE OF MODIS DATA. THIS PARAMETER ALLOWS TO RECOGNIZE THE DIFFERENT CULTIVATION REGIONS THAT IS NOT POSSIBLE TO DETECT BY THE ONLY USE OF THE LANDCOVER .

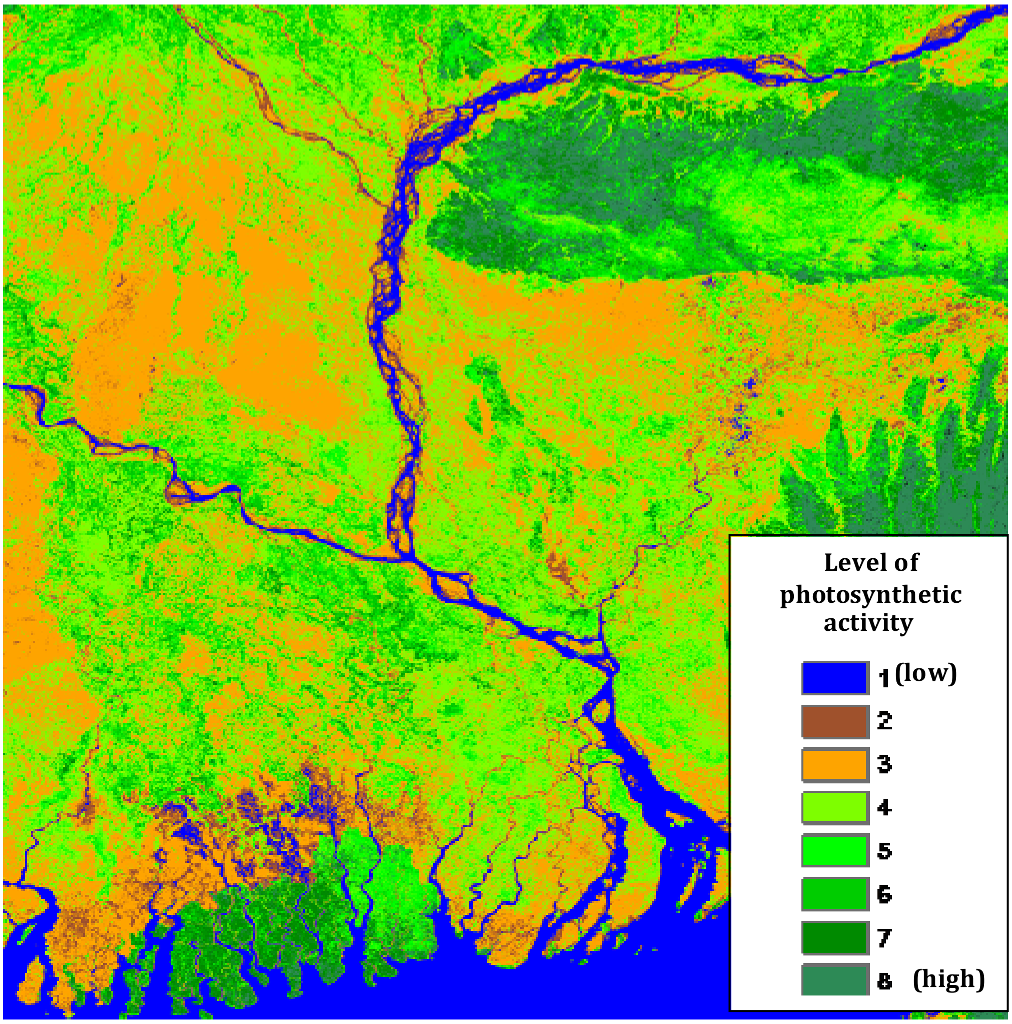

FIGURE 6.11. THE MEAN OF LARGE INTEGRAL OF THE FIRST SEASON OF THE YEAR DERIVED FROM MODIS DATA AT THE RESOLUTION OF 250 M. IN DARK GREEN THE REGIONS WITH THE HIGHER PHOTOSYINTHETIC ACTIVITY WHILE IN BLUE THE ONES THAT HAVE A VERY LIMITED OR EQUAL TO ZERO ONE. IN LEGEND ARE SHOWN 8 LEVELS OF PHOTOSYNTHETICAL ACTIVITY.

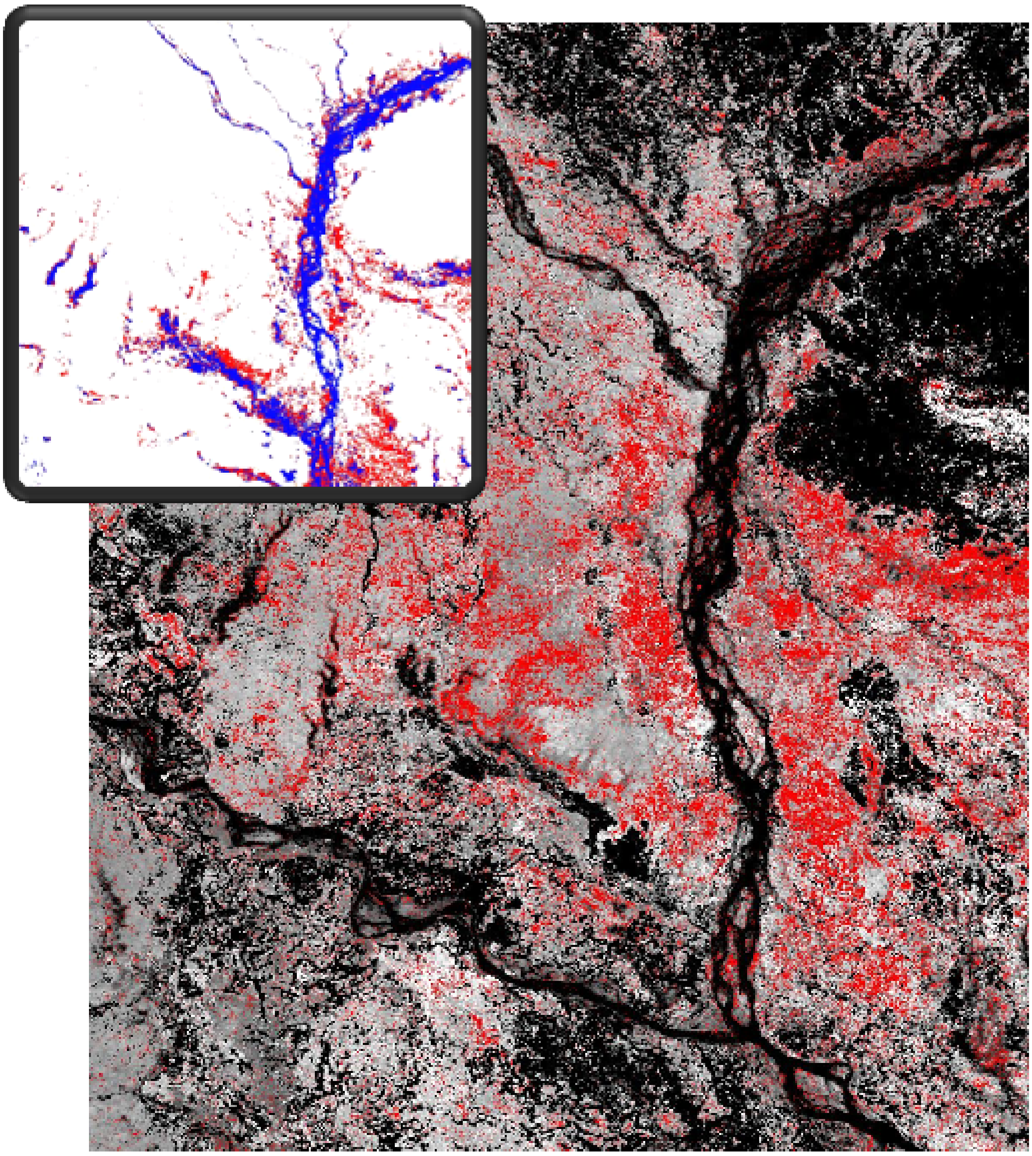

FIGURE 6.12. ANOMALIES FROM THE MEDIUM VALUE OF THE LARGE INTEGRAL OF THE SECOND SEASON IN THE 2004, AFTER THE FLOOD WHICH INVOLVED A GREAT PART OF BANGLADESH. IN LEVEL OF GRAY, ASCENDING FROM BLACK TO WHITE THE MEDIUM VALUES OF LARGE INTEGRAL, THE ANOMALIES, DISPLAYED IN RED, COULD INDICATE THE AREAS IN WHICH THE VEGETATION HAS HAD LESS PHOTOSYNTHETIC ACTIVITY BECAUSE OF THE FLOODS. ON THE UPPER LEFT CORNER OF THE FIGUREOF FLOODED AREAS (RED).

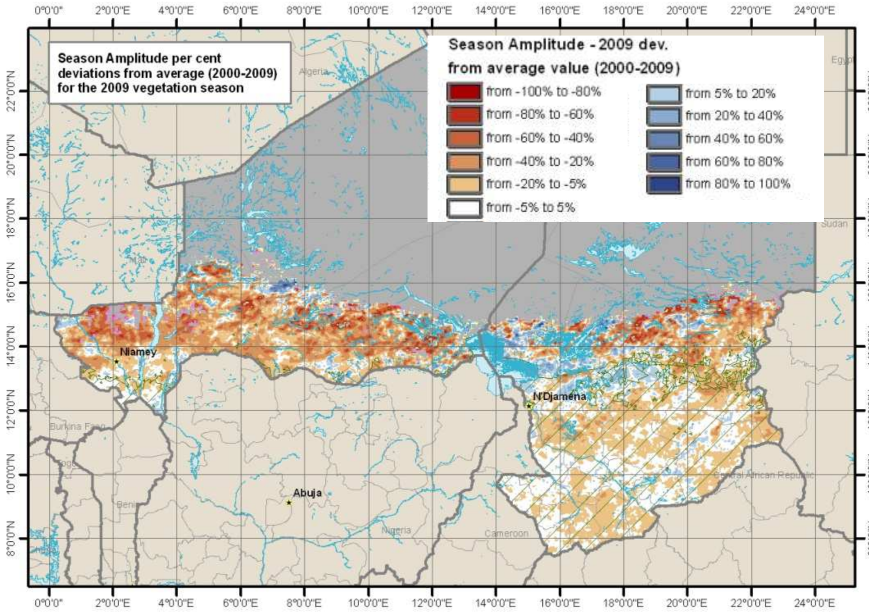

FIGURE 6.13. EXTRACT OF A MAP PRODUCED WITH THE DEVELOPED PROCEDURE FOR THE 2010 DROUGHT IN NIGER AND CHAD. IT REPRESENTS THE ANOMALIES FROM THE MEDIUM VALUE OF THE SEASONAL AMPLITUDE OF THE 2009 VEGETATION SEASON IN NIGER AND CHAD.

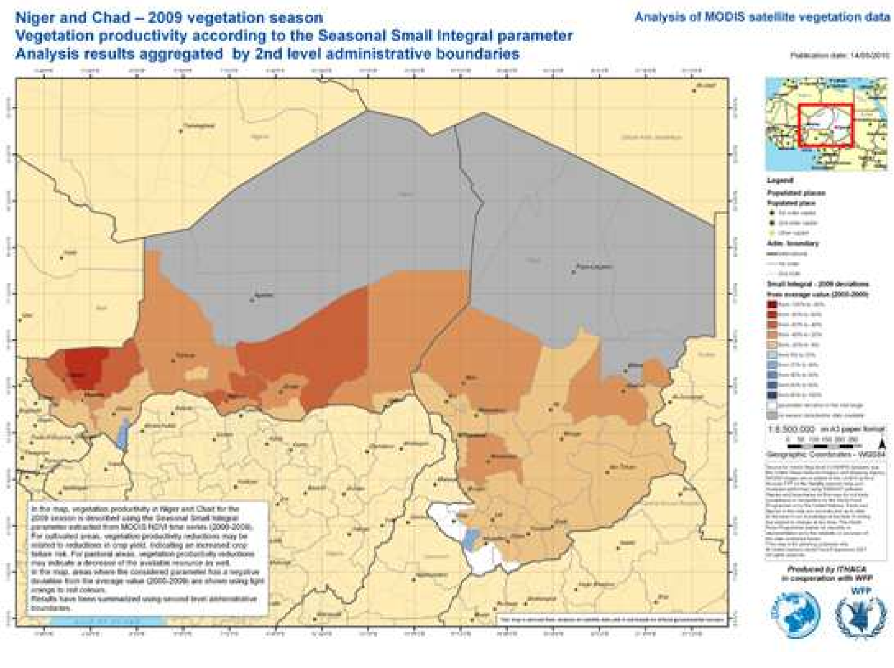

FIGURE 6.14. EXAMPLE OF MAP PRODICED WITH THE DEVELOPED PROCEDURE FOR THE 2010 DROUGHT IN NIGER AND CHAD. IT REPRESENTS THE ADMINISTRATIVE BOUNDARIES AGGREGATION OF THE ANOMALIES FROM THE MEDIUM VALUE OF THE SMALL INTEGRAL OF THE 2009 VEGETATION SEASON IN NIGER AND CHAD.

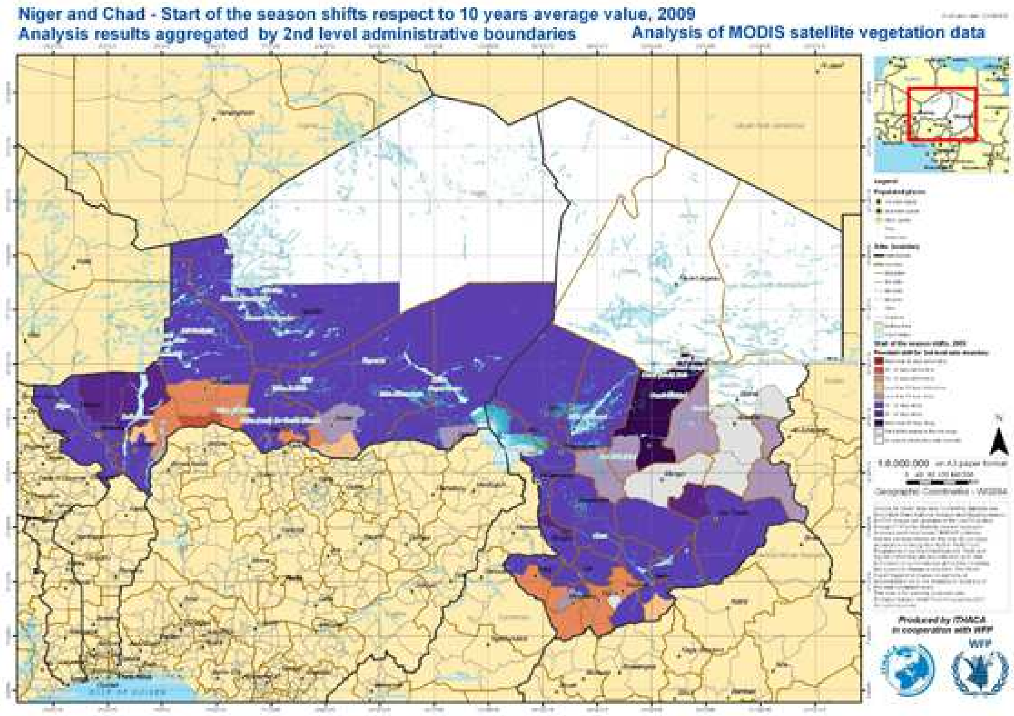

FIGURE 6.15. EXAMPLE OF MAP PRODUCED WITH THE DEVELOPED PROCEDURE FOR THE 2010 DROUGHT IN NIGER AND CHAD. IT REPRESENTS THE ADMINISTRATIVE BOUNDARIES AGGREGATION OF THE ADVANCES AND DELAYS OF THE START OF THE SEASON FROM THE AVERAGE BEHAVIOR OF THE 2009 VEGETATION SEASON IN NIGER AND CHAD.

6.3 FUTURE DEVELOPMENTS

The archive of flooded areas will be soon tested for its use in Ithaca early warning system for floods. According to their period of belonging, flooded areas can be associated to past rainfall anomalies and will allow to produce scenarios according to the registered rainfall value during a flood event caught by the early warning system.

Furthermore after the creation of the archive on flooded historical flooded areas in the region of Bangladesh it will be useful to distribute this information in a web-GIS application that enable the user to choose how to calculate reference water or flooded areas. In this work of thesis compositing periods and day thresholds for classifying a pixel as water bodies have been selected by the author. In future applications it will be possible to visualize flooded areas according to different reference water that can be obtained by grouping historical analysis in 10-days_periods, months, seasons or years. This tool will enable WFP to perform vulnerability analysis.

The procedures for the extraction of phenological parameters will be integrated in a web-GIS application, in order to allow WFP Local Offices to dispose of the produced layers and provide feedback. After the validation, layers of phenological parameter will be combined to the Ithaca drought monitoring system.

Finally the approach for the detection of flooded areas and the instrument for the extraction of phenological parameters are going to be tested also in other countries with different morphological features, in order to test the effective flexibility of the developed applications.

CONCLUSIONS

The purpose of this work of thesis was the development of two applications with the maximum degree of automation, that allow regional scale analysis of the environmental parameters in order to support the WFP food aid activities in developing countries. In order to fulfill this aim, the main steps performed are:

- WFP and Ithaca need assessment and definition of necessary procedures and final products ;
- literary review of the procedures already existing;
- selection of suitable data processing procedures and algorithms; their improvement and integration in final automated applications proper to the production of thematic georeferenced information;
- test of the developed procedures and presentation of their potentials in an area particularly prone to natural disaster, such as Bangladesh.

Particular attention has been paid to two main activities:

- development of procedures proper to the monitoring and historical analysis of the extension of water bodies and floodable areas;
- development of procedures proper to the creation of historical archive and near real time monitoring of thematic georeferenced information about vegetation phenology.

The first conducted activity brought to the development of an historical archive of water bodies extension. A new approach to water bodies' classification has been proposed that showed encouraging results on the Bangladesh area. As a matter of fact, the developed algorithm, based on the compositing of daily classified images, obtained using IR and NDVI thresholds, exhibited producer and user accuracies of the order of 80%, with a complete degree of automation during the processing phase. To correctly work, the algorithm only needs as inputs selection of proper threshold values and the time compositing period. The values of their inputs can be defined in a first preliminary phase of analysis performed using sample images. Mainly, the developed algorithm allows, using proper data processing techniques, to generate an archive containing the extension of flooded water bodies using the whole collection of MODIS data (available from 2000 to present). Starting from this, user can extract flooded areas extension for selected events of interest in a very short time, allowing the support of near real time applications. Furthermore the proposed procedure enables the creation of archive of water bodies that are suitable for different applications: the identification and evaluation of the seasonality of water bodies, the extraction of reference water, support of near real time Ithaca activities devoted to rapid mapping based on MODIS data of areas hit by floods.

If from one side the developed system for the detection of flooded areas can be considered a flexible instrument to adopt in different region of Earth, it is also necessary to consider that the effectiveness of the system is strictly related to the spatial resolution of the product used as input, the 250 m of the MODIS surface reflectance dataset and, for this reason, flood events which present an effect on the soil smaller than the spatial resolution of the data cannot be detected. Afterwards the system requires long processing times for the phase of creation of the archive, disadvantage that could be easily solved by the use of a grid computing system.

The second developed activity, concerning the production of phenological information on vegetation, brought to the definition of an automated procedure which is able to process different datasets. The procedure automatically manage the data downloading operations, the input and output processing and the definition of the final products in the format required by the user. Bangladesh area has been the test site of the implemented procedure. Since it is not possible to reconstruct the seasonality of vegetation in a per-pixel base using land cover data and crop calendars available on line, in order to obtain reference information for a future validation, information will be collected on the ground, with the help of Bangladesh WFP local offices. The processing times are encouraging also for a future use, after validation, of the dataset in a near real time system of drought monitoring.

REFERENCES

CHAPTER 1

Ajmar, A., Boccardo, P., Disabato, F., Giulio Tonolo, F., Perez, F., & Sartori, G. (2008). Early impact procedures for flood events February 2007 Mozambique flood. *RIVISTA ITALIANA DI TELERILEVAMENTO, ISSN: 1129-8596* , pp. 65-77.

Department of Regional Development and Environment, .., Executive Secretariat for Economic Social Affairs, .., & Office of Foreign Disaster Assistance US Agency, f. I. (1991). *Primer on Natural Hazard Management in Integrated Regional Development Planning.* Retrieved January 2, 2010, from Sito Web OAS: http://www.oas.org/usde/publications/unit/oea66e/begin.htm#Contents

HORTON, A. (2008). EARLY WARNING AND EARLY IMPACT ANALYSIS: WFP PREPAREDNESS IN ACTION. In .. EUROPEAN COMMISSION, *FROM EARLY WARNING TO EARLY ACTION?The debate on the enhancement of the EU's Crisis Response capability continues* (pp. 325-328). EUROPEAN COMMISSION EXTERNAL RELATIONS DIRECTORATE-GENERAL.

Rodriguez, J., Vos, F., Below, R., & Guha-Sapir, D. (2009). *Annual Disaster Statistical Review 2008. The numbers and trends .*

CHAPTER 2

http://MODIS.gsfc.nasa.gov/

HDF5 User's Guide: http://www.hdfgroup.org/HDF5/doc/UG/index.html

http://www.nsof.class.noaa.gov/saa/products/welcome http://www.ngdc.noaa.gov/ecosys/cdroms/AVHRR97_d1/AVHRR3.htm

http://wdc.dlr.de/sensors/AVHRR/index.php#NDVI

http://eoweb.dlr.de:8080/short_guide/D-NDVI.html

CHAPTER 4

A.C.T. Pinheiroa, ,. J. (2007, February 15). Near-real time retrievals of land surface temperature within the MODIS Rapid Response System. *Remote Sensing of Environment* , pp. Volume 106, Issue 3, 326-336.

HU Zhuowei, G. H. (August 28-29, 2007). FAST FLOODING INFORMATION EXTRACTION IN EMERGENCY RESPONSE OF FLOOD DISASTER. *ISPRS Workshop on Updating Geo-spatial Databases with Imagery.* Urumchi, Xingjiang, China.

Liam Gumley, J. D. (2010, Jan 14). *Creating Reprojected True Color.* Retrieved Jan 20, 2010, from rapidfire.sci.gsfc.nasa.gov.

Ludmila Samoilenko, L. K. INFORMATION TECHNOLOGY OF FLOOD MONITORING USING REMOTE SENSING DATA.

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, 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*.

CHAPTER 5

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.

Beck et al., 2007 P. Beck, P. Jönsson, K. Høgda, S. Karlsen, L. Eklundh and A. Skidmore, A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula, *International Journal of Remote Sensing 28 (2007), pp. 4311–4330.*

Carreiras, J. M. B., Pereira, J. M. C., Shimabukuro, Y. E. and Stroppiana, D., 2003, Evaluation of compositing algorithms over the Brazilian Amazon using SPOT-4 VEGETATION data. *International Journal of Remote Sensing, 24, 3427-3440.*

Chen et al., 2004 J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita and L. Eklundh, A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter, *Remote Sensing of Environment 91 (2004), pp. 332–334.*

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.

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.

Eklundh, L. and Jönsson, P., 2003, Extracting Information about Vegetation Seasons in Africa from Pathfinder AVHRR NDVI Imagery using Temporal Filtering and Least-Squares Fits to Asymmetric Gaussian Functions. In *Image and Signal Processing for Remote Sensing VIII*. Proceedings of SPIE Vol 4885, edited by Serpico. S.S. Society of Photo-Optical Instrumentation Engineers, pp. 215-225

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.

Filipova-Racheva and Hall-Beyer, 2000 D. Filipova-Racheva and M. Hall-Beyer, Smoothing of NDVI time series curves for monitoring of vegetation changes in time, *Ecological Monitoring and Assessment Network National Science Meeting 2000*, Toronto, Ontario, Canada, January 17–22, 2000. (2000) Available online at: http://www.eman-rese.ca/eman/reports/meetings/national2000/toc_posters.html.

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.

Heumann, B.W., Seaquist, J.W., Eklundh, L. and Jönsson, P., 2007, AVHRR Derived Phenological Change in the Sahel and Soudan, Africa, 1982 - 2005,Remote Sensing of Environment, 108, 385-392.

Hickler, T., Eklundh, L., Seaquist, J., Smith, B., Ardö, J., Olsson, L., Sykes, M. and Sjöström, M., 2005, Precipitation controls Sahel greening trend.Geophysical Research Letters, 32, L21415.

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).

Jennifer N. Hird, G. J. (15 January 2009). Noise reduction of NDVI time series: An empirical comparison of selected techniques, *Remote Sensing of Environment, Volume 113, Issue 1* , 248-258.

Jönsson and Eklundh, 2002 P. Jönsson and L. Eklundh, Seasonality extraction by function-fitting to time series of satellite sensor data, *IEEE Transactions on Geoscience and Remote Sensing 40 (2002), pp. 1824–1832*

Jönsson and Eklundh, 2004 P. Jönsson and L. Eklundh, TIMESAT — A program for analyzing time-series of satellite sensor data, *Computers and Geoscience 30 (2004), pp. 833–84*

Jönsson and Eklundh, 2006 P. Jönsson and L. Eklundh, TIMESAT — A program for analyzing time-series of satellite sensor data: User's Guide for TIMESAT 2.3, Malmö and Lund, Sweden (2006), pp. 3–10.

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. HYPERLINK "http://www.sciencedirect.com/science/journal/00344257" *Remote Sensing of Environment* , HYPERLINK "http://www.sciencedirect.com/science?_ob=PublicationURL&_tockey=%23TOC%235824%232008%23998879987%23700473%23FLA%23&_cdi=5824&_pubType=J&view=c&_auth=y&_acct=C000058882&_version=1&_urlVersion=0&_userid=2823018&md5=95186b29323c08b022bf090cd9cf45da" Volume 112, Issue 12 , pp. 4167-4185.

Ma and Veroustraete, 2006 M. Ma and F. Veroustraete, Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China, *Advances in Space Research 37 (2006), pp. 835–840.*

Olsson, L., Eklundh, L. and Ardö, J., 2005, A recent greening of the Sahel—trends, patterns and potential causes, Journal of Arid Environments, 63, 556-566.

Pettorelli, N., Vik, J. O., Mystrerud, A., Caillard, J. M., Tucker, C.J. and Stenseth, N. C., 2005, Using the satellite-derived NDVI to assess ecological responses to environmental change. *TRENDS in Ecology and Evolution, 20, 503-510*

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.

Seaquist, J. W., Hickler, T., Eklundh, L., Ardö, J. and Heumann, B., 2009, Disentangling the effects of climate and people on sahel vegetation dynamics. *Biogeosciences*, 6, 469-477.

Seaquist, J.W., Olsson, L., Ardö, J. and Eklundh, L., 2006, Broad-scale increase in NPP Quantified for the African Sahel, 1982-1999, *International Journal of Remote Sensing*, 27, 5115-5122.

Stisen, S., Sandholt, I., Norgaard, A., Fensholt, R. and Eklundh, L., 2007, Estimation of diurnal air temperature using MSG SEVIRI data in West Africa,*Remote Sensing of Environment*, 110, 262- 274.

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.

Tottrup, C., Schultz Rasmussen, M., Eklundh, L. and Jönsson, P., 2007, Using 250-meter spatial resolution MODIS data and regression tree modeling to map fractional land cover across the highlands of mainland Southeast Asia, *International Journal of Remote Sensing*, 28, 23-46.

Velleman, 1980 P. Velleman, Definition and comparison of robust nonlinear data smoothing algorithms, *Journal of the American Statistical Association 75(1980), pp. 609–615*

Verbesselt, J., Jönsson, P., Lhermitte, S., van Aardt, J. and Coppin, P., 2006, Evaluating satellite and climate data derived indices as fire risk indicators in savanna ecosystems. *IEEE transactions of Geoscience and Remote Sensing*, 44, 1622.

Vermote, E. F., El Saleous, N. Z. and Justice, C.O., 2002, Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing ofEnvironment,83, 97-111*

Vermote E. F. and Vermeulen A., Algorithm Technical Background Document: ATMOSPHERIC CORRECTION ALGORITHM: SPECTRAL REFLECTANCES (MOD09), Version 4.0, April 1999

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*.