

INTEGRATION AND EXPLOITATION OF EARTH OBSERVATION IMAGERY IN AN OPEN SOURCE BASED GEOGRAPHIC INFORMATION SYSTEM (GIS) TO SUPPORT LITTORAL OPERATIONS

Dr. Peter H. Ranelli, Dr Charles Trees, Dr Farid Askari⁽¹⁾, Ing. Alessio Fabiani, Ing. Simone Giannecchini, Dr. Raffaele Grasso, Ing. Giuliana Pennucci, Ing. Francesco Spina, Mr Roger Stenvoll, Dr Elena McCarthy, Ms Gisella Baldasserini, Ms Ferda Turgutcan, Dr. Michel Rixen

NURC, a NATO Research Centre,
Viale San Bartolomeo 400, 19126 La Spezia, ITALY,
Email: ranelli@nurc.nato.int

⁽¹⁾ Now at Satellite Fusion, San Diego, CA, USA

ABSTRACT

A general net-centric architecture for supporting distributed data fusion and operational oceanography is proposed and discussed. The architecture allows the development of end to end processing systems, automating the data flow from data acquisition to data and products dissemination and visualisation. An integrated approach fusing remotely sensed data, forecast models and *in situ* observations is used for a precise characterisation of the littoral environment and to provide support to operations over the area of interest. Automatic decision aids is employed to assess the impact of environmental conditions on operations. Furthermore, data management and dissemination is provided through Web GIS services which are compliant with standards for interoperability. An example of system implementation in the context of landing operation support is described and discussed.

1. INTRODUCTION

Recent events (Indian Ocean tsunami, Hurricane Katrina) have demonstrated the need to conduct a rapid environmental assessment of littoral regions and use it to support operations. Planning, support and execution of civilian and military operations requires precise and timely, highly resolved 4-D picture of oceanographic, meteorological and geographic conditions. The littoral zone is among the most difficult regions to provide this support due to highly variable and rapidly changing conditions at this interface between sea, air and land.

This paper presents an overview of an end-to-end process (see fig. 1) being developed at the NATO Undersea Research Centre (NURC) for collecting and combing data from remote sensing systems and *in-situ* sensors with computer models to characterize and predict environmental conditions. Remote sensing data from geostationary and polar orbiting satellites are critical, since it can be collected quickly with a large coverage area and will be a principal source of data over

an operational area. The assimilation process provides accurate nowcasts and forecasts of the complex air-sea-land system in the littoral zone on the necessary spatial and temporal resolutions for operational use. The final and critical step is to integrate this data with decision support aids to provide an assessment of potential impact on operations, such as disaster relief and humanitarian operations.

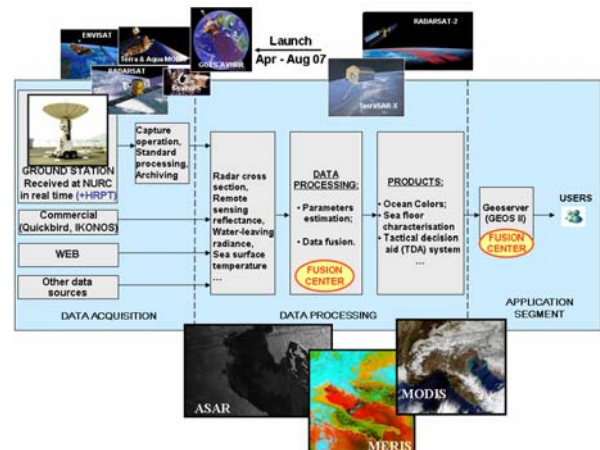


Figure 1: end to end processing system

One of the requirements for a successful end-to-end system is robust data services to provide the many functions needed to manage, store and process data and to operate user applications for decision support. The NURC system is based on an open source client-server Geographic Information System, using an enterprise Java application and a Geospatial database, to store, convert, fuse and deliver the above-mentioned data sets. Applications are run on this server to deliver the tactical decision support products from data stored in the database. Results of this system in both experimental context and operational demonstrations will be presented.

The flexibility of the proposed architecture is achieved thanks to the use of protocols for distributing complex processing chains among several nodes connected in a network. The processing chains are designed following a producer-consumer paradigm in order to automate the data flow. Moreover, the use of ISO and Open GIS Consortium (OGC) standard protocols [1] for the dissemination of data through Web GIS services allows sharing geospatial data and products in an interoperable way among heterogynous units. Automatic decision making systems are employed in order to fuse environmental information from several sources and provide an assessment of the impact on operations. The use of fuzzy logic decision aids gives to the system the needed flexibility to operate in a variety of operational contexts and the capability of efficiently dealing with the data uncertainties and vagueness.

The paper is organized as follows. In section 2 the architecture supporting end to end distributed data fusion is introduced and described in more details. These include the support for distributed processing based on the SOAP Web Service framework and the producer-consumer paradigm, the automatic ingestion engine to assimilate data and products into the geospatial server and the OGC/ISO geospatial services used to disseminate results through the Web. Section 3 is devoted to describe a tactical decision aid (TDA) system based on this architecture for supporting landing operations. This section describes the processing chain used to predict high resolution environmental parameters over the landing area, the fuzzy logic expert system used for automatic inference and the presentation stage based on custom Web applications compliant with the OGC/ISO standards. The conclusions and lines for future work are traced in section 4.

2. DISTRIBUTED DATA FUSION

Architectures for supporting distributed network-centric geospatial data fusion have to integrate data acquisition facilities, processing resources, automatic data analysis processes, higher level reasoning and automatic decision making systems, and geospatial information managing and dissemination platforms into a common framework. Standard service and communication protocols have to be exploited in order to exchange data and information in an interoperable way. Moreover, a high level of automation is needed, going from the data sources to the presentation and visualization stage, in order to process large sets of dynamic and spatially distributed data without human intervention.

Supporting operational oceanography means managing data from variety of sources in a wide range of formats. Common approach is to download the data and then convert it into a single commonly accepted format with some loss of information. This clearly becomes inconvenient and cumbersome as soon as the amount of

data grows. Many times several toolkits are required to allow scientists and developers to perform complex operations.

The solution is represented by the creation of a client-server architecture based on well-accepted protocols and interfaces which will guarantee interoperability both at toolkit level as well as at service level.

The project developed an open source, OGC/ISO TC211 compliant GIS layered architecture for the compilation of the Recognized Environmental Picture (REP). The REP provides a unified picture of the geographic, meteorological and oceanographic operating environments. It is the result of the fusion of a variety of data sets: static and dynamic, acquired by remote sensors or by fixed or mobile sensors or generated by models.

Fig. 2 depicts the architecture to support operation planning. The architecture was designed in order to achieve the above mentioned requirements and provide to the end user a timely and effective data access interface through standard geospatial services and custom client-server applications. The complexity of the system behind the data access interface is totally hidden thanks to the separation between data sources, processing capabilities, geospatial services and application layers.

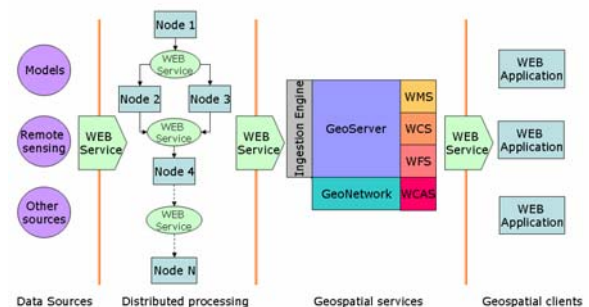


Figure 2: data fusion architecture.

The system makes use of methods based on the producer-consumer paradigm and on network protocols like SOAP-Web Services, to achieve the following goals: 1) build complex distributed processing systems chaining environmental forecast models and processing algorithm for remote sensing observations and in situ data, 2) connecting the data sources to the inputs of the processing chains, 3) providing a link between processing system outputs and data sources, and the GIS web services and applications. Full automation is achieved by using an automatic ingestion engine to assimilate the data and the results in the geospatial server. The geospatial server is based on the open source GeoServer application, which implements a Web Map Service (WMS), a Web Feature Service (WFS) and a Web Coverage Service (WCS) following the specifications provided by the OGC/ISO standards for interoperability. GeoNetwork is a GIS application

developed by FAO implementing a Web Catalog Service (WCaS) which allows the user to search data sets through queries on the associated metadata. Future works will be devoted in integrating this application with GeoServer in order to provide the system with improved search capabilities.

The Java object programming language was chosen for the implementation of the proposed architecture in order to allow the portability of the system.

The following sections will provide an overview of the basic components of the system i.e. the support for distributed processing, the automatic ingestion engine and the OGC geospatial services.

2.1. Support for distributed processing

The connection between heterogeneous processing nodes is based on the producer-consumer paradigm and on protocols for distributing processing chains on the network.

The system prototype implements a modular architecture composed of basic units in cascade. Each unit wraps a particular process in a structure which is devoted to ingest input data and propagates intermediate results, when available, to the input of the next processing node with the suitable format. The processing chain is triggered by the presence of a new input data set following a sort of polling mechanism of the directory storing the data. Each node in the chain is triggered as well by polling the directory storing the output results of the previous node.

The input and output directories can be placed on different physical processing units and the data transfer between them is assured by using the SOAP Web Service framework or similar protocols.

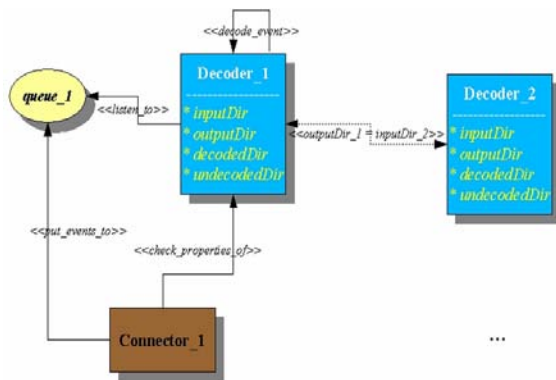


Figure 3: Processing System Classes Interaction Diagram

Three fundamental objects are used to implement a processing chain based on the producer-consumer paradigm: the decoder, the connector and the connector chain (see fig. 3). These objects are implemented, depending on the particular task they have to accomplish, through a set of Java classes derived from the basic abstract classes ModuleDecoder,

ModuleConnector, and ModuleConnectorChain. The three basic classes provide a standard interface for implementing a complex processing system composed of several sub-processes which have to be synchronized in order to share data and processing resources.

Objects of the ModuleDecoder class can be instantiated on different threads (see fig. 3) so that they can run independently from one another. They are the physical units implementing a processing node which has to produce “something” starting from a set of input data. The state of a decoder object can be changed following an event-based logic.

The ModuleDecoder class member *queue* is a first-in first-out (FIFO) buffer where events are stored. A ModuleDecoder object is normally under a “wait mode” of operation. The operative system, the Java Virtual Machine (JVM) in our case, puts the decoder object in the “execution mode” when an event is available from the queue. During this phase, a process is executed several times depending on the number of events stored in the queue. The *inputDir* and *outputDir* members are the places from which the decoder takes the data to process and stores the results, respectively. An event could be generated, for example, when a new data set is available from the *inputDir* source. The objects derived from the ModuleConnector abstract class generate the events to be stored in the event buffer. For each decoder there is a connector looking at the *inputDir* source and producing an event when a new data set is available. The ModuleConnectorChain class allows the user to set up a processing chain composed of several decoder-connector couples. In particular, it provides methods for initializing the properties of each decoder and setting the *outputDir* of a node as the *inputDir* of the next node so as the input data and the results of the processing blocks can be propagated along the chain. This class includes methods for monitoring the processing chain (execution time statistics and current state of a particular node, for instance) and for sending broadcast command to all the nodes of the chain in order to stop, for example, the execution of the system threads.

2.2. Automatic Ingestion Engine

An effective data fusion system has to be designed to deal with data sources having the following characteristics:

- new data are produced in a timely fashion,
- new data produced are generated in separate flow scheme, where a flow represents a pipeline of data with common characteristics.

Such data sources should be managed and configured in a way that reduces operator interventions in order to achieve superior automation and to decrease the potential for errors. The role of an ingestion engine would be to semi-automatically recognize, configure and ingest different flows of data in order to avoid human intervention.

The ingestion engine we foresee would be a separate module of GeoServer and it will be designed and implemented in such a way that it would not impact overall GeoServer performances, as they will run on the same machine.

2.3. OGC geospatial services

This section focuses on describing with more detail the functionalities of the OGC services, namely WMS [2], WFS [3], WCS [4] and WCaS [5], depicted in the Data and Metadata Dissemination Tiers in fig. 4. The needed level of data abstraction is provided by the Data Management Tier which is based on the concepts of Feature [6] and Coverage [7]. These are, in practice, the abstract entities used to represent geospatial vector and raster data, respectively. The services can be described dealing with these concepts without regard of the physical nature of the originating data sources.

Interactions between clients and servers take place on an HTTP or HTTPS link, by employing XML based protocols in order to leverage technologies and protocols which are freely available and widely supported. This allows us to focus attention on the definitions of protocols and architectures which strictly deal with geospatial data and metadata dissemination issues.

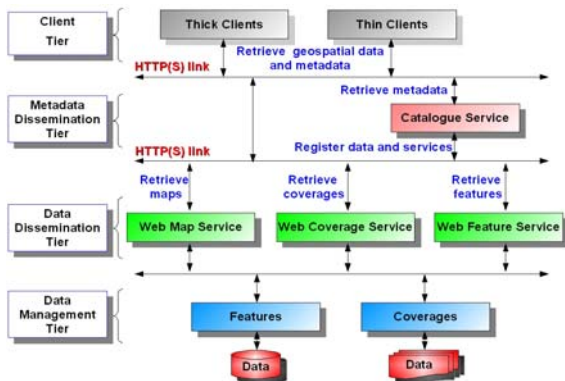


Figure 4: Core geospatial services for supporting distributed geospatial data fusion.

A Web Mapping Service exposes an XML interface to provide clients with a georeferenced map obtained by superimposition of several layers, coverage and/or features, employing rules specified by the user using styled layer descriptor files. A Web Mapping Service standardizes the way through which clients can specify requests for maps, in terms of output format, resolution, coordinate reference system, geographic envelope and style.

A Web Feature Service exposes an XML interface to let the clients extract geospatial features using GML or other equivalent formats. A transactional WFS system might also support data manipulation such as updates, insertions and deletions by providing support for long term transactions.

A Web Coverage Service exposes an XML interface which supports the exchange of coverage. It provides access to intact (unrendered) geospatial information for client-side rendering, or input to numeric processing (e.g. meteorological models).

The metadata dissemination tier comprises of the Web Catalogue Service (WCaS). We introduced a distinction between Data Dissemination Tier and Metadata Dissemination Tier in order to clearly illustrate the semantic and operational differences between them. Services like WCS, WMS and WFS disseminate mainly geospatial data (coverage, features and maps), while the dissemination of metadata is an accessory task whose objective is to enable the users to refine their queries. Moreover this accessory task is focused mainly on service-level metadata which describes rules for accessing the above mentioned services. A Catalogue Service has a different goal. It supports the ability to publish and search collections of descriptive information (metadata) for data, services, and related information. Metadata in catalogues represents resources' characteristics which can be queried and presented for evaluation and further processing by both humans and software. Catalogue Services are required to support the discovery of and the binding to registered information resources within an information community; hence their main activity is the mediation between clients and data dissemination services.

3. SYSTEM FOR LANDING OPERATIONS

The architecture described in the previous section is the foundation of the tactical decision aid (TDA) system for short term planning of amphibious landing operations developed at NURC. The main goal to be achieved by this system is the fusion of dynamic environmental parameters (METOCs) characterising the landing area into an automatic decision making tool which is able to provide a clear and easily interpretable picture of the environmental conditions over the operation theatre. The decision authority may use this simple description to plan the deployment of the operation resources on the place and at the time of most favourable conditions.

Superior decision capabilities can be achieved by means of forecast models to predict waves and currents in the so called surf area and decision making systems which are able to deal with data uncertainties and vagueness.

The surf area is the region delimited by the coastline and a limit just seaward of the wave breaking zone. It is a highly dynamic area for which special modeling capabilities are needed at high spatial resolution.

The support for automatic decision making is provided by a fuzzy logic expert system that mimics the human decision process in presence of strong uncertainties in METOC data, operational constraints and vagueness. The system is a flexible data fusion engine capable of ingesting predicted METOCs and drawing conclusions about the environmental situation over the operation

theatre, synthesizing the huge amount of forecast data (distributed over time and space) by simple traffic light maps and percentage of favorability.

The architectural design of the entire processing system is challenging due to the necessity of connecting software systems which are heterogeneous in terms of input and output data formats and execution modalities. The system implementation scheme is depicted in fig. 5. The following sections describe the system components in more details.

3.1. Prediction of surf area METOCs

Prediction of waves in the surf area is crucial for supporting amphibious landing operations. In particular, the modified surf index (MSI), the littoral current velocity, the significant wave height and the maximum breaker height are needed for assessing the conditions of the surf zone and aiding the operation planning and execution processes. The system is able to predict surf zone environmental parameters from low resolution wave and wind fields and use these forecasts as inputs to the fuzzy logic TDA system.

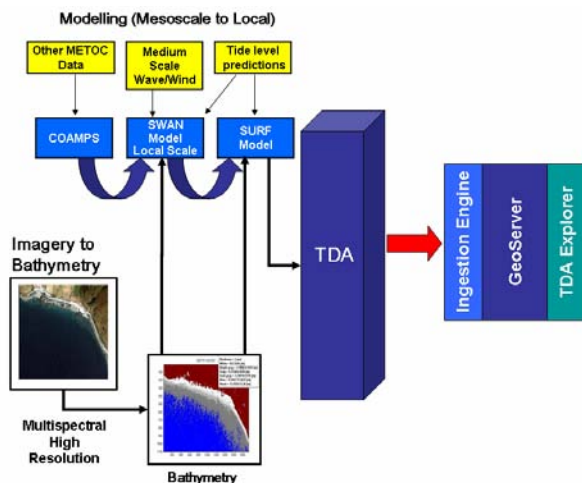


Figure 5: system architecture for supporting amphibious landing operations.

The prediction of the wave field in the surf area needs several intermediate runs of a wave prediction model at different spatial scales, each run producing the wave boundary conditions for the subsequent step. The developed system makes use of the standard one dimensional US Navy SURF model for surf zone predictions being able to predict currents and MSI, besides other standard wave parameters (significant wave height for instance). The Simulating Waves Nearshore model (SWAN) is used for the intermediate scale wave predictions. A four node distributed processing system (the processing chain exploits the producer-consumer principle and the nodes are distributed over the network through the SOAP web service framework as stated in section 2) was developed which includes two SWAN model units for predicting

waves at medium and local scale, one node for the SURF model for prediction in the surf area and one final unit wrapping the TDA for automatic decision making. The system inputs include forecasts of global scale wind and wave fields, bathymetry grids, and tide data. The output wave field at medium resolution is used for setting the boundary conditions of the second SWAN model predicting waves at local scale with a resolution better than 100 m. The SURF model is executed for a chosen number of one dimensional approaching lines to the beach of interest. The wave field predicted by the local SWAN model at the line starting points were used for setting the SURF boundary conditions.

The local SWAN model needs high resolution bathymetry data for the wave field prediction. This may be estimated from a set of multi-spectral passive sensor data acquired from satellite with a resolution of few meters (such as QuickBird or Ikonos). Satellite data in this case has the main advantage of providing high resolution bathymetry estimates even in denied areas. On the contrary, being an indirect measurement method, it must be considered carefully from the calibration point of view.

3.2. Fuzzy logic TDA

The TDA node provides the support for automatic decision making. It is based on a fuzzy logic expert system that is completely customisable depending on the type of operation to be supported. The inputs parameters may be arranged on regular geo-located grids as well as sparse points over the area of interest. The fuzzy system inference engine draws conclusions about the environmental conditions through fuzzy logic operators, comparing the actual METOC inputs, predicted by the models, against a set of rules stored in a knowledge base. The knowledge base is configured depending on the input variables and the operational constraints. The rules are usually provided by experts of the domain and represent the prior knowledge about the operating conditions of the resources involved (vehicles for instance) in the operation. The outputs of the system are the percentage of belief of the three risk categories favourable, marginal and unfavourable over space and time. A traffic light map is built by deciding for the category having the greater percentage. For the amphibious landing operation support, the output is provided for each chosen approaching line. The outputs are stored on files having a format that is suitable for geo-located vector data. The results are provided for each landing point at the end of the approaching lines. Each point has a set of attributes attached storing the information for the decision authority such as the prediction base time and forecast time, the METOC values for the current forecast time, the TDA favourable, unfavourable and marginal output percentage values, and the final maximum rule decision

among the three risk categories (namely, the traffic light map).

The end user has the support of a very simple environmental picture description, through the traffic light map, and the possibility to assess the degree of belief of the decision by means of the risk category percentages. Moreover, once the time window with the best conditions is located, the user can check the consistency of the traffic light map by means of the input METOC variables.

3.3. Presentation through custom web applications

The results of the processing chain are assimilated into the GeoServer through the automatic ingestion engine. The connection between the distributed processing system and the ingestion engine is achieved through the same framework used to distribute the processing nodes i.e. the SOAP Web-Service protocol. The TDA Explore in fig. 5 is a custom web application used to present the results to the end user. The TDA Explorer is intended to be a thin client, based on a web-like architecture in order to be accessed and executed from the standard Internet browsers without installing anything on the local machines, aimed to easily interact with the TDA products stored on a GeoServer based architecture, and distributed through its WFS and WMS OGC Services.

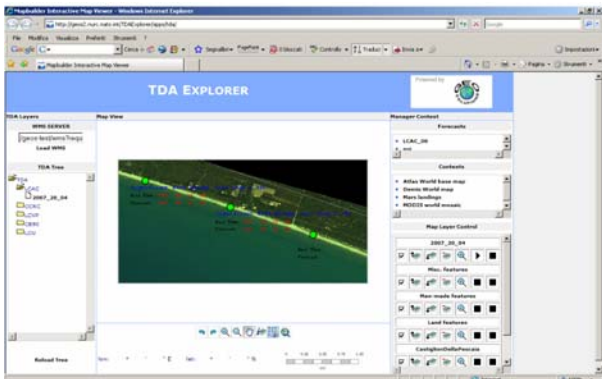


Figure 5: TDA Explorer Web application

4. CONCLUSIONS AND FUTURE WORK

In this paper an end to end processing chain with automated data flow for an effective and affordable preparation and timely distribution of new near-real time environmental pictures was described and an example of its use in the context of landing operation support was examined.

The key points characterising the proposed data fusion support framework are the following:

- integrated approach for fusing remote sensing data, forecast models and in situ observations,
- use of OGC\ISO geospatial services for data and product dissemination through the network in an interoperable way,

- support for chaining heterogenous nodes and build complex automatic data processing schemes through SOAP Web Services and producer-consumer paradigm,
- improved automation through automatic ingestion of data and products into the geospatial server,
- use of expert systems for automatic decision making in order to provide an assessment of the environmental condition impact on operations.

Future work includes improvements on the data management and dissemination tier and on the automatic decision support side.

Support for multidimensional coverage having a spatial domain comprised of 4 dimensions (x,y,z,t) as well as having complex codomain depending on another independent variable (e.g. different bands depending on wavelength) will represent the main research and development effort for improving the capabilities of geospatial services and applications. Part of the work will be used to generate the first prototypes of frameworks for I/O and management of n-dimensional coverages based on the ISO 19123 specification. Furthermore, integration of a WCaaS into the Geoserver application will provide the system with an improved interface between the user and the geospatial services and products.

The exploitation of the predicted environmental parameter uncertainty within the automatic decision making process will provide a reliable assessment of the impact on operations. Moreover, the support for group decision making will allow the system to fuse decisions from several TDA systems using different kind of input data so as to provide an overall picture of the situation over the operation theatre improving the overall reliability of the expert system conclusions.

5. REFERENCES

- [1] Percival, G., "OGC Reference Model, Version 0.1.3", OGC Document 03-040.
- [2] Jeff de la Beaujardiere, W. (ed.) "Web Map Service Implementation Specification (version 1.3.0)" OGC Document 01-109r1.
- [3] Vretanos, P. (ed.) "Web Feature Service Implementation Specification (version 1.0.0)", OGC Document 02-058.
- [4] Evans, J. (ed.) "Web Coverage Service Implementation Specification (version 1.0)", OGC Document 03-065r6.
- [5] Nebert, D., Whiteside, A. "Catalogue Services Implementation Specification (version 2.0.0)", OGC Document 04-021r2.
- [6] Kottman, C., "Topic 5 - Features , Version 4", OGC Document 99-105r2.
- [7] "Topic 6: The Coverage Type and its Subtypes, Version 4", OGC Document 00-106.