Integration of In-situ and remote sensing Data for Water Risk Management

Bernhard H. C. Sputh

Oliver Faust

Lasse H. Pettersson

Torill Hamre

Damiano Vitulli

See next page for additional authors

Presenter/Author Information
Bernhard H. C. Sputh, Oliver Faust, Lasse H. Pettersson, Torill Hamre, Damiano Vitulli, Alastair R. Allen, and Tim Spracklen

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Integration of In-situ and remote sensing Data for Water Risk Management

Bernhard H.C. Sputh[a], **Oliver Faust**[a], **Lasse H. Pettersson**[b], **Torill Hamre**[b], **Damiano Vitulli**[c], **Alastair R. Allen**[a], **and Tim Spracklen**[a]

[a]*School of Engineering, University of Aberdeen, Aberdeen AB24 3UE, UK ({b.sputh, o.faust, a.allen, tim_spracklen}@abdn.ac.uk)*

[b]*Nansen Environmental and Remote Sensing Center, Thormøhlensgt, 47, N-5006 Bergen, Norway ({Torill.Hamre, lasse.pettersson}@nersc.no)*

[c]*Sysmedia S.r.l., Via Paduni 2a, 03012 Anagni (FR), Italy (damiano@sys-media.it)*

Abstract: The combination of environmental measurements from different sensor equipment improves water quality monitoring and associated risk management. The merging of in-situ and remote sensing data is a prerequisite to efficient monitoring of coastal and marine environments, because of their complementary nature and application areas. Both sensor types can monitor the same phenomena, but remote sensing delivers the overall picture and in-situ measuring provides localised, high resolution measurements, over a wide range of parameters. The WARMER (**WA**ter **R**isk **M**anagement in **E**u**R**ope) system combines remote sensing satellite data with in-situ measurements for water quality risk management applications and delivers products to end-users through a web GIS (Geographic Information System) portal. This data combination aspect brings the WARMER system in line with the objectives of the GMES (European Global Monitoring for Environment and Security) initiative.

The ADDP (Aberdeen DISPRO DB Proxy) is a key element in the WARMER system, because it conditions in-situ measurements for the data combination. This paper details the handling of in-situ data within the ADDP, especially its capability to aggregate in-situ data from multiple sources. This aggregation caused problems with data synchronisation and unification, which influenced the overall design of the ADDP.

Keywords: FP6 IST, water quality, risk assessment, remote sensing; in-situ measurements; UMN MapServer; OGC-WMS

1 INTRODUCTION

Coastal waters have many different users, such as: maricultures, fisheries, shipping, and tourism. Each user has specific environmental quality requirements, and is a potential source and victim of pollution. Water quality monitoring is the only way to protect victims and identify the source of pollution. For instance, Pablakis et al. [2001] point out that shipping is a constant source of oil pollution, due to accidental events and deliberate operations. Oil pollution affects fisheries and maricultures by being harmful to or by poisoning sea-living animals and birds. Furthermore, it has a degrading impact on the environment and hence tourism industry. Troell et al. [2003] report that maricultures pollute the environment with feces, antibiotics and nutrients. The high concentration of sea-living animals, inside maricultures, results in a discharge of higher than normal concentration of feces. Furthermore, animals bred in maricultures require additional nutrients. Naturally, not all nutrients get consumed by the caged fish and these surplus nutrients leave the maricultures,

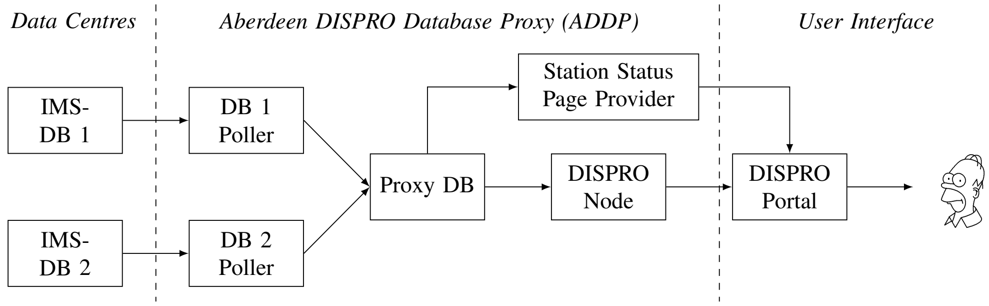

Figure 1: Overview of ADDP and its surrounding elements

either to be consumed by animals outside the mariculture, or to rot on the sea-bed. Finally, maricultures are often mono-cultures, making them vulnerable to diseases and parasites. To cure or protect the animals, antibiotics and other medications are used which leave the mariculture and thus pollute the environment.

The EC WFD (Water Framework Directive) urges the national environment protection agencies to improve the water quality of all their water resources, including coastal waters (see European Commission [2000]). To improve the water quality the level of pollution must be detected and if possibly its origin determined. This can be achieved by regular water quality monitoring. To monitor the quality of coastal waters two methods are available: in-situ measurements using buoys, ships, unmanned vehicles etc. and remote sensing from satellites or aircraft. In-situ measurements are performed by IMSs (In-situ Monitoring Stations) located on buoys or moving platforms such as ships. IMSs monitor the water quality with high resolution, but are limited to the "buoy" location. Monitoring large areas, such as a coastal area, requires a network with a large number of IMSs. This results in high initial deployment and operational maintenance costs. In contrast, remote sensing offers a comparatively cheap way, disregarding the total cost of the satellite, to obtain measurements over large areas and it provides a synoptic mapping of a large surface of the environment. The shortcomings of remote sensing are its limited number of observational environmental parameters and their accuracy, as well as the spatial and temporal resolution of the data.

The WARMER project has developed and implemented advanced IMS sensor technology for high accuracy measurements of a range of "water quality" parameters, using innovative in-situ sensor technologies. Several of the developed sensors are able to measure low concentrations of environmental contaminants present in the natural environment. However, even these low concentrations of environmental contaminants already represent a high risk factor for human health and environmental safety. Several of these environmental contaminants do not have other distinct attributes which would allow indirect detection using e.g. remote sensing — such as water discoloration, which applies for e.g. phytoplankton, sediments and dissolved organic matter, or surface film forming attributes, which applies for e.g. crude oil or biological surfactants.

Apart from advancing the present state of sensor technology, the WARMER project aims to provide a cost effective, all embracing view on the current water quality situation for risk management in aquatic environments, such as coastal areas. We achieve this all embracing view by combining water quality measurements from remote sensing and in-situ sensors. This paper presents our current work on the ADDP (Aberdeen DISPRO[1] DB Proxy), which transforms in-situ measurements to an OGC (Open Geospatial Consortium) compatible format. These data are made available to web-based remote sensing data presentation systems. The presentation of all relevant environmental monitoring information within the same user interface makes this data as easy to access and comprehensible. According to Hepting and Haciag [2006], easy access and comprehensible presentation of environmental data is a vital aspect of such systems. The remote sensing data helps also to identify optimal IMS locations, in order to obtain representative measurements.

[1] DISPRO stands for DISMAR (Data Integration System for Marine Pollution and Water Quality) Prototype.

Accordingly, the system offers cost effective high quality monitoring of coastal areas.

Figure 1 illustrates the information flow within the proposed system. The Data Centres provide in-situ data to the ADDP, see Section 2.1. The ADDP, discussed in Section 3, retrieves the in-situ data, processes it, and provides it to the User Interface. The User Interface (Section 2.3) accesses the remote sensing data held by a DISPRO Node and fuses it with the processed in-situ data.

2 PREREQUISITES

This sections details the sources of water quality information within WARMER: in-situ measurements and remote sensing. The section closes by detailing the DISPRO system we chose to perform the data combination and which also provides the user interface.

2.1 In-situ Measurements

In WARMER a network of instrumented stations obtain in-situ measurements. These IMSs (In-situ Monitoring Stations) measure water quality parameters, such as: temperature, salinity, chlorophyll, oxygen content, and aquatic pollutants. These measurements are taken at fixed locations, given depths, and in specific time intervals. Each measurement is a tuple which consists of time-stamp, location, measured parameter, and measurement value. After obtaining a measurement the IMS transfers it to a data centre, which acts as a central repository and includes automatic quality control routines for the first data assessment. From there the measurement data can be retrieved and assessed by multiple users and applications.

In-situ measurements have a high temporal resolution, with a sampling period from seconds upwards, however at a fixed location. This high temporal resolution prevents missing or misinterpreting, important events, such as high levels of pollution over a short time period, caused by a discharge of pollutants. IMSs are located at critical or representative locations, for instance potential water pollution sources (factories, maricultures, etc.), or where water pollution may cause severe harm (maricultures, water intakes etc.). The IMS measurements may also provide the vertical structure of the water column.

The highly localised monitoring provided by IMSs makes it a challenge to monitor large areas, because the spatial resolution of the field measurement is directly proportional to the number of stations. For example, one possible WARMER field trial is to monitor the water quality of the Boknafjord in Norway. The Boknafjord has roughly $1,000\mathrm{km}^2$ of water surface. Therefore, to achieve a spatial resolution of 1 measurement per km^2 (a typical resolution for remote sensing satellite data) we would need to deploy around 1000 IMSs. This is an impossibly high number of IMS to be deployed, for at least two reasons: (1) high cost of both initial deployment and operational maintenance; (2) deploying a large number of IMSs effectively prevents shipping, which is an undesirable impact on the usability of the area. To monitor an area such as the Boknafjord one is typically only permitted five to ten station locations.

Different applications require access to different measurement parameters. A data centre stores the measurements in a dedicated DB (database), the so-called IMS-DB. In the WARMER system there is only one data centre operated by Sysmedia (Sysmedia [2008]), based in Italy. Sysmedia is responsible for quality control and validation of the stored information.

2.2 Remote sensing

Only a subset of parameters measured by the WARMER buoy system is directly fused with any of the existing (and planned) satellite remote sensing technologies and applications. Accordingly, an integrated monitoring approach using a facet of optical, infrared and microwave satellite remote sensing sensors will be implemented for the demonstration areas. The integrated use of these information sources will form the basis for the assessment of emergency / environmental situations.

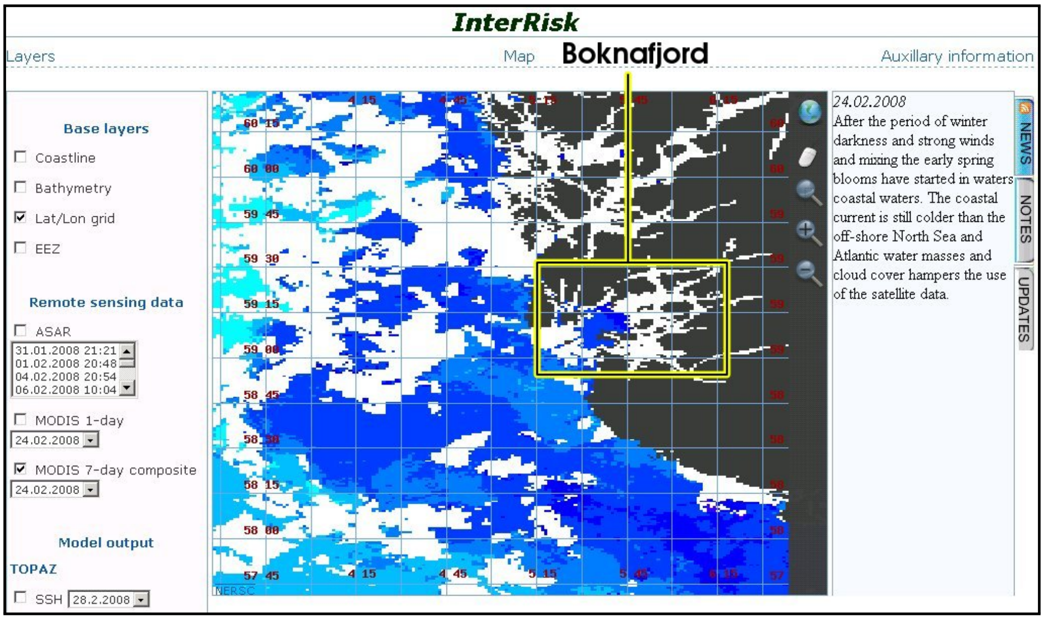

Figure 2: Screenshot of the DISPRO Client user interface, to be used by the WARMER project

The proposed suite of EO (Earth Observation) sensors includes the European Space Agency Envisat sensors - MERIS FR (Medium resolution Imaging Spectrometer 10's meters). These EO sensors however, meet neither the requirements of frequent spatial coverage, nor near real-time availability of the data.

Accordingly, the project demonstrations will take place at coastal locations where one may expect gradients in the level of contaminants. These demonstrations are used to assess the project development. In addition, the information should be supplemented with higher spatial resolution satellite remote sensing information from optical sensors, such as the Ikonos sensor.

Figure 2, shows a 7-day composite MODIS sea surface temperature (in February) of the Boknafjord and abutting coastal areas. Establishing even this, rather low, spatial resolution of the water quality situation in the Boknafjord with IMSs is impractical due to the high cost and its previously mentioned impact on shipping. This clearly demonstrates why combining remote sensing with in-situ measurements is essential for the efficient monitoring of large aquatic environments.

2.3 Data Combination and User Interface

In WARMER we use the DISPRO system to combine remote sensing and in-situ data. DISPRO has been developed by the DISMAR Consortium, coordinated by the Nansen Center (Tuama and Hamre [2007]). We have extended the DISPRO functionality to suit the special needs of the WARMER project.

The DISPRO system includes a set of independent DISPRO Nodes. Each node provides an OGC-WMS (Web Map Service) interface. The DISPRO Portal communicates with the nodes and provides various client applications, such as map viewer, profiler, and metadata catalogue. Figure 1 illustrates the relationship between DISPRO Portal, DISPRO Node and users, in the context of the WARMER application.

The implementation of the DISPRO system relies heavily on Java and XML technologies, with the Open Source UMN (University of Minnesota) MapServer (UMN [2007]), running as a CGI application under the Apache web server, providing the core web GIS (Geographic Information System) client and OGC WMS (both client and server) functionality. The implementation of the main DISPRO components is described in the following.

The metadata catalogue is at the heart of the DISPRO system. It stores data derived from capabilities files returned by participating nodes in response to a GetCapabilities request issued by the portal aggregator. A capabilities file consists of information about the WMS node itself and about each map layer that it can serve. This aggregated dataset underpins DISPRO, and is updated every 15 minutes (configurable). The distributed data structure assures that the latest data from each provider is available in the system.

The web GIS interface (Figure 2) allows a user to view the map layers in a predefined area of interest and map projection. Standard functions, such as select/deselect layer, pan, zoom, and centre are provided, and rely on basic web browser capabilities (no plug-ins required). The DISPRO user interface will be extended with in-situ monitoring specific features, such as the SSP (see Section 3.4).

3 THE ABERDEEN DISPRO DB PROXY

The block diagram in Figure 1 shows that the ADDP acts as an intermediate service between Data Centres and User Interface. The DISPRO Portal provides user interface and data combination capabilities. To increase system performance and reliability the ADDP includes an internal storage service for in-situ data.

The ADDP aggregates in-situ measurement data from the IMS-DBs (In-situ Monitoring Stations-Data Bases) within the data centres. For this aggregation the ADDP utilises multiple DB Pollers, discussed in Section 3.1, which periodically check the IMS-DBs for new in-situ data. The DB Pollers store new measurement data in the Proxy DB, see Section 3.2. The DISPRO Node projects the data stored in the Proxy DB onto a map, which is then fused by the DISPRO Portal with remote sensing data, Section 3.3 gives more details. The S2P2 (Station Status Page Provider), detailed in Section 3.4, provides a SSP (Station Status Page) for each IMS, which, among other information, presents time series plots of in-situ measurements. This allows operators to assess the wider range of in-situ data. These SSPs are available from within the user interface provided by the DISPRO Portal.

3.1 DB Pollers

The DB Pollers have a number of different functions. They obtain all relevant information from the IMS-DB of a Data Centre and store it in the Proxy DB. This involves transforming the obtained information into a format suitable for the Proxy DB. A typical example of such a transformation is the conversion of the station location from latitude/longitude format to PostGIS geometry format (Ramsey [2001]).

In relational databases, such as Proxy DB and IMS-DBs, data sets refer to one another using each others unique IDs. The Proxy DB allows data aggregation from multiple Data Centres by assigning new unique IDs to the data sets. The DB Pollers translate the IDs within the data sets retrieved from the IMS-DBs into Proxy DB IDs. For this purpose each DB Poller has a set of tables which contain the relationship between the IDs of IMS-DB and Proxy DB.

Another problem is that we must assume that the database layouts of the external DB and the Proxy DB differ. To overcome this, we use a special event driven technique to import data from an IMS-DB into the Proxy DB. One aspect of this technique is that the data handling is first sensor and then measurement centric. That means, the DB Poller first determines the available sensors within an IMS-DB. In a second step it checks whether or not these sensors have any measurements unknown to the Proxy DB. When this is the case, these measurements get added to the Proxy DB. The data set that represents a measurement is not atomic, i.e. it refers to other data sets. The measurement data set refers to the data set describing the sensor that took it. The data set, describing the sensor, has therefore to be present in the Proxy DB as well. The system handles this when a DB Poller tries to add a new data set. During this process the system checks whether or not all data sets the new data set refers to are available within the Proxy DB. If a missing data set is discovered the system requests the DB Poller to add it, before the initial data set may be added. This guarantees the consistency of the Proxy-DB contents.

Apart from updating the Proxy DB with new information from the IMS-DBs, the DB Pollers also trigger the generation of mapfile and XML metadata files. This happens once a new sensor-type gets added to the Proxy DB. This includes the generation of a legend image for this sensor type, which maps colours to measurement values, using the same colour scales for similar parameters.

3.2 Proxy DB

The Proxy DB is a spatial enabled (PostGIS, see Ramsey [2001]) data base (PostgreSQL, see Pos [2008]). The ADDP uses it to store all water quality information that the DB Pollers retrieved from the IMS-DBs. This local storage of water quality information increases system performance and reliability.

Presently, the Proxy DB uses four tables to hold the retrieved data: Measurements, Sensors, SensorTypes, and Stations. The rows of the Measurements table hold individual measurements. Each row consists of: Measurement-ID, time, location, value, and the Sensor-ID of the sensor that performed the measurement. The Sensor-ID refers to a row of the Sensors table, where each row has the following columns: Sensor-ID, SensorType-ID, and Station-ID. The table SensorTypes contains detailed information for each different sensor type, the main elements are: SensorType-ID, name, unit, precision, minimum value, and maximum value. The last table of the Proxy DB is the table Stations which contains entries consisting of: Station-ID, name, location. Furthermore, the Proxy DB automatically generates a table which contains the most recent measurement of each sensor. This is achieved using a stored procedure within the Proxy DB. This procedure is called by the PostgreSQL server once a new measurement is added to the Measurements table.

3.3 DISPRO Node

The DISPRO Node component of the ADDP interfaces directly to the DISPRO Portal. It provides the portal with the latest in-situ measurements, available in the Proxy DB. The data gets served by UMN MapServer running as a web application within the Apache HTTP server. The UMN MapServer provides specifications of available data layers within a so-called mapfile. In the case of in-situ measurements, each sensor measurement is represented by a coloured symbol, at the location of the IMS. The symbol colour represents the values of the various measured parameters. A mapfile consists of one or more layers. Each layer groups similar information, allowing for different parameters to be presented within the same mapfile. The ADDP uses this feature by grouping all measurements obtained by the same sensor-type within one layer, i.e. all temperature measurements in one layer, all chlorophyll measurements in another layer, and so on. This allows operators to select between the different parameters.

The UMN MapServer obtains the data it projects directly from the Proxy DB. This is possible with UMN MapServer if the DB is spatially enabled. One example for such a spatially enabled data base is PostGIS (Ramsey [2001]). Using UMN MapServer in conjunction with a spatially enabled DB makes it unnecessary to generate a new mapfile every time new data becomes available. In such a configuration it is only necessary to generate a new mapfile when a new layer gets added. This is the case when a DB Poller adds a new sensor-type to the Proxy DB. This triggers the generation of the new mapfile.

For each mapfile layer the ADDP DISPRO Node provides additional information, such as: title, legend image, and a XML metadata file. The legend image relates station symbol colours and measurement values within the user interface. The XML metadata contains, among other information, source and contact information related to the sensors represented within this layer.

3.4 Station Status Page Provider

The S2P2 (Station Status Page Provider) provides a SSP (Station Status Page) for each IMS in the system. This page contains information about the station, i.e. location, operational status, available sensors, and other parameters. Furthermore, it provides a time series plot over past measurements for each sensor. These time series plots allow operators to assess what happened in the past.

User interfaces such as the SSPs should provide a responsive look and feel. To achieve this the SSPs are implemented using Ajax (Asynchronous Javascript And XML) technology. The Ajax technology allows us to develop the SSP as a GUI application executing within standard web browsers. Similarly to normal applications, the SSP communicates with the S2P2 backend server to obtain or commit information. We implement the S2P2 using the GWT (Google Web Toolkit), (Goth and Costlow [2007]). The GWT allows us to develop the whole application with the Java programming language. The GWT translates the resulting Java program into an AJAX enabled web page, the SSP. The SSP communicates with the S2P2 via a dedicated protocol, which gets generated by the GWT as well.

4 Conclusion & Further Work

Coastal waters have many different users, each user is a potential source and "victim" of pollution. Water quality monitoring is the only way to protect this environment in order to preserve its usability. The WARMER project integrates in-situ and remote sensing data to monitor water quality. After introducing both types of environmental monitoring, the paper discusses how both types of monitoring complement each other for the monitoring of large aquatic environments. The proposed system represents an all embracing tool for water monitoring and risk management of coastal and marine environments. Based on the ADDP structure this paper has outlined a system which combines in-situ and remote data.

The ADDP system is composed of independent components which are arranged around a central data base. This data base aggregates the contents of multiple Data Centres. The data base unifies the data access for the different system components. This makes the DISPRO Node and S2P2 components independent from the data layout of the different data centres. The result is an easy to integrate, extendible, responsive and reliable information handling and assessment system.

Its easy integrability and extendability allows its use outside the WARMER project boundaries, thus allowing others to take advantage of the combination of in-situ and remote sensing data. The ADDP will be released as Open Source, thus a community of ADDP users and developers could emerge. Over time this community will shape the ADDP to suit the different data presentation and data aggregation needs and thus turn the ADDP into an industrial strength tool for environmental monitoring.

The system presented in this paper is only one step towards the combination and presentation of remote sensing and in-situ monitoring measurements. It will be tested in the WARMER project field trials, during 2008. The next ADDP development step is to integrate mechanisms to present and handle alarms generated by IMSs. This will require additions to most ADDP components. Furthermore, to demonstrate the power of the divide and conquer approach behind the ADDP design more ways to access the data of the Proxy DB will be developed. One example would be to implement an OGC-SOS (Sensor Observation Service) component, following OGC-SWE (Sensor Web Enablement) recommendations, detailed in OGC [2006].

Acknowledgements

This work has been supported by the EC FP6 project WARMER (**WA**ter **R**isk **M**anagement in **E**u**R**ope) project (contract no.: FP6-034472). The DISPRO system was first developed in the EC FP5 DISMAR (Data Integration System for Marine Pollution and Water Quality) project (contract no.: IST-2001-37657) and is currently being enhanced in the EC FP6 InterRisk (Interoperable GMES Services for Environmental Risk Management in Marine and Coastal Areas) project (contract no.: FP6-035253).

References

European Commission. Directive 2000/60/EC of the European Parliament and of the council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Communities*, L327(1), 22 December 2000.

Goth, G. and T. Costlow. Currents: The Google Web Toolkit Shines a Light on Ajax Frameworks. *IEEE Software*, 24(2):94–98, 2007. DOI: `http://doi.ieeecomputersociety.org/10.1109/MS.2007.39`.

Hepting, D. H. and T. Haciag. Opening Access to Environmental Software Systems. In *Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software"*. International Environmental Modelling and Software Society, July 2006.

OGC. OGC® Sensor Web Enablement: Overview And High Level Architecture. Internet, July 2006. `http://portal.opengeospatial.org/files/?artifact_id=15540` (last visited 04.03.2008).

Pablakis, P., D. Tarchi, A. J. Sieber, G. Ferraro, and G. Vincent. EUR 19906 EN: On the Monitoring of Illicit Vessel Discharges — A Reconnaissance Study in the Mediterranean Sea. Technical report, EC DG - Joint Research Centre & EC DG - Environment, 2001.

PostgreSQL Global Development Group. *PostgreSQL 8.3.0 Documentation*, 2008. `http://www.postgresql.org/files/documentation/pdf/8.3/postgresql-8.3-A4.pdf` (last visited 27.02.2008).

Ramsey, P. PostGIS Manual. Technical report, Reflection Research Corporation, 2001. `http://postgis.refractions.net/docs/postgis.pdf` (last visited 10.02.2008).

Sysmedia. Sysmedia S.r.l. — Home Page. Internet, March 2008. `http://www.sys-media.it/` (last visited 07.03.2008).

Troell, M., C. Halling, A. Neori, T. Chopin, A. Buschmann, N. Kautsky, and C. Yarish. Integrated mariculture: asking the right questions. *Aquaculture*, 226(1-4):69–90, 2003.

Tuama, É. Ó. and T. Hamre. Design and Implementation of a Distributed GIS Portal for Oil Spill and Harmful Algal Bloom Monitoring in the Marine Environment. *Marine Geodesy*, 30(1):145 – 168, January 2007. DOI: `http://dx.doi.org/10.1080/01490410701296671`.

UMN. UMN Mapserver. Internet Homepage, December 2007. `http://mapserver.gis.umn.edu/` (last visited 14.12.2007).