Developing a GIS-Database and Risk Index for Potentially Polluting Marine Sites

Giuseppe MASETTI, Brian CALDER and Lee ALEXANDER, USA

SUMMARY
The increasing availability of geospatial marine data provides an opportunity for hydrographic offices to contribute to the identification of “Potentially Polluting Marine Sites” (PPMS). These include shipwrecks, oil rigs, pipelines, and dumping areas. To adequately assess the environmental risk of these sites, relevant information must be collected and converted into a multi-scale geodatabase suitable for site inventory and geo-spatial analysis. In addition, a Risk Index – representing an assessment of the magnitude of risk associated with any site – can be derived to determine the potential impacts of these PPMS. However, the successful collection and integration of PPMS information requires some effort to ‘normalize’ and standardize the data based on recognized international standards. In particular, there is benefit in structuring the data in conformance with the Universal Hydrographic Data Model (IHO S-100) recently adopted by the International Hydrographic Organization. In this paper, an S-100 compliant product specification for a PPMS geo-spatial database and associated Marine Site Risk Index is proposed which can be used by national hydrographic offices and marine protection agencies.

Key words: GML, PPMS, Risk Index, shipwreck, S-100

1. INTRODUCTION
In addition to ballast water and non-indigenous or introduced marine species, petroleum-based and chemical pollutants are significant threats to the global marine environment. The increasing number of marine polluting events that have occurred throughout the world have led to an increased focus on the need to look proactively at the risks of oil and other pollutants being released from such submerged sources as shipwrecks, pipelines and dumping areas (Monfils 2005; Gertler et al. 2009). Collectively, these sources can be considered as Potentially Polluting Marine Sites (PPMS).

The prevalence of these sites is constantly increasing, and is likely to continue to do so. For example, the potential for a polluting event to occur at any particular wreck site generally increases with time since the sinking due to the ongoing corrosion of the wreck. An average corrosion rate of 0.1 mm/year might imply that hulls of ~25 mm steel (typical for many modern vessels) should not decay quickly, but internal structures are often considerably thinner, and their collapse can lead to premature release of pollutants even if the main hull remains intact (Schumacher 1979; Macleod 2010). Effects vary with depth and composition of water (Macleod 2002), but the significant number of vessels sunk during World War II (WWII) and in the many regional and local conflicts since (see, e.g., Brown 2005), particularly those that may have been damaged by enemy action as they sank, mean that there is significantly increased risk of major
polluting events in the near future. Improved methods for the collection, analysis, and interchange of information on wrecks are needed.

Successfully managing information about such sites, and making it available for use and exchange in a uniform manner, is critical to effectively supporting a proactive approach to monitoring and remediation. In particular, if a solution is to be effective, it must address three fundamental requirements: 1) it must be generic enough to handle different types of potential polluters and auxiliary information; 2) it must enable easy exchange and re-use of information; and, 3) it must be standards-based to allow for ready adoption into available tools.

Shipwrecks are the most obvious, but by no means the only source of pollution. For example, pipelines or abandoned wellheads can release pollutants, and old munitions or chemical weapons dumping sites are obvious risks to fishermen, divers and the local community. A successful database solution must be generic enough to represent various types of potential polluters, but do so in such a manner to allow specific analyses to be conducted that enable the site to be properly classified. At the same time, the solution must support integrated thinking about how to plan for, and respond to, potential polluters. This was recognized by the International Maritime Organization recommendation “to develop regional co-operation on aerial and satellite surveillance” for problems (IMO 2004). Gathering all relevant data in a sufficiently flexible database is one way of supporting this process.

Determining who is responsible for both the activities and cost of remediation after a polluting event if often complex, and may be exacerbated by national and international law. For example, it is generally held that shipwrecks continue to belong to their nation after they are sunk (Johnson 2008; Aznar-Gomez 2010), but it is unclear whether the owner is responsible for damages caused by pollution related to these wrecks. The U. S. Navy removed oil from the USS Mississinewa after a storm caused leakage of fuel (U. S. Navy 2004) but asserted that this did not constitute a precedent (Guerin et al. 2010). It is likely that many events or potential events will include more than one actor, therefore, and exchange of information in a uniform manner is essential in timely appraisal and response (Woodward 2008). Definition and adoption of a state-neutral database is therefore important in supporting the planning and response goals.

As a consequence of the requirement for interchange of information, it is inevitable that data related to PPMS are going to be used by multiple agencies across multiple software and hardware platforms. Although often dismissed as an implementation problem, it is therefore important to consider requirements for compatibility and standardization when defining the structure of any putative database. In addition, while working within the constraint of a given standard often implies extra effort, this is rewarded by re-use of already available resources (e.g., feature catalogues) and can significantly improve rate of adoption in standard data manipulation packages such as desktop GIS systems. A practical (rather than merely efficient) solution for PPMS must therefore consider the requirement for a standards-based definition.

We propose in this paper a model for the implementation of a PPMS geo-spatial database that attempts to satisfy these requirements. Drawing on previous example databases that were built parochially for specific purposes, core and extension requirements were extracted for a variety of potential polluters, augmented by auxiliary information such as relevant resources (e.g.,
availability and location of pollution response equipment) and complementary information (e.g., sensitivities of coastlines to particular pollutants). To ensure standards compatibility, the database was developed based on the International Hydrographic Organization’s S-100 approach (IHO 2010), while providing generic descriptions of various potential polluters, it is defined through both a UML description (to assist in clear documentation) and uses an XML-based schema to provide a GML-structured computer-translatable description of the model. This paper describes the basic structure of the model and its XML implementation, and illustrates one potential use by considering how a Marine Site Risk Index (MaSiRI) can be created using the components of the database.

2. MODELING THE PPMS’ UNIVERSE OF DISCOURSE

Evaluating the entities required in a PPMS database is complicated by the diversity of objects to be represented. However, some important work has been conducted recently with the aim of cataloging shipwrecks by ocean/basin location. This includes the South Pacific Regional Environment Program (SPREP) (SPREP 2002; Monfils et al. 2006; Talouli et al. 2009) and Barrett Project (Barrett 2011), the Atlantic, Mediterranean and Indian Ocean (AMIO) database (Monfils 2005), a Mediterranean area in the Development of European guidelines for Potentially Polluting shipwrecks (DEEPPI) project in 2005 (Alcaro et al. 2007), a global International Oil Spill Conference (IOSC) study in 2005 (Michel et al. 2005), etc. Collectively, these have been analyzed in regard to the types of information that are fundamental for a PPMS GeoDB. A similar approach for non-shipwreck PPMSs was more difficult conduct since there is less in the literature about this type of information in an integrated environmental-risk framework (Overfield 2005; Aichele 2010).

The conceptual model developed here was documented using an object-oriented notation known as the Unified Modelling Language (UML) which was then realized in practice using the Geography Markup Language (GML).

2.1 S-100 compliance Product Specifications

As defined in IHO S-100, a Product Specification (PS) is “a description of all the features, attributes and relationships of a given application and their mapping to a dataset” (IHO 2010). A PS is different but related to metadata: while metadata describes how a dataset actually is, a data PS describes how it should be, focusing on the requirements. Since S-100 is a set of profiles of the ISO TC 211 standards for Geographic Information, the proposed PPMS GeoDB PS is in compliance also with the ISO 19100 series of geographical information standards.

The proposed PPMS GeoDB PS conforms to the S-100 requirement to be a precise and human readable technical document that describes a particular geospatial data product for hydrographic requirements (IHO 2010). This includes machine readable files that define the structure (XML Application Schemas), and can be converted to a XML Product Specification.

The S-100 prescribed workflow was used to create the PPMS GeoDB PS. Outputs included:
- definition of a vector-only product.
– selection of required features, feature attributes, and enumerates in existing IHO Data Dictionaries.
– identification of some new features that will be submitted for inclusion in an IHO Supplemental Dictionary.

The defined features and attributes were then described in a Feature Catalogue, and geometry types required in the product were determined. Any new geometry types will be not need to be added to the S-100 framework for the proposed PS.

At this point, it was possible to construct an Application Schema. The creation was conducted in two different but related ways: a Logical model, using a conceptual schema language (UML), and a Physical model using an encoding specific language (XML Schema). The resulting UML model indicates how the data are logically organized.

### 2.2 The Conceptual Model

In the proposed PPMS GeoDB PS, any product has a root element instance of the *Root* class. This root element may be related by composition with three types of composite Feature Collections (Figure 1). Each of these composite Feature Collection can have an unbounded number of basic Feature Collections, each one represented by an abstract class.

The entities to model the possible types of PPMS are heterogeneous: from submarines sunk during WWII to oil rigs (Figure 2). Since some of these entities are already present in a basic “safety-of-navigation” form in the IHO Registry, they are enriched with a series of new attributes and enumerations, mainly on the basis of the content of the existing databases previously reported and the classification proposed by a Regional Marine Pollution Emergency Response Centre for the Mediterranean Sea (REMPEC 2004). As example, Figure 3 outlines attributes and relationships proposed for Potentially Polluting Shipwrecks (PPSW).

![Figure 1 - Relationships of Root class.](image)

Dumping areas are another selected entity due to the large quantities of live ammunition, mines, chemical warfare agents (CWA), and other explosives present in a large number of marine sites.
(Plunkett 2003; Beddington and Kinloch 2005; Sato 2010). This situation is the result of the past conviction that the dumping of CWA at sea was the best disposal method rather than to store them or incinerate them (Overfield 2005). Currently, an increasing number of injuries and problems related to these dangerous objects are being reported (Laurin 1991; Simons 2003). Although the position of a large part of these dumping sites is known, many problems come from the buoyancy of containers used to store the waste materials, and the difficulties for the local authorities to supervise the correct position during dumping operations.

Abandoned and exploratory wells also represent a threat for structural failure over time, and the Deepwater Horizon disaster recently highlighted the dangers related to oil rigs and offshore extraction of hydrocarbons (Orth 2011). Even if this last event remains in the memory of public opinion, large platform accidents represent only a limited part of marine oil pollution (Fingas and Charles 2001) when compared to periodic releases of water containing small amounts of oil from offshore oil installations (Espedal and Johannessen 2000; Farmen et al. 2010). Having these represented in the proposed GeoDB allows for spatial analysis to correlate objects with satellite Synthetic Aperture Radar (SAR) or other remote sensing sensors to distinguish between slicks due to hydrocarbon release and natural phenomena (Brekke and Solberg 2005).

![Diagram of sub types and relative relationships of the AbstractPPMS class.](image-url)
Some additional data resources are required to enable useful products to be generated from the GeoDB. These include shoreline, archaeological sites, fishing areas/farms, marine sanctuaries, tourist installations, but are not strictly objects in the PPMS sense. As such, they are organized in two related groups: ResourcesCollection for marine resources directly or indirectly related to the PPMS, and ComplementaryInfoCollection for information auxiliary to the previous two entity clusters.

2.3 Use of GML for Encoding

S-100 does not mandate particular encoding formats (IHO 2010). Among the many possible encoding standards, the PPMS GeoDB provides an informative implementation using the Geography Mark-Up Language (GML).

The GML is an XML-encoding tag language defined by the Open Geospatial Consortium (OGC) to describe geographic objects (Lake 2004). Being built on the Extensible Mark-up Language (XML), it has some advantages of binary file formats (easy to understand by a computer, compact, the ability to add metadata), as well as some advantages of text files (universally interchangeable).
Since it is accepted by the mass of industrial companies and research institutions, GML essentially acts as a *de facto* standard in spatial data processing and exchange. In 2007, version 3.2.1 became an international standard as ISO 19136. This ISO GML provides “[…] an open, vendor-neutral framework for the description of geographical application schemas for the transport and storage of geographic information in XML” (ISO 2007).

2.3.1 From the proposed PPMS GeoDB PS to the GML Application Schemas

The steps followed in the creation of several GML Application Schemas for a Potential Polluting Marine Sites Geospatial Database are:

- Provide the declaration of a target namespace;
- Import the appropriate GML Core Schemas;
- Derive directly or indirectly all objects and object collections from the corresponding GML abstract types;
- Define properties (as global or local elements) for each object’s content model;
- Define attributes for all of these objects and properties;
- Define Metadata Schemas as a function of the schema-defined objects.

2.3.2 PPMS GeoDB data validation and manipulation with XML Data Binding

Because GML is a markup data format (i.e., data without instructions) and not a programming language, the application of any operation to the information stored has to be implemented in an application written in a suitable programming language. Thus, in order to apply some data validation and manipulation on GML document based on the PPMS GeoDB PS, a basic C++ application is being developed.

Commonly, a program working with data stored in an XML format adopts either the Document Object Model (DOM) or Simple API for XML (SAX) method. Both DOM and SAX work at a raw representation of the XML structure (elements, attributes, and text). Thus, the developer has to write a substantial amount of bridging code to transform information encoded in XML to a representation more suitable for the application. For the PPMS GeoDB application we chose instead an approach called XML Data Binding. This approach, skipping the raw representation of XML, delivers the data in an object-oriented representation generated by a compiler from an XML schema (Kolpackov 2007; Surhone et al. 2010). XML Data Binding is a more efficient way to handle the GML documents, given the complexity of the PPMS GeoDB Application Schemas.

3. A RISK INDEX AS A PPMS GEODB TOOL

Health, environmental and safety issues are an inherent part of industrial societies. Their public opinions are always more sensitive to risks of major accidents happening near populated, environmentally and economically sensitive areas. The impacts of natural or technological disasters can be prevented, or at least bounded, through an integrated approach to environmental risk assessment and safety management to identify the elements of risk and to prioritize actions (Fedra 1998; Goodchild 2010). While many studies are present in fields like floods, earthquakes
and forest fires, a limited number are centered on the detection, study and analysis of risk from oil spill and other marine pollutants incidents (Sofotassios et al. 1997; Kassomenos 2004; Castanedo et al. 2009; Pincinato et al. 2009). The information collected by the proposed PPMS GeoDB represents a contribution to this issue at global and sub-national scale; nevertheless the development of some tools and indicators structured on this product is desirable to better manage and monitor the risk of a large number of PPMSs.

Although the main target of the PPMS GeoDB Application is a PPMS inventory (Figure 4), its implementation can be a tool for each phase of the disaster management cycle: emergency response, recovery, development, mitigation, and preparedness (Figure 5).

The possibility to identify potential risks before the release of pollutants is a key element for a proactive approach. This approach could permit evaluation of each shipwreck site in order to decide on a direct intervention (i.e. the removal of the threat sources), the isolation of the threat, the preparation of a release management plan before the event, or the definition of a monitoring protocol, etc.
At the same time, a PPMS GeoDB permits inventory of possible assets and responders present in the area in case of a release notice. In the case of an unidentified source of oil (or any other pollutant) the PPMS GeoDB could return a list of suspected sites, possibly on the basis of the results from an analysis of oil samples recovered that permits determination of the type and age of the oil.

In the remaining of this paper, a possible risk assessment tool is presented.

4. DESCRIPTION OF THE MARINE SITE RISK INDEX

The Marine Site Risk Index (MaSiRI) is a tool for risk assessment based on the information collected using the PPMS GeoDB Product Specification. It is independent from the GML implementation previously described. However, in the project described in this paper, it is integrated in a ‘pilot’ PPMS GeoDB implemented using GML.

Its main aim is to provide an index that evaluates possible environmental impacts of each PPMS on the surrounding area and shoreline, analyzing and ‘weighting’ some of the information present in the proposed Product Specifications. For instance, the proximity of historical and archeologically significant sites increase the risk related to PPMSs being close to the area.

In order to reduce the intrinsic subjectivity of any risk assessment process, the proposed risk index follows a fixed series of steps and defines look-up tables to identify resources at risk, additional threats, distance of available assets, coastal observations, etc. The results of this automated process may be partially modified by a registered experienced user (with an increment or decrement Expert Correction Factor), e.g., on the basis of external information such as records of similar incidents or by analyzing previous events on the same PPMS stored in the GeoDB.

An issue related to any risk index is the need to have some indications of the reliability of the index value, in other words a meta-risk value to concisely describe the quality of the risk assessment process. The consequence of a lack of similar information can erode the users’ trust in the validity of the risk index, besides having potential dangerous effects in the decision makers’ choices. For instance, identification of the ‘most potentially polluting shipwreck’ in the...
DEEPP database mobilized public opinion to push for an intervention, but when investigated, the Italian Coast Guard found it was empty of any carburant (Alcaro et al. 2007; Baccicalupi 2008). This represents a clear example of the need to provide information about the risk assessment process (e.g., about the statistical approach adopted to estimate the amount of remaining carburant). The solution adopted in MaSiRI is to provide a level of confidence that tends to increase along the risk assessment process (outlined in Figure 6).

![Figure 6 - Main steps in the MaSiRI calculation based on the information collected in a PPMS GeoDB.](image)

The filtering step avoids the application of the MaSiRI to a given PPMS in cases where not sufficient information is present in the GeoDB (the index is set to “unknown”) or they do not match threat criteria related to size, time and material, e.g. nuclear waste (the MaSiRI is “not applicable”). The filtering approach used here is based on (Lindström 2006; Overfield and Symons 2009).

The basic idea of the core criteria step is to provide a rough indication of the risk index based on three elements common to each typology of PPMS: distance from the shoreline, the volume and the type of pollutant. Since the potentially polluting volume is often unknown, the estimation approach suggest by the DEEPP Project (Alcaro et al. 2007) is adopted if required, although we assign this a low level of confidence.

The hazard represented by the polluting volume is scaled as a function of the distance from the coast. The resulting Scale Factor (from 0.1 to 10) is combined with four pollutant categories
(related to hydrocarbons and chemical materials) providing a first evaluation of the risk index as presented in Figure 7.

The refinements are mainly PPMS type specific. They have the double role to evaluate a series of elements related to the PPMS (e.g., corrosion effects, marine biodiversity, distance from emergency facilities, site depth, shoreline characteristics, socio-economic indexes, meteor-oceanographic elements, etc.) and to increase the level of confidence in the resulting MaSiRI.

![Figure 7 - Risk Index derived from the Scale Factor and the type of pollutants.](image)

Each new PPMS event requires a review of the MaSiRI, and the result of this iterative process is the definition for each PPMS of minimum and maximum values used to catalogue the risks of PPMSs present in the whole database. The range of these two values tends to reduce with the amount of information.

5. CONCLUSIONS

Recent pollutant releases from PPMSs have resulted in significant impacts, including loss of marine life, economic impacts to coastal areas, and high cost to mitigate the effects. Standardized collection of information about these sites can contribute to monitoring and reducing these events.

The described PPMS GeoDB, developed in the S-100 framework, is a possible solution by providing a georeferenced picture of hazardous sites and related marine resources. At the same time, the Marine Site Risk Index, built on the GeoDB enables a risk assessment of undersea threats.

The adoption of an S-100-compliant GeoDB standard can become an important global contribution from the hydrographic community to reduce or at least better manage environmental and economic risks related to Potentially Polluting Marine Sites.
REFERENCES


Alcalo, E., E. Amato, F. h. Cábioch, C. Farchi and V. Gouriou, 2007, DEEPP Project Development of European guidelines for Potentially Polluting shipwrecks, Rome (Italy) / Brest (France), EU D.G. Environment - Civil Protection Unit.


IHO, 2010, S-100 - Universal Hydrographic Data Model, Monaco (Montecarlo), International Hydrographic Bureau.


Marine Spatial Data Infrastructure and Management 2
Giuseppe Masetti, Brian Calder and Lee Alexander
Developing a GIS-Database and Risk Index for Potentially Polluting Marine Sites

CHC 2012
The Arctic, Old Challenges New Approaches
Niagara Falls, Canada 15-17 May 2012


Plunkett, G., 2003, Chemical Warfare Agent (CWA), Sea Dumping off Australia, Australian Government Department of Defence.


Sato, R., 2010, Sea-Dumped Chemical Weapons: Japan, Washington (USA), Global Green USA.


Surhone, L. M., M. T. Timpledon and S. F. Marseken, 2010, XML Data Binding, VDM Verlag Dr. Mueller e.K.
Talouli, A., T. Gilbert and R. Monfils Gilbert, 2009, Strategic Environmental Assessment and potential future shoreline impacts of the oil spill from WWII shipwreck Hoyo Maru Chuuk Lagoon - Federated States of Micronesia, SPREP.

BIOGRAPHICAL NOTES

Giuseppe Masetti has served as hydrographer with Italian Navy since 1999 and, in 2008, graduated in Marine Geomatics. After a period onboard the hydrographic vessel ITN Aretusa as Operation Officer, he achieved the FIG/IHO Cat. A certificate (2010). As a research assistant at CCOM-JHC and PhD candidate at University of Genoa, he is working on various issues related to potentially polluting marine sites.

CONTACT

Center for Coastal and Ocean Mapping & Joint Hydrographic Center
24 Colovos Road
Durham, NH 03824
USA
Tel. +1 (603) 862-1417
Fax +1 (603) 862-0839
Emails: gmasetti@ccom.unh.edu (Masetti) – brc@ccom.unh.edu (Calder)
Web site: http://ccom.unh.edu