2011 2nd Intl Conf on Innovative Computing & Communication and
2010 Asia-Pacific Conf on Information Technology & Ocean Engineering

Geomatrix model as new tool for improving oil spill surveillance*

Michele Cocco[1], Sara Colantonio[2], Mario D'Acunto[2], Massimo Martinelli[2], Davide Moroni[2], Gabriele Pieri[2], Ovidio Salvetti[2], Marco Tampucci[2]

[1]*Coordinator, ARGOMARINE Project, Tuscan Archipelago National Park, Loc. Enfola - 57037 Portoferraio-Isle of Elba (LI), Italy*

[2]*Istituto di Scienze e Tecnologia dell'Informazione, Consiglio Nazionale delle Ricerche Via Moruzzi 1 I-56124, Pisa Italy*

{name.surname}@isti.cnr.it

Abstract **– The ability to remotely detect and monitor oil spills at sea is becoming increasingly important due high demand of oil based products. As a consequence, shipping routes become much busier and the likelihood of slicks occurring will also increase. If applied correctly, an integrated remote sensing system can act as a beneficial monitoring tool. The integrated system should monitor ship traffic and marine operators using a sensing capability made of electronic sensors, geopositioning tools, and a communication infrastructure network. In this paper, based on the concept of dynamic risk, we propose a new model that should account in an unique scenario two different classes of data correspondent, in one case to possible sources of oil spill pollution events, and in the other one case to real time sensing monitoring. We name this model Geomatrix. Its successful implementation could involve a reduction of the costs for an effective real time monitoring of large marine areas, including Oceans.**

Keywords, Oil Spill, Remote monitoring, Geopositioning, dynamic risk, Bayesian reasoning.

I. Introduction

The recent Deepwater Horizon oil spill has received unprecedented media attention as an ecological disaster. The potential for future spills is huge, but technologies for mitigating spills are not improving fast enough, and the disaster warning is not being learned by the governments of the countries with strong oil industry sustain A. Jernelöw [1]. In this paper, we present a new model for the automatic oil spill recognition. The monitoring is realized by means of sophisticated electronic, geopositioning, and communications tools connected through a high speed network along with data transmission through suitable data links. Data collected from different sensing sources in an independent and remote fashion are sent to a main acquisition and elaboration central unit. Motivation and scope of the new model is the safe detection, notification and interventions on vessels in emergency situation and the protection of sea and coast environment, endangered by heavy and continuous activities, mainly due to intense ship traffic, generating a consistent pollution risk. All the data and the information obtained will be merged and elaborated in a Marine Information System (MIS) that is an information system where remote sensing data, field experiment results and estimates from simulation models must be integrated, and tools for data storage and retrieval, data manipulation and analysis, as well a for presentation, is made available to the community through a common computer interface. In the next section, we will present the basic properties and definition of the model.

II. The ARGO-GeomAtrix model

The Geomatrix model has been developed within the *Argomarine* project, [2] (in the following denoted as ARGO-G). It has been thought as a smart system that should conjugate the ability to monitor in real time large marine areas, involving a reduction of costs due to the surveillance. The main quantities introduced into the system are three:

The Load function **L**, taking in account all the possible source of oil spill pollution events that can be the ship traffic, shore activities, or oil platforms, when present in the marine area or nearby.

The Monitoring function **M**, taking in account all the skills and facilities for the oil spill monitoring included communication infrastructure.

Such two functions define the cost function **S**, defined as **S**=**M**-**L**. **S**=0 represents the ideal situation occurring when the monitoring activity is adequate to the local load and the costs incidence is optimal.

The third main function denoting the ARGO-G model is the dynamic risk, *R*. The dynamic risk is based on the Kaplan theory on the risk, [3], and modified for being used for marine oil spill surveillance [4]. The functions **L**, **M** and **S** are time dependent, depending locally by the ship density and traffic, so the risk is dynamic in the sense not only that the functions can changes in time, but that same configurations of ships on the same marine area could involve different dynamic risks. So defined, the dynamic risk is strictly connected to inferential methods [5].

The type of risk under consideration is the risk of oil spill pollution (OSP) event and subsequent environmental damage. The risk is commonly measured in units of tons of oil spill per ship year, referred as pollution risk in the literature [6]. However, since the impact of oil spills may be very considerably depending on what, when and where it is spilled, it may be more useful to include some measure on the environmental impact of the OSP, thus assigning a risk unit of OSP impact for ship year. This factor is commonly recognized

* This work is supported by the Argomarine Project, 7th Seventh Framework Program Grant # 234096

and defined as *environmental risk* [6]. Through this is a less stringent measure in terms of absolute risk levels, it is convenient for a comparison of the risk connected to different ships. It must be noted that the proposed unit for the risks R is not informative in any context outside of the specific applications of the presented model. However, since the objective is to produce a decision support tool to aid in the priorization of ships, the units used need to be accurately checked when used in different environments. Since the pioneering work by Kaplan, [3], provides a definition of system risk as a complete set of triplets $\{\alpha,\beta,\gamma\}$, where α describes the context of an accident scenario, β is the likelihood of an accident occurring in the scenario and γ is a description of the consequences associated with it.

Any generated accident scenario, α, is recorded to a database describing its accident and consequence descriptors, accident type and incident type preceding the accident producing OSP. Accident types considered can be collisions and groundings with subsequent OSP. In addition, voluntary pollution of oil involving washing tanks must be added. The likelihood β of a pollution event involving oil spill in a scenario α can be evaluated also using comprehensive historical oil spill pollution data. While depending on data availability, different sources for OSP formation, such as incidents or intentionally (human voluntary, washing oil tanks, etc.) must be considered. A metric to measure consequences depends on preferably a preset definition of risk suited for the problem context in question. Essentially, the main consequence is the impact of oil spill slick on the coast, so the near real time localization of the oil slick, its potential impact on the coast can be evaluated by the degree of gravity of the same oil slick. In our ARGO-G model, we have simulated the consequence using an advection-diffusion model that gives the velocity of the oil slick coast reaching as a function of the chemical parameters when known. The results of the oil spills analysis may be further separated into multiple categories, such as, crude oil, refined products, bunker fuel, and diesel fuel. Crude oil and bunker fuel are less volatile and typically display a more environmentally persistent behavior than refined products and diesel fuels.

In principle, one arrives at a metric of overall baseline system risk using the complete set of triples $\{\alpha,\beta,\gamma\}$ by evaluating

$$R_\alpha = \sum_i \beta_i \times \gamma_i \qquad (1)$$

where the summation (1) is conducted over the various incident types and accident types being considered in any specific scenario α. The variation of some specific parameters changes the scenario α and, as a consequence, the likelihood of OSP events.

The ARGO-G model presents a particular effort in the evaluation of the dynamic risk using geopositioning tools such as Automatic Information System (AIS) and radar data to follow the oil tanker routes as well as any other crossing ship. AIS has been used for quite some time in aviation, but its use is becoming more prevalent as a navigational tool on board of vessels as well. At set intervals, it automatically transmits the position of the vessel along with a time stamp and vessel identification to an AIS data repository. Hence, already available radar data is more frequently supplemented with AIS data. However, as with any data recording process, raw data, being it radar or AIS, has errors within it that either occur at the transmission end or at the receiving end. As a consequence, a specific algorithm must be constructed to follow the route localization and manage a great amount of data. Each ship transit may include thousands of points, and the computational effort required to calculate movements of vessels in the simulation increases with the number of n points along a route. Hence, we must attempt to keep n pairs data per transit as low as possible while maintaining a reasonable curvature of vessel routes along the waterways. In addition, the real time localization of the ships crossing in a given moment the marine areas must be synchronized with the information characterizing any ship, so that a risk factor for oil spill pollution can be associated to any ship using inferential statistical methods as mentioned before. All the details including the algorithms will be available on the Argomarine website as soon [2]. In the next section we present a Bayesian approach for the definition of an OSP.

III. A Bayesian approach for the definition of OSP event Probability

According to Bayes rule, the probability of occurrence of a specific event X is affected by the fact of another event to have happened or not. Thus, it is necessary to calculate the occurrence of X conditioned to the previous occurrence of Y, denoted by $p(X|Y)$ (probability of X given Y), where $p(X|Y)$ is given by the following expression

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)} \qquad (2)$$

where the implementation is made considering the dependency of measures between two branching levels. At this point, we have to build a cumulative density function relative to $p(X|Y)$, once that the occurrence of X depends on the Y. If one is interested in paired comparison of accident risk between two different functions, **M** and **L**, it is sufficient to estimate the parameter vector, **C**, as the relative possible OSP probability that can be defined as

$$P(\mathbf{S}|\mathbf{C}) = \exp\{\mathbf{C}^T\mathbf{S}\} \qquad (3).$$

Now, we have to build the likelihood of a response on the possibility of OSP when both **M** and **L** are known. Two possible scenarios described by the supplemental variables **S** are introduced, and let X defined as

$$X = P(\mathbf{C}|\mathbf{S}_1)/P(\mathbf{C}|\mathbf{S}_2) \qquad (4).$$

The response to the level of the knowledge on the possible production of OSP in a date location can be considered as normally distributed such that $(Z|\mu,r)\sim N(\mu,r)$, where $Z=\log X$ and $r=1/\sigma^2$ is the precision confidence and σ is the standard deviation of the normal distribution in (4) and $\sigma>0$. We can redefine the **C** vector as $\mu=\mathbf{S}^T\mathbf{C}$, so that the likelihood of the confidence knowledge can be rewritten as:

$$L(z)\propto\sqrt{r}\exp\left\{-0.5r(z-\mu)^2\right\} \quad (5).$$

Suppose to have n different levels of knowledge of the **M** function, so that the decision support require the interrogation of the skills linked to the **C** vector, so that we need to define a vector $\boldsymbol{q}_j=(\boldsymbol{M}_j-\boldsymbol{L}_j)$, $\boldsymbol{j}=1,\ldots,n$ and a matrix $p\times n$, $\mathbf{Q}=[\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n]$. As a consequence the likelihood (5) becomes [7]:

$$L(z_1,\ldots,z_n|\mathbf{C},r,\mathbf{Q})\propto r^{n/2}\cdot \exp\left\{-0.5r\left(\sum_{j=1}^{n}z_j^2-2\sum_{j=1}^{n}\mu_j z_j+\sum_{j=1}^{n}\mu_j^2\right)^2\right\} \quad (6)$$

To allow for a conjugate Bayesian analysis, a prior distribution is proposed for the joint distribution of (**C**, r). Following the West and Harrison approach for similar problems, such as ship incidents in limited harbor systems, a multivariate normal-gamma prior can be proposed [7]. Then the distribution of $(\mathbf{C}|r)$ is assumed to be multivariate normal with a prior $p\times1$ dimensional mean vector $\boldsymbol{m}$ and a $p\times p$ matrix:

$$\Pi(\mathbf{C}|r)\propto r^{n-1/2}\exp\left(-\frac{r}{2}\right)\times\exp\left\{-\frac{r}{2}(\mathbf{C}-\mathbf{m})^T\Delta(\mathbf{C}-\mathbf{m})\right\} \quad (7)$$

where $(r\Delta)^{-1}$ is the variance covariance. Applying Bayes theorem utilizing the likelihood (6) to the prior distribution, (7), it follows that the posterior distribution $\Pi(\mathbf{C},r|L,\mathbf{Q})$ is proportional to (8). The posterior distribution (8) can be updated and represented on a grid, where to any cell can be assigned a numerical value that can be considered as the β factor, i.e. the OSP likelihood. To obtain the dynamic risk, we have to calculate the impact factor of an OSP, γ, when a pollution event is occurred. A possible evaluation of γ will be made in the next section.

$$\Pi(\mathbf{C}|r)\propto r^{n-1/2}\times\exp\left\{-\frac{r}{2}\left(\sum_{j=1}^{n}z_j^2-2\left[\sum_{j=1}^{n}q_j z_j\right]^T\mathbf{C}+\mathbf{C}^T\left[\sum_{j=1}^{n}q_j q_j^T\right]\mathbf{C}\right)^2\right\}\times\exp\left(-\frac{r}{2}\right)\exp\left\{-\frac{r}{2}(\mathbf{C}-\mathbf{m})^T\Delta(\mathbf{C}-\mathbf{m})\right\} \quad (8)$$

IV. EVALUATION OF THE IMPACT FACTOR FOR OSP EVENTS

The dynamic risk, R, requires the quantitative knowledge of the impact of an OSP on the coasts, that in (1) is given by the parameter γ. The consequence of an OSP, γ, must be defined considering a series of specific variables meanly, the size of the soil spill slick, the chemical properties of the oil, the distance from the coastal, the environmental conditions. All such ingredients must be inserted in a stochastic model because the random nature of the variables mentioned before. The definition of impact consequence can be handle basing our analysis on the well-known advection-dispersion mathematical model. One possible choice it to start considering the rate of change of the oil spill concentration under some time changing environmental conditions (marine currents, wind intensity, etc.) on the marine surface (bi-dimensional case). If the OSP event is large size, we can divide its dimension in different k portions, where any portion subject to slightly different environmental conditions, if these are known. The advection-dispersion model for the evolution of the oil spill concentration slick χ can be described by the following partial derivate equation (PDE) [8]

$$\frac{\partial\chi_i}{\partial t}+u\frac{\partial\chi_i}{\partial x}+v\frac{\partial\chi_i}{\partial y}=D_x\frac{\partial^2\chi_i}{\partial x^2}+D_y\frac{\partial^2\chi_i}{\partial x}+E_i \quad (9)$$

where χ_i is the oil concentration in the i-th segment, where $i=1,\ldots,k$, u and v are the water velocities (m/s) in the x,y directions, D_x and D_y are the dispersion coefficients in the x and y directions (m^2/s), and E_i describes the chemical reaction of the oil spill with water. Generally, one limited information is available on the oil spill concentration and dispersion coefficients. This type of uncertainty can be taken into account by considering both such quantities like fuzzy numbers. This implies that such unknown quantities at any time and position will behave as fuzzy numbers [8]. They will follow the advection dispersion partial differential equation (9). Although derivatives of fuzzy variables exist, there is no unique solution to (9) with fuzzy variables, because fuzzy numbers take different values at different levels of confidence. The solution is given taking any fuzzy numbers represented by a discrete set of h-level cuts, and for every confidence level h, we look only for the lower and the upper limiting values of the unknown fuzzy variables. A solution of (9) in the ordinary intervals for the fuzzy variables can be found using finite differences and finite elements [8]. Here, we write a possible solution that can be expressed as:

$$\chi(x,y,t)=\chi_0\left[e^{T_1}\left[1-erf\left(Z_1\right)+e^{T_2}\left[1-erf\left(Z_2\right)\right]\right]\right] \quad (10)$$

with $T_{1,2}=\left(u\mp\sqrt{u^2+4iD_x}\right)x\Big/2D_y$ and $Z_{1,2}=\left(u\mp\sqrt{u^2+4iD_x}\right)t\Big/\sqrt{2D_y t}$, where the number indexes, 1,2 are respectively correspondent to the signs - and +. Once χ is known, the parameter γ can be calculated taking in account the capacity of intervention by the Coast Guard or other correspondent institution. United to the β factor calculated with the Bayesian approach, the dynamic risk can be quantified, figure 1.

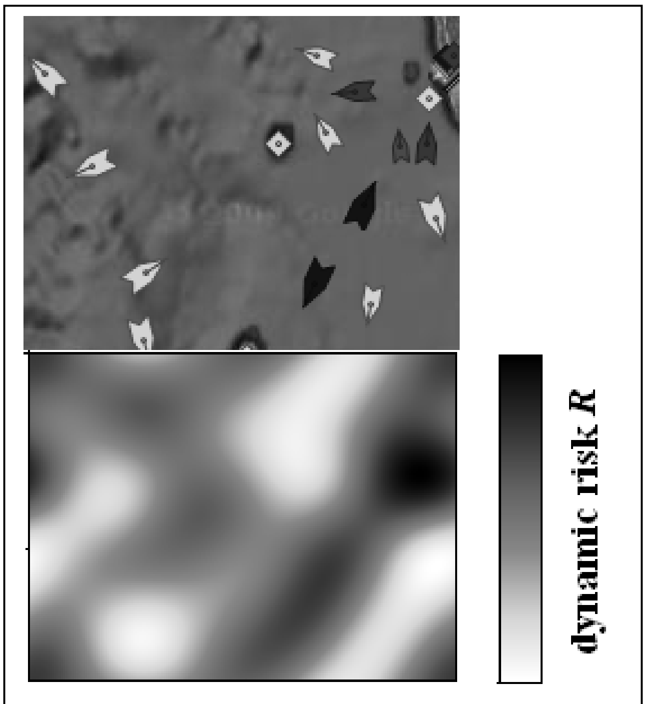

Fig. 1 Example of operative results obtained mapping the dynamic risk *R* factor as given by the ARGO-G model. On the upper image it is represented the marine area to be monitored with the representation of the load function, **L**. Such function is obtained taking in account the density of ships crossing the marine area. On the lower image, the correspondent scheme produced by the ARGO-G predictive algorithm. The dark domains on the marine area represent the localization of the critical sub-areas where an OSP event is more probable (and high damage produced) and efforts on the monitoring activity are required.

In figure 1, we give a schematic sketch of the dynamic risk mapped using the Bayesian approach and the advection model for the propagation of OSP. The cargo tankers are considered as the main sources for OSP, nevertheless, coastal activities could have an important role for the production of OSP, and their inclusion in the present model is one of the future steps for improving the ARGO-G model.

IV. Conclusive Remarks

A near real time surveillance of oil spill on large marine areas requires a continuous monitoring power able to combine remote sensing with Decision Support or Decision Making methods for active actors operating in the marine areas. In this paper, it has been presented and described a Geomatrix model (ARGO-G) within the activity of the Argomarine project. The model has been jointed with the main geopositioning systems for monitoring marine areas for a real time surveillance of oil spill pollution events. The effective optimization of remote sensing power focused on the higher oil spill risks is the main basic requirement for the ARGO-G model. Based on the concept of dynamic risk, the ARGO-G model used inferential statistical tools and its versatility make it suitable for an immediate application to general contexts of environmental control and monitoring.

Acknowledgment

The Argomarine Project, 7th Seventh Framework Program Grant # 234096, is warmly acknowledged.

References

[1] A. Jernelöv, “How to defend against future oil spills”, *Nature* 466, pp. 182-183, July 2010

[2] www.argomarine.eu

[3] S. Kaplan, “The words of risk analysis,” *Risk Analysis*, vol. 17, no. 2, pp. 407-417, March 1997.

[4] J.R. van Dorp, J.R.W. Merrick, “On a risk management analysis of oil spill risk using maritime transportation system simulation,” *Annals of Operations Research*, available on line 12 December 2009, *http://www.seas.gwu.edu/~dorpjr/Publications/JournalPapers/ANOR2009-2.pdf.*

[5] P. Szwed, J.R. van Dorp, J.R.W. Mererick, T.A. Mazzuchi, and A. Singh, “A Bayesian paired comparison approach for relative accident probability with covariate information,” European Journal of Operative Research, vol. 169, pp. 157-177, 2006.

[6] M.S. Eide, Ø. Endresen, P.O. Brett, J.L. Ervik, K. Røang, “Intelligent ship traffic monitoring for oil spill prevention: Risk based decision support buildings on AIS,” *Marine Pollution Bulletin*, vol. 54, pp.145-148, 2007.

[7] M West and J. Harrison, *Bayesian Forecasting and Dynamic Models*, Springer-Verlag, 1989.

[8] H. Mpimpas, P. Anagnostopoulos, and J. Ganoulis, “Modelling of toxicants distribution in Thermaikos Gulf,” in Proceedings of the Fifth International Conference on Hydroinformatics, Cardiff, UK, IWA Publishing, pp. 457-463, 2002.

[9] A. Pownuk, “Numerical solutions of fuzzy Partial Differential Equations” in Fuzzy Partial Differential Equations and Relational Equations and its application in Computational Mechanics,” Nikravesh, Zadeh, Kroptkin, eds, Springer-Verlag, vol. 142, pp.308-347, 2004.