The Two C's of the Risk-Based Approach to Goal Based Standards: Challenges and Caveats

Christos A. Kontovas [1)], Harilaos N. Psaraftis [2)], Panos Zachariadis [3)]

[1)] Laboratory for Maritime Transport, National Technical University of Athens
Athens, Greece, xkontovas@gmail.com
[2)] Laboratory for Maritime Transport, National Technical University of Athens
Athens, Greece, hnpsar@deslab.ntua.gr
[3)] Atlantic Bulk Carrier Management Ltd.
Piraeus, Greece, engineering@atlanticbulk.gr

Abstract

While it is generally accepted that the overall level of maritime safety has improved in recent years, further improvements are still desirable. It is fair to say that much of maritime safety policy worldwide has been developed in the aftermath of serious accidents (such as 'Exxon Valdez', 'Estonia', 'Erika' *and* 'Prestige'*). Industry circles have questioned the wisdom of such an approach. The safety culture of anticipating hazards rather that waiting for accidents to reveal them has been widely used in other industries such as the nuclear and the aerospace industries.*

The international shipping industry has begun to move from a reactive to a proactive approach to safety through "Formal Safety Assessment" (FSA) and "Goal Based Standards" (GBS). FSA was introduced by the IMO as "a rational and systematic process for accessing the risk related to maritime safety and the protection of the marine environment and for evaluating the costs and benefits of IMO's options for reducing these risks". The recent GBS approach aims to be another proactive instrument, and there has been recent discussion in the IMO on the possible links between FSA and GBS.

This paper attempts to clarify some widely used, but confusing to many, notions such as Risk Based Rulemaking vs. Risk Based Design, and IMO's GBS Traditional Approach vs. Safety Level Approach, and the implications of their use, or misuse, to future ship rulemaking, design and safety. The paper elaborates on some identified weaknesses of the risk based approach which must be corrected, with an emphasis on environmental risk evaluation criteria, which is an area in which further research is deemed necessary.

Keywords: Formal Safety Assessment; Goal Based Standards; Risk-Based GBS; Safety-Level Approach to GBS; Environmental Risk Evaluation Criteria.

1. Introduction

Much of the recent debate at the IMO and in other regulatory fora centres on a set of questions that deal with the possible use of the so-called "Safety Level Approach" (SLA) in modern rule-making and design. SLA is also known as the Risk Based approach, and involves the use of probabilistic tools and techniques in the formulation of regulations and in the actual design of ships. Examples of questions that are raised within the SLA debate are: Should SLA be used within the new Goal Based Standards (GBS) framework? Should GBS be risk based? Should Formal Safety Assessment (FSA) be used within GBS? Should Structural Reliability Analysis (SRA) be used within GBS? And so on.

Such questions, if posed this way, do not address the right issue. As there is no doubt that modern maritime safety rulemaking should use the concept of risk, along with all tools developed to study it, most of the above questions do not really concern "if", but rather, "how" and "when". This paper attempts to shed some light on these issues, by clarifying some widely used, but confusing to many, notions such as Risk Based Rulemaking vs. Risk Based Design, and IMO's GBS Traditional Approach vs. Safety Level Approach, and the implications of their use, or misuse, to future ship rulemaking, design and safety.

The paper elaborates on some identified weaknesses of FSA and the Risk Based approach which must be corrected. It further cautions on the over eagerness of some rule makers and designers to drop all prescriptive rule formulations and haphazardly adopt risk based formulations borrowed from other industries which may not be appropriate for ships. A reliable risk based approach involves avoidance to cut corners and thus avoidance on relying on a large number of arbitrary assumptions. To be applied properly, the risk based approach requires a significant amount of future research in order to reliably link from first principles the ship risk model with the desired acceptable Risk or Safety level.

To do so, the rest of the paper is structured as follows. Section 2 provides some background and focuses on proactive regulation and Risk-Based rulemaking and design. Section 3 describes FSA and GBS within IMO's regulatory process. Section 4 discusses the issue of a tolerable risk level and Section 5 highlights some environmental criteria deficiencies. Section 6 refers to a possible environmental risk index. Section 7 focuses on environmental risk acceptance criteria and Section 8 on the cost of averting a spill criterion. Finally, Section 9 presents the conclusions of the paper.

2. The Need to Be Proactive; Risk-Based Rulemaking vs Risk-Based Design

While it is generally accepted that the overall level of maritime safety has improved in recent years, further improvements are still desirable. However, it can be argued that much of maritime safety policy worldwide has been developed in the aftermath of serious accidents (such as *'Exxon Valdez', 'Estonia', 'Erika'* and *'Prestige'*. The safety culture of anticipating hazards rather that waiting for accidents to reveal them has been widely used in other industries such as the nuclear and the aerospace industries. The international shipping industry has begun to move from a reactive to a proactive approach to safety through what is known as 'Formal Safety Assessment' (FSA). The recent 'Goal Based Standards' (GBS) approach aims to be another proactive instrument, and there has been recent discussion at the IMO on the possible links between FSA and GBS (see, for instance, IMO document MSC 81/6/16, among others[1]).

GBS started as an attempt of IMO to better structure its regulatory process by use of a tier system where high level goals are at the top and the functional requirements necessary to achieve the goals follow. The first development started with the subject of hull design and construction of bulk carriers and oil tankers for two reasons. a) IMO wanted to have a stronger input into the regulations for the construction of ships, which traditionally were left to the classification societies. b) tankers and bulk carriers were chosen first due to their increased structural defects.

Soon a difference of opinion ensued with regard to how these standards should be developed. Many argued that the standards should follow the risk based approach for which FSA is suited and which specifies a safety level to be achieved and the proper methodology to be followed. Within the proponents of the risk approach there are further differences of opinion as to whether the method should include specific acceptance criteria or not and who will develop these; IMO or the classification societies which write the rules in detail? The proponents for few criteria argue that this aids design innovation without posing many restrictions. The opponents argue that just specifying the methodology without enough specific requirements (criteria) allows unlimited latitude so that even unsafe designs can appear to comply.

Those not favouring altogether the Risk Based approach, argued that at least for tankers and bulk carriers the huge accumulated practical experience should be the primary guide, with the standards developed being the direct result of such experience. They also argued that the problems to be fixed on these types of ships are urgent whereas the risk level approach needs many years to be developed and is more appropriate for "high technology" ships whose design has not solidified over the years. Therefore they urged to continue the "traditional" rulemaking approach which includes a mix of statistical formulations, formulations from first principles and empirical prescriptive formulations. In the end, recognizing the urgency to improve the construction standards of tankers and bulk carriers, it was decided that both approaches are developed in parallel and independently.

It should be noted however that in practice the two approaches are related and closer than most people think. The requirements that one group considers necessary "from experience" should also be evident following the risk based approach, *provided it is done properly*.

The above phrase "provided it is done properly" encompasses some large and serious issues going to the core of the debate between the proponents of the two approaches. The above described debate within IMO concerned Risk Based *Rulemaking* (not Risk Based *Design*, which is a different concept). Namely, the formulation of regulations by using the Risk Based –or Safety Level- approach (using FSA, Risk analysis, setting goals but also clear and specific functional requirements and criteria that must be complied to achieve the goals –while not totally abolishing prescriptive rules which are known to be effective).

While such approach to set rules and regulations has been successfully used in rulemaking in the past and continues to expand by the use of FSA and Goal Based Standards, the leap to Risk Based *Design* of ships is a long one. Risk Based Design involves designing ships or their arrangements in alternate to standard "Rule acceptable" methods and results and "proving" that they are equivalent or better to standard rule designs. To do this the Safety level of the new design must be shown to be equivalent to the Safety level of the standard Rule. At this stage, (where, at least for structures, neither the Rule safety level nor the alternate design safety level are

[1] In this paper we cite IMO documents by using the standard code for MSC (MEPC) publications: MSC (MEPC) x/y/z, where x: session; y: agenda item; z: document number of agenda item.

known or can be safely assessed) this seems to many like wishful thinking. The road to assess current rules and designs safety levels and come up with widely acceptable methods to "compare" new innovative designs to standard ones is a long one involving years of research in order to devise a reliable model and link the top tier "safety level" of the design to its individual detailed structural arrangements. The urge therefore to shorten that process and cut corners is strong and manifests itself in many ways, from haphazardly borrowing inapplicable methods and results from other industries to utilizing a huge number of arbitrary assumptions to avoid the research.

It should be noted that both the Traditional and the Risk or Safety level approaches can be, and have been, misused. With the traditional or empirical approach, in the late 80's, scientific or engineering assessments were not always rigorously performed in favour of ever lighter and super-optimized ships, whereas many recurring design defects becoming apparent soon after building were being ignored and not corrected (by regulation or otherwise) mainly due to builders' resistance to redesign or impose "more steel" on their designs. The new IACS Common Structural Rules are also based on the traditional approach of rulemaking (i.e. combination of first principles formulations, probabilistic and semi-experiential formulations etc.) but in many areas it seems that not all lessons of the past have been learnt.

Prescriptive regulations have also been misused in the past by applying different interpretations amongst the various classification societies while designers, always looking for loopholes to construct their ships faster and cheaper, many times are able to bypass even totally prescriptive rule requirements. Sometimes a missing or misplaced comma could make a world of difference.

Critics of the Risk Based approach argue that if a totally prescriptive regulation can be misused or bypassed by smart designers, this will be 1,000 times more likely with a non-prescriptive approach calling it "the ultimate loophole". FSA can, willingly or not, point to false directions and many times it is difficult, even for FSA experts, to verify its correctness or validity. Furthermore, the Risk Based approach *to design* aims at opening the door to designers to demonstrate an "equivalent Safety Level" to that prescribed by the rules. As mentioned above, critics call that "official bypassing of the Rules" and admittedly it would be currently impossible to transparently demonstrate "equivalency" when the actual rule Safety Levels are unknown and cannot even be determined at present prior to extensive research toward that goal. Thus assumptions, conjectures, misinterpretations, oversimplifications and a lot of grey areas are prevalent in these first steps of the Risk Based approach.

There is no doubt however that the Risk Based approach is the way of the future to provide effective alternatives and aid innovation. It just has to be developed properly and carefully so it does not develop into "the ultimate loophole". It must always be checked and compared against gained practical experience and it should result in clear "prescriptive-like" formulations which the designers can easily use.

In fact, there are five challenges to which any risk based approach to modern maritime safety regulation must respond. It has to be:

- Proactive – as mentioned above, anticipating hazards, rather than waiting for accidents to reveal them which would in any case come at a cost in money and safety (of either human life or property i.e. the ship itself)
- Systematic – using a formal and structured process
- Transparent – being clear and justified of the safety level that is achieved
- Cost-Effective – finding the balance between safety (in terms of risk reduction) and the cost to the stakeholders of the proposed risk control options.
- Where possible calibrated to known experience.

The need for proactivity has been argued extensively time and again (among others, see Psaraftis (2002) before *'Prestige'* and Psaraftis (2006) after *'Prestige'* for an analysis of the main issues). FSA has long been considered the prime scientific tool for the development of proactive safety regulation. GBS is another, more recent tool, toward the same goal.

3. Formal Safety Assessment (FSA) and Goal Based Standards (GBS)

FSA was introduced by the IMO as *"a rational and systematic process for accessing the risk related to maritime safety and the protection of the marine environment and for evaluating the costs and benefits of IMO's options for reducing these risks"* (see FSA Guidelines in document MSC Circ. 1023, MEPC Circ. 392). In MSC 81 (May 2006), an FSA 'drafting group' proposed some amendments to these guidelines (see Annex 1 of document MSC 81/WP.8). These amendments have been approved by the MSC and were subsequently sent on to the MEPC for approval, which happened at MEPC 55 (October 2006).

To achieve the above objectives, IMO's guidelines on the application of FSA recommended a five-step approach, consisting of:

1. Hazard Identification
2. Risk Assessment
3. Risk Control Options
4. Cost-benefit Assessment
5. Recommendations for decision making

The notion of GBS was introduced in IMO at the eighty-ninth session of the Council in November 2002 through a proposal by the Bahamas and Greece (IMO document C 89/12/1). This document suggested that IMO should play a larger role in determining the standards to which new ships are built. After consideration by MSC, the biannual IMO Assembly, at its twenty-third session, adopted the Strategic Plan for the Organization –the six year period 2004 to 2010- and, inter-alia, resolved that "the IMO would establish goal based standards for the design and construction of new ships" (resolution A.944(23)). Following this resolution, the Maritime Safety Committee established a correspondence group to progress work.

Discussions in plenary and in the working group during MSC79 and MSC80 (see document MSC80/24) resulted in agreement on the basic principles as follows:

"IMO goal-based standards are:

1 broad, over-arching safety, environmental and/or security standards that ships are required to meet during their lifecycle;

2 the required level to be achieved by the requirements applied by class societies and other recognized organizations, Administrations and IMO;

3 clear, demonstrable, verifiable, long standing, implementable and achievable, irrespective of ship design and technology; and

4 specific enough in order not to be open to differing interpretations."

Following a proposal by the Bahamas, Greece and IACS at MSC 78 (document MSC 78/6/20) a five-tier system was agreed. It was, also, agreed that the first three tiers constitute the goal-based standards to be developed by IMO, whereas Tiers IV and V contain provisions and detailed rules developed/to be developed by classification societies, other recognized organizations and industry organizations. It has to be noticed that, traditionally, the determination of the standards to which new ships are built is the responsibility of classification societies and shipyards. IMO wants to play a larger role in this determination, however, acknowledges the roles of Class and shipyards.

• Tier I: *Goals*
A set of goals to be met in order to build and operate safe and environmentally friendly ships.

• Tier II: *Functional requirements*
A set of requirements relevant to the functions of the ship structures to be complied with in order to meet the above-mentioned goals.

• Tier III: *Verification of compliance criteria*
Provides the instruments necessary for demonstrating that the detailed requirements in Tier IV (Rules) comply with the Tier I goals and Tier II functional requirements.

• Tier IV: *The detailed rules which apply the functional requirements to satisfy the goals.*
The detailed mandatory requirements developed by IMO, National Administrations and/or classification societies and applied by national Administrations and/or classification societies acting as Recognized Organizations to the design and construction of a ship in order to meet the Tier I goals and Tier II functional requirements.

• Tier V: *Industry standards, guidelines, recommendations, codes of practice and safety and quality systems for shipbuilding, ship operation, maintenance, training, manning, etc.*
Industry standards and shipbuilding design and building practices that are applied during the design and construction of a ship.

Goal-based standards are not a completely new concept in the work of IMO. Over the past years, the IMO has introduced goal-based standards for certain subjects such as fire protection, even though not in a systematic way. The revised SOLAS Ch, II-2 contains a regulation on alternative design and arrangements (regulation 17) which allows deviation from the prescriptive requirements:

" 2.1 Fire safety design and arrangements may deviate from the prescriptive requirements set out in parts B,C,D,E or G, provided that the design and arrangements meet the fire safety objectives and the functional requirements".

However, the proposal of the Bahamas and Greece brings IMO rule-making to a new era. The high importance of GBS is understood by the fact that GBS is included in the Strategic Plan for the IMO (Assembly resolution A.944(23)) and in the Long-term Work Plan (Res. A.943(23)).

It is also understood that basic principles of GBS were developed to be applicable to all goal-based standards and not only goal-based new ship construction standards, in recognition that, in the future, IMO may develop goal-based standards for other areas, e.g. machinery, equipment, operation, maintenance, fire-protection, etc.

As stated before, much of the recent debate in many regulatory fora centers on a set of questions that deal with the possible use of the Risk-Based or "Safety Level" Approach (SLA) in modern rule-making and design. Also it was mentioned that questions such as whether SLA should be used within GBS, or whether FSA will be used within GBS, and so on, are misleading, as most of these questions do not really concern "if", but rather, "how" and "when".

Given that FSA is currently used for proposed new rules and will be eventually used within the Risk Based Approach to GBS, one question is, are there potential deficiencies that should be corrected before anything like this is attempted. The answer is surely yes. For a detailed analysis of possible deficiencies and what can be possibly done to overcome them see Kontovas (2005), Kontovas and Psaraftis (2006), document MSC82/INF.3 (submission of Greece to MSC 82 on FSA), Zachariadis et al (2007) and document MEPC 56/18/1 (submission of Greece to MEPC 56)[2]. These papers elaborate on a variety of identified weaknesses of FSA and the Risk Based approach which must be corrected. They further caution on the over eagerness of some rule makers and designers to drop all prescriptive rule formulations and haphazardly adopt risk based formulations borrowed from other industries which may not appropriate for ships. A reliable risk based approach involves avoidance to cut corners and thus avoidance on relying on a large number of arbitrary assumptions. To be applied properly, the risk based approach requires a significant amount of future research in order to reliably link from first principles the ship risk model with the desired acceptable Risk or Safety level.

In the remainder of this paper we focus on some issues that in our opinion merit attention.

4. What is a Tolelable Risk Level?

The ultimate step of FSA aims at giving recommendations to the relevant decision makers for safety improvement taking into consideration the findings during all four previous steps.

The risk control options (RCOs) that are being recommended should

- ✓ Reduce Risk to the “desired level”
- ✓ Be Cost Effective

The IMO Guidelines suggest that both the Individual and Societal Types of risk should be considered for crew members, passengers and third parties. Individual Risk can be regarded as the risk to an individual in isolation while Societal Risk as the risk to the society of a major accident – an accident that involves or affects more than one person. In order to be able to analyze further these categories of risk and their acceptance criteria, we must have a look at the levels of risk.

According to Health and Safety Executive’s (HSE, United Kingdom) Framework for the tolerance of risk, there are three regions in which risk can fall into (HSE, 2001). Unacceptable Risk (for example resulting from high accident frequency and high number of fatalities) should either be forbidden or reduced at any cost.

Between this region and the Acceptable Risk region (where no action to be taken is needed) the ALARP (As Low As Reasonable Practicable) region is defined. Risk that is falling in this region should be reduced until it is no longer reasonable (i.e. economically effective) to reduce the risk. Acceptance of an activity whose risk falls in the ALARP region depends on cost-benefit analysis.

These regions are illustrated in the following figure.

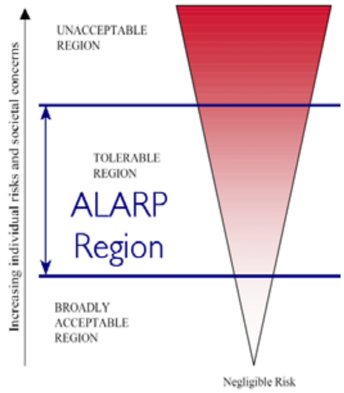

Fig. 1: The ALARP Concept

Incredible as it may seem, there is still no single universal level of acceptable individual risk, either at IMO, or at any other rule-making body. IMO’s guidelines provide no official Risk Acceptance Criteria and currently decisions are based on those published by the UK Health & Safety Executive (HSE, 1999). HSE’s criteria define the intolerable and the negligible risk for a single fatality as follows:

Maximum tolerable risk for crew members	10^{-3} annually
Maximum tolerable risk for passengers	10^{-4} annually
Maximum tolerable risk for public ashore	10^{-4} annually
Negligible risk	10^{-6} annually

Risks below the tolerable level but above the negligible risk (for crew members, passengers and third parties) should be made ALARP by adopting cost-effective RCOs.

We first note that in the recently adopted amendments to the FSA guidelines (see Annex 1 of document MSC 81/WP.8), it was made clear that all of these numbers are only indicative. Therefore, the crucial issue of what are acceptable risk criteria for the safety of maritime transport is still very much open.

More fundamentally, we further note that the expression of these risk limits on an *annual* basis (instead, for instance, on a *per trip* basis) does not account for the number of trips per year undertaken by a person who travels by ship, a number that may vary significantly and one that surely would influence the level of risk someone is exposed to (if someone does not travel by ship, obviously this risk is exactly zero). The ratio of 10 to 1 between the maximum tolerable risk for crew members vis-à-vis the equivalent risk for passengers implicitly assumes that the former category makes roughly 10 times more trips than the latter, for the acceptable risk to be equivalent on a per trip basis (also the crew takes the risk of their job willingly, so pre-

[2] In addition, a shorter version of this paper will be presented at the PRADS 2007 conference in Houston (Kontovas et al, 2007).

sumably they would be willing to tolerate risk more than others).

Another comment is that these risks, as formulated this way, seem to compare unfavorably to air transport, in which the most recently estimated probability of being involved in a fatal air crash is about 1 in 8 million per flight for ‘First World’ airlines (Barnett, 2006). This means that a maritime transport passenger is allowed an annual risk which is 100 times higher than that of an airline passenger who takes an average of 8 flights during the year (or, one roundtrip every 3 months), or even more than 100 times higher, when comparing with less frequent air travelers. Among some, such a comparison might raise the question if maritime transport travelers are second-class citizens as compared to air transport ones.

In any event, it is clear that additional analysis is necessary to define risk acceptance criteria and to ascertain if a better ‘risk exposure variable’ can be found in maritime transport. If the expression of tolerable risk on an annual basis may present problems, as noted above, the fact that the number of flights (trips) was chosen as the most appropriate exposure variable for air transport does not necessarily mean that this should be adopted for maritime transport as well. Variables such as journey length or journey time may be more relevant for shipping, and this is something that should be examined.

5. Environmental Criteria Deficiencies

We now come to a subject that is very important for maritime safety but for which the current state of knowledge is lacking and there is urgent need for new knowledge and analysis, and particularly for the SLA approach to GBS.

In all recent FSA studies, cost effectiveness is limited to covering fatalities from accidents and implicitly, also, injuries and/or ill health from them. However, thus far no FSA study has tried to assess environmental risk. Lately, the IMO tried to deal with this aspect (see for instance documents MSC 81/18 and MEPC 55/18) and made reference to a recent report from project SAFEDOR co-funded by the European Commission (Skjong et al, 2005). Much analysis is reported there, and the report properly identifies the difficulties to arrive at a single environmental criterion. Environmental damage and clean up costs vary tremendously depending on which part of the world the spill occurred and furthermore data is available mostly from spills in developed areas of the world where of course clean up costs are high. But in the end this report implies a figure as high as \$60 000 as the so-called ‘Cost of Averting one Tonne of Spilled oil’ (CATS), for which more below. However, as a broad multitude of factors enter into damage estimation of oil pollution, the adoption of any single figure as the per tonne cost of oil spills is bound to be problematic, particularly as regards regulatory policy formulation. For more comments on this see Kontovas and Psaraftis (2006) and the initial reaction of Greece to this approach in MSC 81, urging caution on the matter (document MSC81/18/2). Also Japanese submission document MSC81/6/3 includes the results of several prior studies as reported by the International Ship and Offshore Structure Congress which would shed serious doubt on any metric that consists only of volume of oil spilled and reported clean up costs.

The IMO has adopted a similarly cautionary stance on this issue, with MSC 81 turning the matter over to MEPC. In MEPC 55, an invitation was issued to *“members and international organizations to consider the draft environment risk evaluation criteria during the intersessional period and submit comments thereon to MEPC 56, for further consideration prior to referring the agreed text to the MSC for appropriate action.”* (see also IMO documents MEPC 55/18, MEPC 55/23, MSC 82/24 and MEPC 56/18). In response to this invitation, Greece submitted document MEPC 56/18/1 on FSA, with a focus on environmental risk evaluation criteria. After discussion, in MEPC 56 (July 2007) it was agreed to form a ‘correspondence group’, coordinated by the second author of this paper, and tasked to look into the matter in more detail and report back in time for MEPC 57 (April 2008).

The rest of this paper draws much from Greece’s recent submission to the MEPC (document MEPC 56/18/1), highlighting what we believe are some of the most important issues that are at stake so as to adopt sensible criteria relevant to the protection of the marine environment.

6. Environmental Risk Index

In FSA, the explicit consideration of the frequencies and of the consequences of hazards is typically carried out by the so-called risk matrices. These are used to rank risk in order of significance. A risk matrix divides the dimensions of frequency and consequence into categories. Each hazard is allocated to a frequency and consequence category and the risk matrix then ranks the risk that is associated with that hazard.

Frequency Index

FI	FREQUENCY	DEFINITION	F (per ship year)
7	Frequent	Likely to occur once per month on one ship	10
5	Reasonably probable	Likely to occur once per year in a fleet of 10 ships, i.e. likely to occur a few times during the ship’s life	0.1
3	Remote	Likely to occur once per year in a fleet of 1000 ships, i.e. likely to occur in the total life of several similar ships	10^{-3}
1	Extremely remote	Likely to occur once in the lifetime (20 years) of a world fleet of 5000 ships.	10^{-5}

The above currently used Frequency Index (MSC circ. 1023, MEPC circ. 392) can also be used for assessing environmental risk. However, the current Severity Index deals only with the effects on human safety (injuries or fatalities) and the ship itself. In order to be able to measure the effect on the environment, a proper Severity Index that measures effects on the environment has to be defined. This is not a trivial task.

It is our opinion that all FSA steps should be looked at carefully from an environment protection standpoint, and this includes the adoption of a proper environmental Risk Index, and specifically one that produces no distortions in ranking environmental hazards. In fact, the objectives of the first step of every FSA (Hazard Identification- HAZID) are (a) to identify all potential hazardous scenarios which could lead to significant consequences, and (b) to prioritize them by risk level.

Annex 3 of document MEPC 55/18 offers no explicit proposal for an environmental Risk Index, or something equivalent, but it refers to the aforementioned SAFEDOR report. As a basis for further discussion, we propose that the following Severity Index, which, as stated in the SAFEDOR report, is based on NORSOK Standards Z-013, be analyzed and debated:

SI	Severity	Effect on the environment (recovery time)
1	Minor	Between 1 month and 1 year
2	Moderate	Between 1 and 3 years
3	Significant	Between 3 and 10 years
4	Serious	In excess of 10 years

By combining the Frequency and Severity Indices, the Risk Index could be defined the same way as in the current FSA Guidelines:

Risk Index = **F**requency **I**ndex + **S**everity **I**ndex.

Note that if the aforementioned four category Severity Index is used, the resulting risk matrix is the same as the one currently in use (MSC circ. 1023, MEPC circ. 392).

7. Environmental Risk Acceptance Criteria

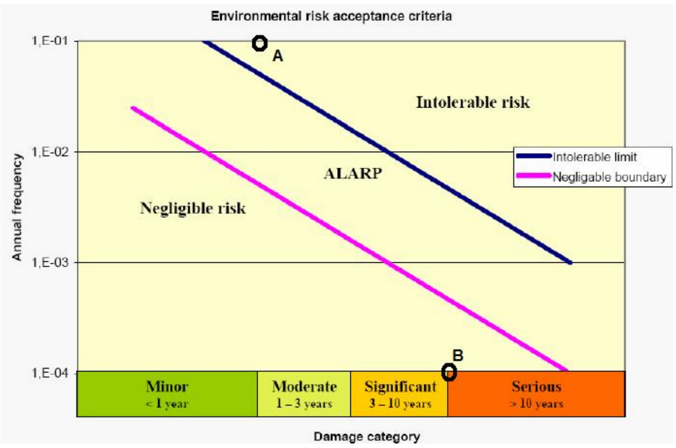

Figure 2: Environmental risk acceptance criteria (from document MEPC 55/18, Annex 3, p. 4, Fig. 3)

The ALARP concept mentioned earlier is the standard one already used in the FSA guidelines, using a slope of minus one in the F-N diagram. We are of the opinion that the use of a slope of minus one needs to be justified before adopted for environmental criteria. This can be seen by the following example, taken from Fig. 3 of document MEPC 55/18 (Annex 3, page 4), which depicts a possible ALARP region for environmental criteria. Let us add two points in this figure, A and B, as follows (see Fig. 2 above). Point A: Spill of minor-to-moderate severity (recovery time 1 year) that may occur once a year with a probability of 1/10. Point B: Spill of significant-to-serious severity (recovery time 10 years) that may occur once a year with a probability of 1/10,000.

A pertinent question is this: Is society really sure that point B, which refers to a rare but catastrophic spill scenario, should be ranked so much less below point A, which refers to a more frequent but much less serious spill scenario? Note that point B lies in the 'negligible risk' region, whereas point A lies in the 'intolerable risk' region, meaning that a series of measures, rules, or even legislation to prevent type-A spills might receive much higher priority over equivalent actions to prevent type-B spills. Of course, pertinent recommendations for decision making would be made later in the FSA analysis. However, a hazard ranking that assigns so much lower importance to a rare but environmentally catastrophic event as opposed to a more frequent but less serious one runs the risk of distorting the picture. It should be noted that if the slope of the F-N diagram is not equal to minus one and if the ALARP region is different, point B may not necessarily be ranked below point A (see Fig. 3 below).

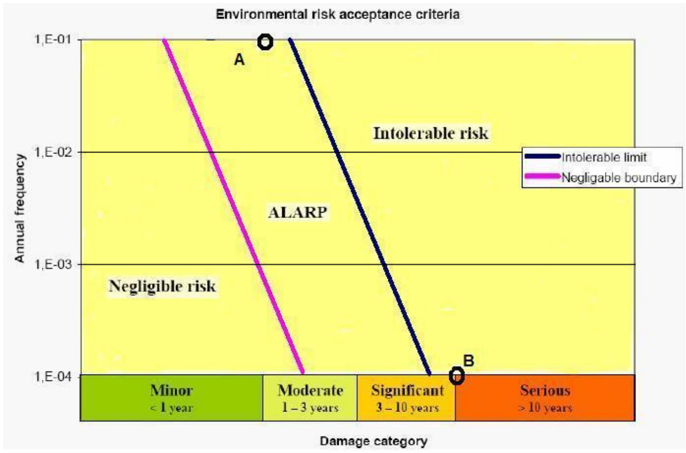

Figure 3: Alternate ALARP region if F-N slope is not equal to -1.

We therefore are of the opinion that the issues of Risk Matrices, F-N curves and ALARP regions for environmental criteria should receive thorough attention and debate before any of the numbers proposed in document MEPC 55/18 are adopted.

8.The Cost of Averting a Spill Criterion (CATS)

A major topic on this subject is the definition and analysis of risk evaluation criteria for accidental releases to the environment. In fact, mention of the "cost of averting a spill" is made (equation (1), page 6 of Annex 3 of document MEPC 55/18). In the SAFEDOR report, the criterion of CATS is defined as a 'per tonne of spilled oil' environmental criterion equivalent to CAF, the Cost to Avert a Fatality. The latter criterion is widely used in FSA studies in which risk to human life is assessed and Risk Control Options (RCOs) to reduce such risk are contemplated. According to the CATS criterion, a specific RCO for reducing environmental risk should be recommended if the value of CATS associated with it is below a specified threshold, otherwise that particular RCO should not be recommended. The equivalent threshold for the CAF analysis is $3 million (document MSC78/19/2).

Document MSC 81/18 cites the SAFEDOR report and claims that the latter concludes with a $19,000 per tonne value as the CATS threshold. In fact, in page 60 of the SAFEDOR report a $63,000 per tonne value is given for CATS, based on a series of assumptions. One can immediately note that if this figure is used in some actual past accidents, the resulting damages come out astronomical: The damage of the *"Prestige"* oil spill would be $5.1 billion and that of the *"Atlantic Empress"* $20.7 billion. If one actually translates these figures in terms of equivalent fatalities, and assuming the $3 million per fatality yardstick, the latter spill would be considered as catastrophic as 6,900 deaths!

The question what is an appropriate threshold value of CATS is an interesting one, but in our opinion sidesteps a more general question, whether the CATS criterion itself, that is, formulating an environmental index of costs averted *on a per tonne of spill basis*, is appropriate.

To arrive at a single threshold figure for such a criterion, in document MEPC 55/18 (page 5 of Annex 3) the following assumptions are made for simplicity purposes. Per tonne cleanup costs are thus assumed (a) constant with spill size, (b) independent of oil type, ie, a generic oil type is assumed, (c) constant within certain locations, and (d) independent of all other factors.

Although the need for simplicity is understood, it is very hard to justify these rather drastic assumptions, particularly given there is ample reference in the literature (see for instance Etkin (1999) and even Annex 3 of document MEPC 55/18 itself) that the cost of oil spills on a dollar per tonne basis depends on a variety of parameters and has a broad variance. This is in agreement with document MSC 81/6/3 by Japan, which includes, among others, statements such as *"as mentioned above the quantity of oil outflow is not a good measure of the impact of the spill, since it does not have a linear relationship with the risks to people and the environment. By concentrating on the quantity of the oil spilled the real risks are not being investigated."* (from ISSC 2000, Annex of document MSC 81/6/3, page 16).

In fact, according to ITOPF (White and Molloy, 2003), factors that determine the clean-up cost of spills include (a) type of oil, (b) amount of oil spilled and rate of spillage, (c) physical, biological and economic characteristics of spill location, (d) weather and sea conditions, (e) time of the year and (f) effectiveness of clean-up. And in general, costs involved in oil spill incidents include (i) clean-up costs, (ii) indemnification of the owner and (iii) compensation costs to third-parties.

In page 6 of Annex 3 of document MEPC 55/18, it is suggested that risk reduction measures are to be implemented if the costs of averting a spill are less than the costs of an occurred spill multiplied by F, where F is an "assurance parameter" postulated to be between 1 and 3 ($1<F<3$). That $F>1$ may be a plausible hypothesis (society should be willing to pay more to avert a spill than incur the cleanup cost of the spill itself). However, this hypothesis is not universally documented and factually one may witness situations where the opposite may be the case (society is complacent and unwilling to invest in averting spills and eventually ends up paying more to clean them up when they occur). A fortiori, postulating that $F<3$ seems rather arbitrary, and it may be the case that values of F higher than 3 are warranted, particularly for some pollution situations and measures to prevent it (for instance, tanker double hulls). Greece is of the opinion that if such an assurance parameter F (different from 1) is introduced, its appropriate value should be ascertained after a quantitative assessment of society's willingness to pay to avert pollution. The value of F should not be inferred 'in reverse', that is, to certify that previous legislative action to prevent pollution has been correct.

A point of primary importance is the inadequacy of using any single dollar per tonne figure as an environmental criterion. According to the SAFEDOR report and its references (see again Etkin (1999)), some average cleanup cost values in 1997 USD per tonne are: 1,600 (Africa), 12,700 (Europe) and 36,200 (USA). More recent data suggest the following average cleanup costs in 1999 USD per tonne: 6.09 [six USD] (Mozambique), 438.68 (Spain), 3,082.80 (UK), 25,614 (USA) and even the extreme value of 76,589 for the region of Malaysia (Etkin, 2000). Furthermore, the same report in page 54 states that *"ITOPF claims that as every oil spill is different with its own unique set of conditions, it is impossible to give, even within a limited geographic area, a reliable average cost per tonne spilt"*. All of the above testify to the broad variation of values on a per tonne basis, which makes the use of any single dollar per tonne figure questionable (see also Kontovas and Psaraftis (2006)).

We therefore are of the opinion that a non-linear function of spill volume is more realistic, one for which cleanup cost per volume spilled is a decreasing function of spill volume and is also a function of oil type and spill location. Environmental FSA would most likely be used to evaluate different regulatory options mostly on design issues. Having different non linear functions for quantity of different oils is therefore a reasonable way of proceeding as already MARPOL treats dirty oils, clean products and chemicals differently. Each of these could have a different CATS function reflecting the different behavior of each of these oils.

9. Conclusions

As regards formulating environmental risk evaluation criteria, this is an important task that should be pursued by research organizations and policy makers alike. However, the extreme variability of the per tonne cost (cleanup and damage) of oil spills worldwide cannot be overlooked, as a great number of factors other than volume are important. This means that one cannot produce a single number, to be applied worldwide, which, of all variables, uses oil spill volume alone as the main determinant of environmental pollution cost. If a regression analysis were made, it could very well establish that volume is not the most important variable in oil spill cost.

The need of IMO (and other regulatory bodies) to assess environmental risk and formulate relevant policy necessitates the development of a risk matrix to assess effects on the environment. The use of risk matrices is crucial in Formal Safety Assessment (and, by extension, to the Safety Level Approach). After gaining the needed experience, quantitative criteria to evaluate cost effectiveness could be discussed. In any case, any environmental risk evaluation criterion should have a strong theoretical background and should be based on assumptions that can be justified. The recent formation of an MEPC Corresponce Group to look into this matter in more detail can be seen as supportive of this basic premise.

As far as GBS itself is concerned, two Correspondence Groups have been formed after MSC 82 (December 2006). The first to continue the development of GBS for bulk carriers and oil tankers with the prescriptive approach (with a view to propose draft amendments for the incorporation of GBS in SOLAS chapter II-1) and the second to develop GBS using the Safety Level Approach. Both groups have submitted their reports for discussion at MSC 83 which will be held in October 2007 (documents MSC 83/5/2 and MSC 83/5/3 respectively). Time constraints prevent us to substantively comment on these documents, except to note that work in this area is far from over.

Whatever developments in this area take place, we believe that this paper has provided sufficient arguments that caution is necessary before the Safety Level Approach is fully integrated within the rule making process for maritime transport safety. Ongoing IMO work on the GBS methodology aspires to remove many of the current shortcomings of the scientific approach to maritime safety. In particular, the debate of how to bring the "safety level" (or "risk based") approach within the GBS framework is only just starting. While it is still early to draw conclusions, maybe the recommendations of this paper can be useful in such a process. From our part, caution is recommended, as we think it would be a mistake to rush through the GBS process before potential deficiencies in FSA and other Risk Based methodologies such as those identified in this paper are dealt with successfully.

Acknowledgments

This paper is the result of unfunded research collaboration among the three authors.

References

Barnett, A. (2006) "World Airline Safety: The Century So Far," Flight Safety Digest, in press.

Etkin, D.S.(1999), "Estimating Cleanup Costs for Oil Spills", Proceedings, International Oil Spill Conference, American Petroleum Institute, Washington, DC.

Etkin, D.S. (2000), "Worldwide Analysis of Marine Oil Spill Cleanup Cost Factors", Arctic and Marine Oil Spill Program Technical Seminar

HSE (1999) "Reducing Risks, Protecting People", Discussion document, Health & Safety Commission, 1999.

HSE (2001) "Reducing Risk, Protecting People :HSE's decision-making process", Health & Safety Commission, 2001.

Kontovas, C.A., (2005) "Formal Safety Assessment: Critical Review and Future Role", Diploma Thesis supervised by H.N. Psaraftis, National Technical University of Athens, July 2005 (available at www.martrans.org).

Kontovas C.A., H.N. Psaraftis (2006), "Assessing Environmental Risk: Is a single figure realistic as an estimate for the cost of averting one tonne of spilled oil?", National Technical University of Athens, Working paper NTUA-MT-06-01 (available at www.martrans.org).

Kontovas, C.A., H. N. Psaraftis, P. Zachariadis (2007), "Improvements in FSA as a Prerequisite for Risk-Based GBS", PRADS 2007 Conference, Houston, USA, October.

Skjong, R., E. Vanem, Ø. Endresen (2005), "Risk Evaluation Criteria" SAFEDOR-D-4.5.2-2005-10-21-DNV, October (available at www.safedor.org).

White, I.C., F. Molloy (2003), "Factors that Determine the Cost of Oil Spills", International Oil Spill Conference 2003, Vancouver, Canada.

Zachariadis, P., H.N. Psaraftis, C.A. Kontovas (2007), "Risk Based Rulemaking and Design: Proceed with Caution", RINA conference on Developments in Classification and International Regulations, London, UK, January.