Formal Safety Assessment: an updated review

Harilaos N. Psaraftis

Journal of Marine Science and Technology
Official Journal of the Japan Society of Naval Architects and Ocean Engineers (JASNAOE)

ISSN 0948-4280

J Mar Sci Technol
DOI 10.1007/s00773-012-0175-0

J Mar Sci Technol
DOI 10.1007/s00773-012-0175-0

REVIEW ARTICLE

Formal Safety Assessment: an updated review

Harilaos N. Psaraftis

Received: 2 January 2012 / Accepted: 9 April 2012

Abstract The method of Formal Safety Assessment (FSA) was devised several years ago with the intent to help the International Maritime Organization (IMO) and other policy makers formulate policies and regulations by the proper use of the scientific method in matters pertaining to maritime safety and the protection of the marine environment. A host of FSA studies have been submitted over the last several years and have been reviewed by the IMO. Also, progress has been recently accomplished as regards incorporating environmental risk evaluation criteria within FSA. On the basis of these developments, revisions of the FSA guidelines have been proposed and adopted. This paper presents a review of the FSA method in light of these developments. This review updates an earlier review of FSA by Kontovas and Psaraftis [Marine Technol 46(1):45–59, (2009)]. It also takes this opportunity to identify some deficiencies of FSA, either due to an incorrect application of the method or to the method itself, and makes some suggestions for further action in this area.

Keywords Formal Safety Assessment · International Maritime Organization · Safety · Environmental protection

Abbreviations

ALARP	As low as reasonably practicable
CAF	Cost of averting a fatality
CATS	Cost of averting a tonne of spilled oil
CBA	Cost benefit analysis
CEA	Cost effectiveness analysis
CO_2	Carbon dioxide
CTX	Center for Tankship Excellence
EEDI	Energy Efficiency Design Index
F–N	F–N curves
FSA	Formal safety assessment
GCAF	Gross cost of averting a fatality
HSC	High speed craft
HSE	Health and safety executive
IMO	International Maritime Organization
IOPCF	International Oil Pollution Compensation Fund
LMT	Laboratory for Maritime Transport
LNG	Liquefied natural gas
LOWI	Loss of watertight integrity
MEPC	Marine Environment Protection Committee
MSC	Maritime Safety Committee
NCAF	Net cost of averting a fatality
NPV	Net present value
NTUA	National Technical University of Athens
RI	Risk Index
RCO	Risk control option
RoPax	Ro/ro passenger ferry
UK	United Kingdom
USA	United States of America
USD	United States dollar
V	Spill size in tonnes
VHL	Value of human life (threshold on)
ΔB	Benefit of the RCO
ΔC	Cost of the RCO
ΔR	Risk reduction
ΔTSC	Total spill cost reduction
ΔV	Spill tonnes averted

H. N. Psaraftis (✉)
Laboratory for Maritime Transport, National Technical University of Athens, Athens, Greece
e-mail: hnpsar@mail.ntua.gr

Published online: 11 May 2012

1 Introduction

Formal Safety Assessment (FSA) was introduced by the International Maritime Organization (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 the IMO's options for reducing these risks" (see FSA Guidelines in MSC circ. 1023, MEPC circ. 392).[1] In the 81st session of the Maritime Safety Committee (MSC), an FSA 'drafting group' proposed some amendments to these guidelines (see Annex 1 to document MSC 81/WP.8).[2] These amendments have been approved by the MSC and were subsequently sent on to the Marine Environment Protection Committee (MEPC) for approval, something that happened at its 55th session (October 2006). As a result, there is now an amended set of 'consolidated' FSA guidelines, incorporating all recent revisions (this can be found in the Annex to document MSC 83/INF.2). The process of further amending the FSA guidelines is ongoing.

The topic of FSA has also been the object of research leading to several academic papers, even before its formal adoption by the IMO. For instance, we refer to the works of Wang [2], Soares and Texeira [3], Rosqvist and Tuominen [4], among others, for reviews, studies and analyses on the subject. RINA, the Royal Institution of Naval Architects, has also published a collection of some 15 papers on the subject, covering various contexts of the problem [5]. Devanney [6] has been much more critical, arguing that FSA is fundamentally flawed as a method and should be scrapped altogether. Last but not least, this author and his colleagues have been active in recent years in reviewing FSA, proposing ways to improve FSA, and extending it to environmental criteria. See for instance Kontovas et al. [7, 8], Zachariadis et al. [9], Kontovas and Psaraftis [1, 18]. See also Psaraftis [10], Kontovas et al. [11, 12] and Kontovas [13] as regards environmental criteria in FSA.

At the IMO level, recent developments as regards FSA have been numerous. Already MSC 83 had agreed to convene an FSA Expert Group with the purpose of reviewing FSA studies submitted to the IMO. MSC 85 invited Member Governments and international organizations to submit comments on several FSA studies submitted for review. This Expert Group met for the first time during MSC 86 (2009) and met again intersessionally between MSC 86 and MSC 87 (in 2009) and then again at MSC 87 (2010) and just before MSC 89 (2011).

Most of the recent FSA studies considered by the IMO were submitted by Denmark and consisted of the following documents:

- MSC 83/21/1 and MSC 83/INF.3 (FSA on LNG carriers).
- MSC 83/21/2 and MSC 83/INF.8 (FSA on container vessels).
- MEPC 58/17/2 and MEPC 58/17/INF.2 (FSA on crude oil tankers).
- MSC 85/17/1 and MSC 85/INF.2 (FSA on cruise ships).
- MSC 85/17/2 and MSC 85/INF.3 (FSA on RoPax ships).
- MSC 87/18/1 and MSC 87/INF.2 (FSA on transport of hazardous cargoes for open top containerships).

In addition, the International Association of Classification Societies (IACS) submitted an FSA on general cargo ship safety (MSC 88/19/2, MSC 88/INF.6 and MSC 88/INF.8).

The review of the FSA studies by the Expert Group covered the following points (see also doc. MSC 86/26, para. 17.23.1):

1. consider whether the methodology was applied in accordance with the FSA guidelines and the Guidance on the use of HEAP and FSA;
2. check the reasonableness of the assumptions and whether the scenarios adequately addressed the issues involved;
3. check the validity of the input data and its transparency (e.g., historical data, comprehensiveness, availability of data, etc.);
4. check whether risk control options and their interdependence were properly evaluated and supported by the assessment;
5. check whether uncertainty and sensitivity issues have been properly addressed in the FSA study;
6. check whether the scope of the assessment was met in the FSA study and propose any recommendations for re-analysis and re-calculation; and
7. check whether expertise of participants in the FSA study was sufficient for the range of subjects under consideration.

The Expert Group submitted a series of reports on the reviews of these FSA studies (see for instance docs MSC 87/18, MSC 87/WP.7 and MSC 89/WP.3). Based on these reviews, the Expert Group also went on to recommend some revisions in the FSA guidelines, a process that is ongoing. The FSA study on crude oil tankers has not yet

[1] Joint MSC and MEPC 'circular' on FSA, adopted on 5 April 2002. This document is now superseded by document MSC 83/INF.2.

[2] In this paper we cite IMO documents 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. IMO documents do not appear in the reference list of this paper, but a non-encyclopedic set of IMO documents relevant to this paper, including all recently submitted FSA studies, can be downloaded from this link: http://www.martrans.org/limo.htm.

been reviewed by the Expert Group for reasons that will be explained in Sect. 3.

In parallel to the MSC, the MEPC has been considering since MEPC 56 (2007) the issue of inclusion of environmental risk evaluation criteria within FSA. Progress on this subject has been rather slow, as divergence of opinion among stakeholders was more frequent than agreement. However, after 4 years of discussion by various Correspondence and Working Groups, MEPC 62 (2011) reached an agreement on the subject, and relevant amendments to the FSA guidelines were recently proposed.

As a result of the above noteworthy recent activities, this author feels that an updated review of FSA as a method is warranted, and this paper attempts to present such an update. In that sense, the paper can be considered as a sequel to Kontovas and Psaraftis [1] which had presented such a review, covering developments until circa 2008. The paper also takes this opportunity to identify some deficiencies of FSA, either due to an incorrect application of the method or to the method itself, and makes some suggestions for further action in this area.

The rest of this paper is organized as follows. Section 2 focuses on FSA deficiencies, either due to an incorrect application of the method or to the method itself. Section 3 discusses the environmental dimension of FSA. Section 4 discusses what may be ahead and draws the paper's conclusions.

2 FSA deficiencies

In this section, some of what we call FSA deficiencies can be attributed to an incorrect application of the FSA method. Some other deficiencies can be attributed to the FSA method itself. Both types of deficiencies are important and need to be avoided and/or fixed. In the following, we highlight some of the relevant points.

2.1 Transparency

The IMO has gone to great lengths to discuss, produce and amend (as appropriate) detailed guidelines for FSA. FSA guidelines are described in MSC/Circ.1023-MEPC/Circ.392, and have been consolidated in the Annex of MSC 83/INF.2. More amendments are due, none the least because of recent developments in the environmental part of the FSA (see Sect. 3). To quote from the aforementioned Annex (section 1.2),

> "These Guidelines are intended to outline the FSA methodology as a tool, which may be used in the IMO rule-making process. In order that FSA can be consistently applied by different parties, it is important that the process is clearly documented and formally recorded in a uniform and systematic manner. This will ensure that the FSA process is transparent and can be understood by all parties irrespective of their experience in the application of risk analysis and cost benefit assessment and related techniques."

In that sense, it would seem self-evident that any bonafide FSA study should conform to these guidelines. Yet, there have been a number of instances in the recently submitted FSA studies in which conformance with the FSA guidelines is lacking. Non-transparency is perhaps the most important example.

"Timely and open access to relevant supporting documents" is a central requirement of the IMO FSA guidelines (MSC 83/INF.2, Annex, section 9.2.1). This is a fundamental requirement of any scientific analysis. But in most of the cases, the casualty databases that were used in the recently submitted FSA studies are not only non public, but even worse, subject to strict non-disclosure agreements signed between vendors of such casualty data (typically LMIU or IHS Fairplay) and those who purchase such data for analysis. And even if someone else is to purchase the same casualty data, that someone can also not disclose the data underlying his or her analyses. If the conclusions of two studies using the same data are different, which is one to believe? For a study to be able to withstand scientific scrutiny, all data that was used must be available to whoever wants to replicate the analysis. So even though one would think that hidden, black box data cannot be allowed, this happens all too frequently.[3]

2.2 Root causes, initiating events and consequences

Another issue common in many of the recently submitted FSAs has been confusing cause and effect. Collisions, groundings, fires and explosions are consequences, not causes. Something else happened before any of these events took place. Yet, in many of these studies these events are euphemistically referred to as 'initiating events', with little or no analysis as to what prior event really caused them. But a collision or a grounding can be caused by other 'higher-level' (or 'root cause') events, such as a blackout, a steering gear failure, or other.

Doc. MSC 86/19/1 by Germany is cited here, which recommends that the *root causes* of accidents should be

[3] See http://www.c4tx.org/ctx/job/cdb/flex.html for an example of an open casualty database for tankers and bulk carriers (maintained by the Center for Tankship Excellence—CTX, USA). Another example of an open database is the "consolidated oil spill database" jointly developed by Germany, Greece, Japan and the US in 2011 (see Sect. 3). The latter database is not a casualty database, but will be in the public domain following a decision by MEPC 62.

investigated *before* RCOs are identified in an FSA. In our opinion, the main question is, what RCOs are put in place to prevent such higher-level events from happening? If this is not done, the focus is on RCOs to mitigate the consequence rather than prevent the cause from happening. Such a confusion may also skew the risk analysis that follows, and in fact, a typical pattern of the analyzed RCOs in the recently submitted FSAs deal with accident prevention and most deal with accident mitigation, namely what can be done after the accident occurs (when it is usually too late).

Perhaps the most characteristic example of this can be found in the crude oil tanker FSA (docs MEPC 58/17/2 and MEPC 58/INF.2). In that study, machinery failures are not included in the risk model, allegedly because they are not among the causes that can lead to what is termed 'Loss of Watertight Integrity' (LOWI). This is patently false, as a machinery failure can be the root cause event that eventually leads to a grounding or collision, and eventually to LOWI and an oil spill (the *Amoco Cadiz* being a prime example).

Unfortunately, many of the recently submitted FSA studies suffer from similar deficiencies, which in our opinion violate the fundamentals of risk analysis. As a result, the RCOs that are recommended are mostly in the accident mitigation rather than accident prevention domain. The structure of commercially available casualty databases is often blamed by FSA analysts for such a state of affairs, as root cause information is typically missing from these databases. Whereas this may be true, in our opinion it is not a good excuse for conducting an FSA study in which root causes are not included in the risk model.

This situation has been recognized by the FSA Expert Group when reviewing the recent FSA studies and some recommendations to amend the FSA guidelines to properly emphasize this point have been made. But of course this point will do nothing to rectify the FSA studies that have been conducted already.

2.3 Reliance on expert opinion

Most of the submitted FSA studies rely on expert opinion in critical steps of the procedure. Expert opinion is used to make estimates on the values of a broad spectrum of parameters, constants and variables that can be critical in the outcome of the analysis. These include probabilities of accidents, costs, benefits and a great number of other input data. Even though there are techniques designed to deal with groups and different opinions (such as Delphi and the like), and even though the FSA guidelines specify details on how concordance among experts can be measured, the usual practice is that these methods are rarely followed and hence relying too much on expert opinion is an inherent deficiency of the FSA method. Also sometimes the qualifications of 'experts' are found questionable.

Some examples of assumptions on critical input data many of which seem to rely on expert opinion can be found in the FSA study on LNG carriers (MSC 83/21/1 and MSC 83/INF.3). Pages below refer to doc. MSC 83/INF.3:

On page 40:

> "The second part of the damage extent model determines whether the grounding damage is critical or not in terms of damage stability. First, the damage needs to crack the outer hull, and the probability for this is estimated to be 0.76 for passenger ships [49]. For the purpose of this study, the same value will be used for LNG carriers."

An issue is, why LNG ships are assumed to be similar to passenger ships in that respect? How did the experts arrive at that conclusion?

On page 42:

> "For the purpose of this study on LNG carriers, it is assumed that the fire fighting systems have a similar success rate to that of HSC and passenger ships and the average will be used, i.e. 85 % chance of controlling the fire and 15 % chance of escalating fire. These values are inserted in the event tree."

An issue is, how can such an assumption be made, given that firefighting systems and the potential of fire escalation of drastically different ship types (HSC/passenger vs. LNG) are not necessarily similar?

On page 133:

> "For the purpose of this study, assuming that the effect of a 20 % hull strength increment is equivalent to increase the double hull width of the vessel by 20 %, the following has been considered: 1 % probability reduction of critical damage in collision and contact scenarios and 2 % probability reduction of critical damage in grounding scenarios."

Again, how did the experts find these numbers?

The situation is similar with most of the recently submitted FSA studies: they are typically full of many other similar assumptions, many based on expert opinion, but which are not justified, even though they could have a significant impact on the outcome of the study. If a number such as probability or other is not known, pulling it out of thin air is not the solution. There are other methods, such as using Bayesian analysis or other first principles that can be used. Assembling a group into a room and asking them to come up with numbers for many critical variables is questionable.[4]

[4] This author has been in a HAZID session (as an observer) in which half the people did not really know what was going on, and the discussion was dominated by a couple of experts who ventured their opinion and everybody else agreed without really knowing what was really agreed on.

2.4 Connection between Step 1 and the rest of the FSA

Another general issue with some of the submitted FSAs is that Step 2 does not follow Step 1 in the manner prescribed by the IMO FSA guidelines. These guidelines (MSC 83/INF.2, Annex, section 6.1.1) state that

> "The purpose of the risk analysis in Step 2 is a detailed investigation of the causes and consequences of the more important scenarios identified in Step 1. This can be achieved by the use of suitable techniques that model the risk. This allows attention to be focused upon high risk areas and to identify and evaluate the factors which influence the level of risk."

However, in many instances, the most important hazards identified in Step 1 (Hazard identification, or HAZID) were simply not examined in the rest of the FSA.

One example can be found in the recent FSA study for container vessels (MSC 83/21/2 and MSC 83/INF.8). Among the various hazards in Step 1, it is stated that three have Risk Index (RI) = 9 and four have RI = 8 (MSC 83/INF.8, Annex, page 10). But these are not identified nor mentioned anywhere else in the FSA. The study goes on to consider as top hazards those with RI = 7.4 or below. The hazard with RI = 7.4 is "bad working conditions during lashing (icy, wet floor)". However, this hazard is also eliminated from the rest of the FSA, on the grounds that such accidents have their causes in loading/unloading operations in container terminals and are sometimes associated with port personnel rather than the ship's crew. Perhaps another reason for this omission is that accidents associated with lashing do not appear in the casualty databases used in Step 2.

Eliminating these hazards from the rest of the FSA (along with any possible RCOs to mitigate them) not only makes this FSA non-conformant with FSA guidelines, but, substance-wise, it shifts regulatory focus to RCOs that cannot do anything for the hazards that the study itself recognizes as most important. A major cause for this seems to be the lack of coverage of these hazards in the used database. However, in such cases, other means of risk analysis than just analysis of the historical database, should be followed in the FSA, such as first principles or modelling.

2.5 Probability estimates

As said earlier, Step 2 of the FSA is the risk analysis step, the step which tries to evaluate risk, following Step 1. It would seem self-evident that this step would be based on sound casualty data and accepted scientific principles. However, let us see a recent FSA study to see to what extent these principles have been applied in that study.

Reference is made to the FSA study on cruise ships (MSC 85/17/1 and MSC 85/INF.2). A puzzling feature in the risk analysis of this FSA (Annex II of MSC 85/INF.2) is the use of fatality data of *ferries and RoPax vessels* to formulate worst-case scenarios for cruise vessels. In fact, Table 7-2 of Annex II of MSC 85/INF.2 (page 20) contains only accidents of ferries and RoPax vessels, including the *Herald of Free Enterprise*, the *Estonia*, and the *Al Salam Boccaccio 98*. But some of the very accident scenarios that have occurred on these ferries and RoPaxes, including water ingress via the bow door if the latter is left open (*Herald of Free Enterprise*) or is detached (*Estonia*) simply cannot occur on a cruise ship. Note that the *Al Salam Boccaccio 98*, whose capsizing took 10 min, was a RoPax in which two extra passenger decks were added in a conversion and was very different in design from a cruise ship.

Much of the probability and consequence data that populates the various event trees used extensively in the analysis seems arbitrary or difficult to justify. For instance, given a collision, the study assumes that the cruise ship is the striking ship with probability 50 % or the struck ship with probability 50 %. It is not clear whether the 50–50 chance is documented by accident statistics or seems like a convenient assumption. It is conceivable that cruise ships may be more prone to get struck by another ship than strike another ship, due to differences in navigation equipment, manoeuvering ability, etc. But if the probabilities are not 50–50, the results of the analysis may be different.

Also, if the cruise ship is the striking ship, the FSA study states that this will result to impact only with probability 85 %, to flooding with probability of 5 % and to fire with probability of 10 %. If it results to fire, it will result in minor damage with zero fatalities (with a probability of 45 %), to major damage with five fatalities (with a probability of 42 %), and to total loss with 20 fatalities (with a probability of 13 %). It is not clear how all of these numbers were estimated, although it is understood that some were based on expert opinion. Certainly the scenario of a cruise ship striking a tanker may result in a much higher number of fatalities due to fire.[5]

All of the above, which are certainly non-encyclopedic, show in our opinion, that FSA guidelines are not always followed. Other examples abound, for instance as regards the examination of RCO interdependencies. Psaraftis [14] provides more details.

[5] It is also not clear whether the RCOs recommended in the cruise ship FSA would have prevented accidents such as the *Costa Concordia*, in which human error seems to have been the prevalent factor.

If the above are deficiencies that are conceivably attributable to an incorrect application of FSA, the next deficiencies are attributable to the method itself.

2.6 Step 4 of FSA: GCAF and NCAF

The use of Gross Cost of Averting a Fatality (GCAF) and Net Cost of Averting a Fatality (NCAF) in FSA constitute, in our opinion, a structural deficiency of the FSA method itself. The IMO has gone at great lengths to suggest GCAF and NCAF as evaluation criteria in Step 4 of the FSA (cost-benefit assessment). Appendix 7 of the Annex to MSC 83/INF.2 outlines these criteria in detail. Even though it is stated very clearly (introduction of Appendix 7) that these criteria *may* be used in FSA studies, it is also stated that

> "the use of these cost effectiveness criteria would enable the FSA studies to be conducted in a more consistent manner, making results and the way they were achieved better comparable and understandable."

The last statement is, in our opinion, tantamount to requiring that the GCAF/NCAF criteria *should* be used in an FSA, and in fact with only one exception (see last part of this section), all FSA studies that we are aware of have used these criteria.

Both GCAF and NCAF are *ratio* criteria and are defined as the cost of the RCO (gross or net) divided by the expected reduction of fatalities because of the RCO (ΔR). GCAF refers to the *gross* cost of the RCO (GCAF = $\Delta C/\Delta R$) and NCAF refers to the *net* cost of the RCO, defined as gross cost minus any economic benefit caused by the RCO (NCAF = $(\Delta C - \Delta B)/\Delta R$). According to these criteria, RCOs are rank-ordered according to either GCAF or NCAF (or both) and the lower such criteria are, the more cost-effective the RCO is considered.

For both GCAF and NCAF the same threshold value of USD 3 million (the postulated average economic value of human life) is used as a cutoff point, implying that an RCO that has GCAF or NCAF above that threshold is not cost-effective. Leave aside for the moment what happens if these two criteria lead to different rankings among the RCOs (this is conspicuously left vague in IMO's FSA guidelines and is a basic methodological flaw in our opinion, as the two criteria can be used interchangeably, at the FSA analyst's discretion). Several of IMO's important regulatory decisions have been based on the use of these criteria and the above threshold value.

Indeed, it is interesting that even though the IMO has *not* officially adopted the USD 3 million threshold as the yardstick to be used in FSA and suggests this value is only illustrative, this threshold value has been used time and again in most FSA studies and is (at least de facto) the yardstick to be used in Step 4 of the FSA. Some suggestions to update this value have been voiced (even in MSC 83/INF.2), and such an update would indeed seem as self-evident, but this has not happened yet.

Table 1 Two hypothetical RCOs

	ΔR	ΔC (USD)	ΔB (USD)	GCAF (USD)	NCAF (USD)
RCO1	0.10	100,000	90,000	1,000,000	100,000
RCO2	0.05	45,000	42,500	900,000	50,000

Whatever its value is, let us call VHL the value of the above GCAF or NCAF threshold, assuming of course this value is known (we shall come back to this point later).

If this is so, note that mathematically, and as long as ΔR is not zero, the *ratio inequality*

$$\Delta C/\Delta R \leq \text{VHL (RCO has a GCAF below the threshold)} \quad (1)$$

is *exactly* equivalent to the *difference inequality*[6]

$$\text{VHL} \times \Delta R - \Delta C \geq 0 \quad \text{(RCO entails a positive absolute benefit).} \quad (2)$$

However, the *RCO rankings* implied by these two inequalities can be very different. An example of choosing between two hypothetical RCOs is shown in Table 1 above (see [1] for more examples).

Assume VHL = USD 3 million. In this example, both RCOs produce GCAF and NCAF well below the USD 3 million threshold. Among them, RCO2 is ranked above RCO1 with respect to both GCAF and NCAF. Yet, one can see that the expected reduction of fatalities of RCO1 is double that of RCO2. Also, one can compute the absolute economic benefit of each RCO as per Table 2.

Both RCOs yield positive benefits, and therefore can be considered as acceptable from a cost-benefit viewpoint. However, one can observe that, in addition to reducing fatalities more than RCO2, RCO1 is better than RCO2 in terms of absolute benefit.

It is clear that FSA would favor RCO2 if either GCAF or NCAF is used as criteria. That is, in this case FSA would favor an RCO that

a. would reduce expected fatalities by half of what another RCO would reduce;
b. would entail a gross absolute benefit of USD 95,000 less than what would be achieved by another RCO;
c. would entail a net absolute benefit of USD 142,500 less than what would be achieved by another RCO.

[6] The same is true as regards NCAF: $(\Delta C - \Delta B)/\Delta R \leq$ VHL is mathematically equivalent to VHL $\times \Delta R + \Delta B - \Delta C \geq 0$

Table 2 Absolute benefits of RCOs

	Gross benefit (USD)	Net benefit (USD)
RCO1	0.10 × 3,000,000 − 100,000 = 200,000	200,000 + 90,000 = 290,000
RCO2	0.05 × 3,000,000 − 45,000 = 105,000	105,000 + 42,500 = 147,500

The situation becomes even more complicated if GCAF and NCAF produce different rankings.

So both GCAF and NCAF are flawed as criteria to be used in FSAs, at least those FSAs dealing with fatalities. If they are proposed in the FSA guidelines to be used *only as examples*, they do not seem like very good examples. It turns out that a similar situation pertains to their use in environmental criteria, of which more in Sect. 3.4.

If these ratios exhibit such deficiencies, what should be used instead?

It is clear that using *differences* rather than ratios would have none of these deficiencies, as was done in Eq. 2 and Table 2. Differences, expressed on an NPV basis, can capture the *absolute scale* of the cost-benefit impact of an RCO, which is not possible by either GCAF or NCAF. And for a decision maker or investor such as a ship owner who would be asked to foot the bill for an RCO, the NPV of the RCO is much more sensible as a criterion than a ratio.

As stated earlier, the above analysis assumes that the value of the VHL is known. However, it may be that policy makers may be unwilling or reluctant to put a figure on the value of human life. If the value of VHL is not known, then the difference test of (2) would break down. So would the ratio test of (1), but still such test would produce some rankings.

One might consider that in such instances, and with all caveats expressed above, a dual use of both these tests can be contemplated. For environmental criteria (see also Sect. 3.4) this was explained in MEPC 62/18/2 by Japan. In fact, interchangeable or joint usage of difference and ratio criteria is not new. In economics parlance, the former typically belong to 'Cost Benefit Analysis' (CBA) and the latter (such as GCAF and NCAF) belong to 'Cost Effectiveness Analysis' (CEA). For a discussion of how CBA and CEA can be used for policy making, see [15].

Another alternative if the value of VHL is not known would be to run ranges on VHL and the decision maker can be presented with a sensitivity analysis.

As mentioned earlier, not all FSA studies have used the GCAF/NCAF criteria. In the context of the FSA studies submitted to the IMO, reference is made to the FSA study on dangerous goods transport with open-top containerships (MSC 87/18/1 and MSC 87/INF.2).

According to the study, recommended RCOs are shown to be suitable to *achieve an equivalent level of safety* for the open-top vessel, as compared to conventional transport with respect to individual classes of dangerous goods. According to this criterion, several RCOs are chosen, and among them is one that has a GCAF of close to USD 3 billion, that is, some 1,000 times higher than the USD 3 million threshold.

In this study GCAF and NCAF are *only secondary criteria*. In fact, the Expert Group that reviewed this study at MSC 87 (2010) expressed the opinion that the fact that criteria in this FSA are different than GCAF/NCAF should be no problem, so long as these criteria can be justified. This of course may raise the following question: If an FSA analyst can propose any criteria he or she deems appropriate, so long that proper justification is provided, why do we need any FSA guidelines to start with, and how will consistency and uniformity in the application of FSA be assured? If the answer is that the IMO has presented the GCAF and NCAF criteria only as examples and any other criteria can be used with proper justification, history has shown that the GCAF/NCAF criteria have been used in most FSA studies.

2.7 F–N curves and the ALARP region

Both F–N curves and the ALARP (for 'As Low As Reasonably Practicable') region are currently integral parts of the FSA methodology. Both are supposed to address what is called 'societal risk'. Neither the IMO or any other regulatory body have officially decided what are the appropriate risk tolerance levels of the ALARP region, in terms of actual frequencies or probabilities. Nor has the IMO officially decided that the slope of the F–N curve should be minus one, as is being widely practised. But some standard values for these parameters are routinely used by analysts, perhaps by inertia. For instance, the UK's Health and Safety Executive (HSE) maximum tolerable annual fatality risk level for passengers that is adopted by default for FSA is 0.0001.

There are also some deficiencies in the F–N/ALARP concept. Calculating risk *on an annual basis*, as is assumed by F–N/ALARP, is clearly questionable, as risk depends on how many times someone travels by ship. If John does not travel by ship, his risk to die in a ship accident is zero. If Mary takes a boat trip every 2 days, her risk is much higher than George's, who takes the boat twice a year. So how can it be that for both George and Mary, tolerable annual fatality risk is the same? Tolerable risk should be expressed *on a per trip basis* and in terms of an appropriate *exposure variable*, be that the number of ship trips per year, or something else, to be investigated (e.g, sea miles travelled, tons of oil moved, etc.). One can look at the airline industry and the studies conducted there, for instance [16]. For an airline passenger, the actual risk per flight is 1 in 8 million, meaning that if one takes 20 flights a year, his or her annual risk is 0.0000025, several orders of magnitude lower than the equivalent maximum tolerable risk for maritime transport.

3 Extending FSA to environmental criteria

3.1 Background

A big chapter in FSA that has opened in recent years concerns environmental criteria. At MEPC 55 (2006), the IMO decided to see how FSA could be extended to cover such criteria. A major topic in Annex 3 of document MEPC 55/18 was the definition and analysis of risk evaluation criteria for accidental releases to the environment, and specifically for releases of oil. Discussion on this matter was sparked to a significant extent by a report by EU project SAFEDOR [17], which defined the criterion of CATS (for 'Cost of Averting a Tonne of Spilled Oil') as an environmental criterion equivalent to CAF. If an RCO costs ΔC and averts ΔV tonnes of oil that would be spilled, CATS is defined as the ratio of $\Delta C/\Delta V$. According to the CATS criterion, a specific RCO for reducing environmental risk should be recommended for adoption if the value of CATS associated with it is below a specified threshold, otherwise that particular RCO should not be recommended. In the SAFEDOR report, a threshold value in the neighborhood of USD 60,000 per tonne of spilled oil was postulated for CATS, based on a series of modelling and other assumptions (see [18] for a discussion of these assumptions).

The issues of primary importance that triggered the debate at the IMO on environmental criteria were both the CATS criterion and its suggested threshold value of USD 60,000/tonne. By extension, the adequacy or inadequacy of using any single dollar per tonne constant figure as an environmental criterion was also a critical issue to be discussed. Various spill cost data over the years suggested the following average cleanup costs worldwide (USD/tonne, 1999 dollars): 6.09 (Mozambique), 438.68 (Spain), 3,082.80 (UK), 25,614 (USA) and even the extreme value of 76,589 for the region of Malaysia [19]. The *Exxon Valdez* 37,000-tonne oil spill had a cleanup cost of USD 107,000/tonne (2007 dollars), whereas the cleanup cost of the *Braer* 85,000-tonne oil spill was as low as USD 6/tonne. At least all of the above testify to the broad variation of values on a per tonne basis, which would make the use of any single dollar per tonne figure questionable.

The delegation that brought this set of considerations to the IMO was Greece, with document MEPC 56/18/1 which drew attention to these and other related issues. MEPC 57 (2007) noted that further work, including more research, was needed on the subject, and agreed to establish a Correspondence Group, under the coordination of Greece, in order to review the draft Environmental Risk Acceptance Criteria in FSA, and submit a written report to the 57th session of MEPC. The author of this paper was assigned the task to chair the Correspondence Group, something that lasted 4 years, until MEPC 60 (2010) (see docs MEPC 57/17, MEPC 58/17, MEPC 59/17 and MEPC 60/17).

3.2 Recent progress

Discussion in the Correspondence Group was rather difficult and divergence of opinion was very common on many issues. One of the first issues that was recognized was that spill size (typically represented by the volume of oil spilled) was not the only determinant of the severity of an oil spill accident. In addition to spill size, there are a number of other parameters that can have a significant impact on the severity of an oil spill, such as oil type, location, weather conditions, season, geomorphology of the shoreline, and others. Even so, and after considerable discussion, it was decided to use spill volume as the variable on which estimates of oil spill cost could be made, realizing that there could be considerable variance (scatter) in these estimates.

In that context, and endorsing the deliberations of a Working Group, MEPC 60 reached an important conclusion, that a volume-dependent non-linear spill cost function should be used, and recommended one such function proposed by Greece for further analysis (MEPC 60/17/2). Kontovas et al. [11] provide more details on this approach, which, after regression analysis of International Oil Pollution Compensation Fund (IOPCF) spill data, derives a function of the form:

Total spill cost (in 2009 USD) = $51{,}432\ V^{0.728}$

where V is spill size in tonnes. This particular function was chosen over those proposed by Norway [20] and Japan [21] because it was considered as more conservative (higher cost for same spill size).

It is not within the scope of this paper to discuss the degree of adequacy of using IOPCF cost data to represent the various components of the cost of an oil spill, noting only that IOPCF is perhaps the only global source in the public domain that systematically records such cost components, as well as compensation amounts paid to claimants. At the same time, various limitations of this data source have been recognized (see [11], among others).

In the spring of 2011, under the initiative of Germanischer Lloyd, researchers from Japan, the USA, Greece and Germany developed a 'consolidated oil spill database', incorporating updated IOPCF data, data from the US and data from Norway, and performed a new set of regression analyses with a view to submit the results to MEPC 62. This author and his colleagues at NTUA-LMT N. P. Ventikos and C. A. Kontovas participated in this process.

Table 3 Non-linear total spill cost functions, based on consolidated oil spill database (*V* is spill size in tonnes)

Spill dataset (IOPCF, USA, Norway)	Total spill cost (2009 US dollars)
All spills	67,275 $V^{0.5893}$
$V > 0.1$ tonnes	42,301 $V^{0.7233}$

The outcome of this effort led to several joint submissions to MEPC 62, by Germany, Japan and the US, and to an independent submission by Greece, leading very much to very similar results (see docs MEPC 62/18, MEPC 62/18/1, MEPC 62/18/2, MEPC 62/18/3, MEPC 62/18/4 and MEPC 62/INF.24).

Perhaps the most interesting (and quite unexpected) result of the new regressions was that even though data from the USA and Norway were added (and these are quite expensive spills), the new total spill cost functions obtained are *well below* the original one by Greece, based on IOPCF data alone. The new functions are given in Table 3.

Following the deliberations of a working group, MEPC 62 (2011) endorsed the consolidated database and the above functions, although it made clear that FSA analysts are free to use other formulae, so long as these are well documented by the data. MEPC 62 also decided to put the consolidated database in the public domain.

The absolute and per tonne total spill costs implied by the above functions are in Table 4.

MEPC 62 also agreed to package the main recommendations of the discussion on this topic in the form of an amendment to the FSA guidelines, and forwarded this to IMO's MSC for further action (see doc. MEPC 62/WP.13 for the report of the Working Group).

3.3 Open issues

Even though certainly MEPC 62 reached an important milestone in terms of converging on a difficult topic, at the same time there are several open issues. In our opinion these mainly stem from the considerable if not excessive latitude that the amended FSA guidelines provide to the FSA analyst in performing an environmental FSA. For instance, there is room for the following 'tools' to be used, all at the FSA analyst's discretion:

The first of these tools is not new and it concerns the so-called assurance factor. This factor is supposed to represent society's willingness to pay to prevent an oil spill instead of sustaining its damages. For instance, an assurance factor of 2.0 means that society would rather spend two dollars to prevent an oil spill than pay one dollar in the form of spill cost if the spill occurs. Thus far there has been no agreement on what this factor might be, even though there is a clear belief by some IMO delegations that this factor should be well above 1.0. On the other hand, some other delegations have suggested that this factor should not be the object of an FSA study but should be left for policy makers to decide. As things stand, the value of this factor is open, and FSA allows any value, so long as it can be well documented.

To this author, even the perhaps self-evident position that this factor should be above 1.0 is not supported by evidence. Thus far there has been no use of an equivalent assurance factor in the safety-related FSA, even though one could very well argue on the very same ground that in order to avert a fatality one might be willing to pay more than the economic value of the human life which would be lost in an accident. Yet, in none of the conducted FSAs there has been any consideration of this kind, and if there were, the outcomes could very well be different.

The second of the 'tools' in which the FSA analyst is given considerable latitude to use in an environmental FSA is the so-called uncertainty factor. This is a concept that has come up only very recently, and is supposed to represent the fact that if the recorded costs of a spill are (say) USD 100, the real costs of that spill may be higher (say USD 150), the 50 % difference being the uncertainty factor. In other words, what this factor does is reflect the fact that some spill costs cannot be captured and are uncertain.

It is not yet clear how any uncertainty factor can be computed with a reasonable degree of confidence. There is even an inherent contradiction in the term, in the sense that if any uncertain quantity can be estimated it is no longer uncertain. Also, if any uncertainty in the estimation of the real cost of an oil spill can be computed, this can be reflected in an update of the relevant cost information without the need of such a factor. In fact, as spill claims are typically inflated, the uncertainty factor can even be less than 1.0. But it was clear that some IMO delegations want to use an uncertainty factor much higher than 1.0. As things stand, any uncertainty factor can be

Table 4 Spill costs implied by functions of Table 3 (2009 US dollars)

V (tonnes)	Total spill cost		Per tonne total spill cost	
	All spills	$V > 0.1$	All spills	$V > 0.1$
0.01	4,459	N/A	445,917	N/A
0.1	17,320	7,999	173,202	79,993
1	67,275	42,301	67,275	42,301
10	261,309	223,692	26,131	22,369
100	1,014,971	1,182,907	10,150	11,829
1,000	3,942,337	6,255,336	3,942	6,255
10,000	15,312,768	33,078,868	1,531	3,308
100,000	59,477,637	174,924,497	595	1,749

used in an environmental FSA, so long as it is well documented.

The third potential 'tool' that came up at MEPC 62 (and in that sense, it too came up very recently) is the possibility to perform a regression analysis not at 50 % (as is standard in all regressions), but at a level different from (and likely higher than) 50 %. This is another new idea, unseen before anywhere in FSA studies in the past, and which, if it were to be used, would conceivably change their outcomes. This means that the FSA analyst would be able to choose a higher regression line (for instance at the 80 or 90 % level) if properly justified. In the oil spill context, a higher regression line would also raise spill costs, as only high cost spills would be left in.

Last but not least, and as already alluded to, the spill database can be amended and the FSA analyst is free to use alternate cost functions, to the extent these can be documented. In that sense, if the FSA analyst comes up with a database that leads to a constant CATS, he or she very well can do it.

Collectively, the latitude allowed to an FSA analyst by all of the above 'tools' may make the results of the recent MEPC deliberations something between not very useful to totally useless. One is left wondering what would happen if *for the same problem* different teams of FSA analysts choose different factors, different functions, different regression lines or databases, and use different arguments to support these choices. The same latitude would also likely increase the chance of FSA studies being manipulated to lead to any a priori desired outcome. If the objective was for IMO to recommend spill cost functions that *should* be used in an FSA, the IMO recommended spill functions that *could* be used, *but only so long as the FSA analyst wants to use them*. If someone does not like these functions, for whatever reason, including less than desirable implications, they can choose several paths to bypass them, by 'tools' such as the assurance factor, the uncertainty factor, regressions at a higher than 50 % level, or even 'custom made' spill databases. Of course, the analysts will have to provide adequate justification and documentation for all this, but this is another story.

The IMO, after about 4 years of deliberations, decided against a constant CATS, much less one whose threshold is USD 60,000/tonne or above, but instead endorsed databases and spill cost functions that yield much lower per tonne spill costs (as per Table 3 above). However, the advocacy of 'tools' such as those listed above, some of which were proposed only after the results of the relevant cost regressions became apparent, gives the clear impression of being pursued so as to neutralize or even reverse the potential implications of these results. It is perhaps no coincidence that the main advocates of all such 'tools' are delegations that have long advocated high CATS thresholds (USD 60,000/tonne and above) and a constant per volume spill cost.[7]

What may be ultimately at stake here is not entirely clear. But one can make some conjectures. Suffice it to say that the Danish FSA on crude oil tankers (see docs MEPC 58/17/2 and MEPC 58/17/INF.2) has recommended mandatory adoption some RCOs that would drastically change the design of tankers in the future. These include increased side tank widths and double bottom heights. It is critical to point out that these RCOs were found cost-effective based on a CATS threshold of USD 60,000/tonne and a constant per volume oil spill cost. As discussed in Hamman et al. [22] and Yamada and Kaneko [23], these RCOs would not be cost-effective if the non-linear oil spill total cost functions proposed by Japan [21], Norway [20] and Greece [11] were applied. A fortiori, the same outcome would be the case if the functions of Table 3 are used.

However, if suitably chosen assurance factors, uncertainty factors, regressions higher than 50 %, and different, 'tailor made' spill databases are used, this outcome could very well be reversed. The Danish crude oil tanker FSA has still not been reviewed by the FSA Expert Group, due to the fact that up until recently the CATS issue had not been resolved. But it would be interesting to see the outcome of that review when the study is taken up by the FSA Expert Group. This is anticipated to take place after MSC hands over this FSA study to the Expert Group. At the time this paper was being finalized, the outcome of this process was not known.

3.4 Using NCAF in environmental criteria

Upon recommendation of the relevant Working Group, MEPC 62 also endorsed a *ratio test* in case a certain RCO reduces both fatality and environmental risk. This ratio test is based on NCAF and was chosen by the Working Group after an alternative formulation, proposed by one of the delegations and based on a *difference* test, was rejected. The main argument for the rejection of the difference test was that since NCAF provides a way to proceed, no alternative is needed.

Of course, FSA guidelines (see Appendix 7 of the Annex to MSC 83/INF.2) only *suggest* the use of NCAF (together with GCAF), even though as mentioned earlier, this is tantamount to *requiring* their use in an FSA. But if

[7] As there is an 'assurance factor' of 1.5 embedded in the CATS threshold of USD 60,000/tonne, its implied per tonne total spill cost is USD 40,000/tonne, constant for all spill sizes. Psarros et al. [20] argued for an even higher CATS threshold (more than USD 80,000/tonne) and the USA has argued for a value in the range of USD 100,000/tonne. These are way above the per tonne figures of Table 4, except perhaps for very small spills.

Table 5 Choosing among 3 hypothetical RCOs that reduce both fatality and environmental risk

	ΔR	ΔC (USD)	ΔTSC (USD)	NCAF (USD)
RCO1	0.1	500,000	100,000	4,000,000
RCO2	0.2	400,000	200,000	1,000,000
RCO3	0.4	900,000	300,000	1,500,000

NCAF is to be used for environmental criteria, it can be shown that it does *not* provide a full-proof way to proceed. In fact, the NCAF of an RCO that reduces both fatality and environmental risk is as follows:

$\text{NCAF} = (\Delta C - \Delta\text{TSC})/\Delta R$

where

ΔC = Expected cost of the RCO

ΔTSC = (Expected total spill cost *without* the RCO) − (Expected total spill cost *with* the RCO) = Expected benefit of the RCO, and

ΔR = Expected reduction of fatalities due to the RCO (assumed nonzero).

In the example shown in Table 5 it can be seen that NCAF suffers from the same deficiencies as those shown in Sect. 2.6 for GCAF and NCAF in safety-related FSAs.

If one were to choose among the three RCOs shown above, which one should be preferred?

According to the NCAF criterion, RCO1 has an NCAF above the USD 3 million threshold and should be rejected if this threshold applies. On the other hand, both RCO2 and RCO3 have an NCAF under the threshold, so both are a priori acceptable. Among them, RCO2 has a lower NCAF and, as such, should be preferred to RCO3.

Notice however that, in choosing RCO2 over RCO3:

a. we get a lower reduction in expected fatalities (0.2 versus 0.4),
b. we get a lower environmental benefit (USD 200,000 versus USD 300,000), and
c. we get a lower absolute total net benefit as well!

Indeed, the absolute total net benefit of RCO2 is $0.2 \times 3{,}000{,}000 + 200{,}000 - 400{,}000$ = USD 400,000, whereas the one for RCO3 is $0.4 \times 3{,}000{,}000 + 300{,}000 - 900{,}000$ = USD 600,000.

This, in this author's opinion, confirms that the NCAF criterion is a poor criterion, not only for the safety-related FSA (as per Sect. 2.6, along with GCAF) but also when it comes to choosing an RCO that reduces both fatality and environmental risk. A difference criterion does not suffer from such deficiencies.

Naturally, the same considerations as those expressed in Sect. 2.6 on the possible joint use of difference and ratio criteria also pertain here, particularly in case policy makers are not willing to put a value of VHL. Caution should be exercised however, so as to avoid situations such as the one described above.

4 Way ahead and conclusions

As said earlier, FSA is currently under review by the IMO. Given the outcome of MEPC 62 on environmental criteria, MEPC's role in FSA is over, at least for the time being [8]. The buck is passed over to the MSC, which the only responsible committee in the IMO to handle FSA in the foreseeable future. To that effect, at least the following issues are scheduled to be discussed in the context of revision of FSA guidelines:

1. description/discussion of experts participation in FSAs;
2. description of the structure, selection and composition of the project team, HAZID team and any other team, if established for taking any decision making;
3. information and analysis on root causes and details of casualties, with a view to obtaining RCOs focused on prevention rather than mitigation;
4. development of risk models;
5. unification of terminologies;
6. reporting the method and justification for the final selection of RCOs;
7. indices for cost-benefit analysis for risks other than safety of life;
8. clarification on the use of NCAF and GCAF;
9. methodologies to analyze possible side effect of RCOs;
10. methodologies for sensitivity and uncertainty analysis;
11. consideration of the human element (to have more detailed and specific guidance);
12. methodologies to reach the consensus or agreement as well as reporting the degree of agreement, or concordance;
13. how to present reports; and
14. how to review FSA studies.

In addition, amendments to FSA guidelines are forthcoming as a result of MEPC 62 (environmental criteria).

Moreover, some recommendations have been made by the FSA Expert Group regarding casualty data; that is, to the effect that:

[8] Note that MEPC has only examined *oil pollution* in the context of FSA. At the same time, there are a number of other important environmental issues that merit inclusion, such as hazardous substances, residues, ballast water, and emissions, to name just a few. Inclusion of these within FSA may reopen MEPC's role into FSA at some point in the future.

- databases shall not confuse cause and effect, shall contain information on the root causes of accidents, and allow for multiple causation factors;
- databases shall be able to describe the casualty in complete detail whenever possible, but be flexible enough to also handle casualties for which limited information is available;
- all sources used shall be clearly stated, and there should be a distinction between the 'factual' fields (e.g, date) and the 'judgmental' fields (e.g., causal factor);
- search engines for the IMO GISIS casualty database should be further developed to allow searching by types of ship, casualty date and place, ship particulars, initiating event, and in general the main information contained in the casualty report forms;
- cooperation with the European Maritime Safety Agency (EMSA) regarding casualty databases could be considered; and
- it is necessary to devise a new and efficient method to utilize the huge source of data available inside the seafaring community which up to now remains mainly untapped (e.g., near-miss data), data on psychological factors, as well as adequateness of man–machine interface of hardware systems based on practical on-board experiences should be taken into account.

The database issue is very serious in our opinion and should be looked at with a high sense of urgency. Open databases such as the one maintained by CTX should be looked upon as possible models. It is clear that unless the casualty data of an FSA study is sound and adequate, the FSA itself will suffer.

At the same time, we believe that even though many of the above IMO actions are steps in the right direction, the updated review of FSA presented in this paper supports the thesis that it is high time to seriously retool FSA as a method. A method whose main evaluation criteria and other methodological aspects exhibit deficiencies (serious in our opinion) and whose official guidelines purport to serve consistency and uniformity but in practice are either not followed or are encouraged to be interpreted and used at will, can not be of much assistance in the formulation of maritime regulation.

If any approach can be used to retool FSA, we believe that it should be based on the following principles:

First and foremost, all data used should be available for scrutiny and inspection, and all algorithms, models and tools should be open. Any analyst should be able to replicate the analysis, and no black boxes should be allowed. This is a fundamental ingredient of scientific inquiry and should be respected fully and unequivocally. To FSA's credit, such a principle is part of the current FSA guidelines (MSC 83/INF.2, Annex, Sect. 9.2.1). But thus far it has seldom been followed, or respected in full.

Second, expert opinion should not be stretched to concoct elusive data. If certain data does not exist, it should not be pulled out of thin air. There are scientific methods that can be used in the absence of data, such as Bayesian analysis and others based on first principles. Even if no data whatsoever exists, ranges can be run for this data and sensitivity analyses can be performed. One could even go as far as to say, using data *based on expert opinion alone* should not be allowed.

Third, and as mentioned earlier, the GCAF/NCAF framework for the cost-benefit assessment should be radically restructured. In particular, the use of any ratios that ignore scale should be avoided, and the analysis should be based on cost-benefit *differences* on a Net Present Value (NPV) basis.

Fourth, assurance factors should be left to policy makers to decide and not be included in the analysis. To help policy makers decide what an appropriate assurance factor might be, a comprehensive programme of R&D to ascertain society's risk aversion to maritime accidents should be launched, and that should not be limited to environmental issues.

Fifth, uncertainty factors should not be included in the analysis. Any updated information on spill costs and other data should be reflected in the corresponding cost databases. A continuing effort to develop and enhance such databases should be undertaken.

Sixth, as regards F-N curves and the ALARP region, these concepts should be retooled so that risk is expressed in terms of an appropriate 'exposure variable', such as number of ship trips per year, sea miles travelled, or tons of oil moved.

Last but not least, a full cross-disciplinary approach should be adopted. This should link safety and environmental aspects, should not be limited to oil pollution and should also include emissions, in a seamless integrated fashion. Among other things, this would imply that the impact of Energy Efficiency Design Index (the ship EEDI) on safety should be examined, and conversely the impact of safety measures on CO_2 emissions and EEDI should also be examined. This aspect seems to be recognized in the current attempt to revise the FSA guidelines, even though no analysis to connect these two aspects has been attempted yet. Whenever it is, we think it should be coupled with all other suggested changes, otherwise it would be ineffective.

Acknowledgments All opinions in the paper are the author's alone. The author is a member of the IMO FSA Expert Group, tasked to review the FSA studies submitted to the IMO (2009–2011). He has also coordinated on behalf of Greece as many as four IMO Correspondence Groups on FSA-environmental risk evaluation criteria from MEPC 56 to MEPC 60 (2007–2010). Last but not least, he chaired two IMO Working Groups on the same subject at MEPC 60 (2010) and MEPC 62 (2011). The author would like to thank the Editor and three anonymous reviewers for their comments on a previous version of the paper. He is also indebted to Jack Devanney, Christos Kontovas, Nikos Ventikos, Erik Styhr Petersen and Panos Zachariadis for assembling some of the material and for their comments on previous versions of this paper.

References

1. Kontovas CA, Psaraftis HN (2009) Formal Safety Assessment: a critical review. Marine Technol 46(1):45–59
2. Wang J (2001) The current status and future aspects in formal ship safety assessment. Saf Sci 38(1):19–30
3. Soares Guedes C, Teixeira AP (2001) Risk assessment in maritime transportation. Reliab Eng Syst Safety 74(3):299–309
4. Rosqvist T, Tuominen R (2004) Qualification of Formal Safety Assessment: an exploratory study. Saf Sci 42(2):99–120
5. RINA (2002) Formal Safety Assessment. Conference proceedings, Royal Institution of Naval Architects, London, UK
6. Devanney J (2008) Formal Safety Assessment in Wonderland. Center for Tankship Excellence (CTX) working paper (available at http://www.c4tx.org)
7. Kontovas CA, Psaraftis HN, Zachariadis P (2007) The two C's of the risk based approach to goal based standards: challenges and caveats. International Symposium on Maritime Safety, Security and Environmental Protection (SSE07), Athens, Greece, September
8. Kontovas CA, Psaraftis HN, Zachariadis P (2007) Improving FSA as a prerequisite for risk-based GBS. PRADS 2007—The 10th International Symposium on practical design of ships and other floating structures, Houston, October
9. Zachariadis P, Psaraftis HN, Kontovas CA (2007) Risk based rulemaking and design: proceed with caution. RINA Conference on developments in classification and international regulations, London, January
10. Psaraftis HN (2008) Environmental risk evaluation criteria. WMU J Maritime Aff 7(2):411–430
11. Kontovas CA, Psaraftis HN, Ventikos N (2010) An empirical analysis of IOPCF oil spill cost data. Mar Pollut Bull 60(9): 1455–1466
12. Kontovas CA, Psaraftis HN, Ventikos N (2011) Estimating the consequence costs of oil spills from tankers. 2011 SNAME Annual Meeting & Expo, Nov. 18–19, Houston
13. Kontovas CA (2011) Quantitative risk management framework for maritime safety and environmental protection. Unpublished PhD thesis, School of Naval Architecture and Marine Engineering, National Technical University of Athens, December
14. Psaraftis HN (2009) Comments on FSA studies by Denmark. Unabridged version of MEPC 86/17/2, submitted to IMO FSA Group of Experts. Available online at http://www.martrans.org/documents/2009/sft/86-17-2%20UNABRIDGED.pdf
15. Boardman AE, Greenberg DH, Vining AR, Weimer DL (2001). Cost-benefit analysis. Concepts and practice. Second edition. Prentice Hall, Upper Saddle River
16. Barnett A (2006) Global passenger-mortality risk decreased substantially in accidents from 2000 to 2005. Flight Safety Digest 25(1–2):14–19
17. Skjong R, Vanem E, Endresen Ø (2005) Risk evaluation criteria. SAFEDOR-D-4.5.2-2005-10-21-DNV; 21 October 2005 (report available at http://www.safedor.org)
18. Kontovas CA, Psaraftis HN (2008) Marine environment risk assessment: a survey on the disutility cost of oil spills. SNAME Greek Section's 2nd International Symposium on Ship Operations, Management and Economics, Athens, Greece, September
19. Etkin DS (2000) Worldwide analysis of marine oil spill cleanup cost factors. Arctic and Marine Oil Spill Program Technical Seminar
20. Psarros G, Skjong R, Vanem E (2011) Risk acceptance criterion for tanker oil spill risk reduction measures. Mar Pollut Bull 62(1):116–127
21. Yamada Y (2009) The cost of oil spills from tankers in relation to weight of spilled oil. Marine Technol 46(4):219–228
22. Hamann R, Loer K, Buettner M (2010) Risk based optimization of oil tanker cargo holds. SNAME Greek Section's 3rd International Symposium on Ship Operations, Management and Economics, Athens, Greece, October
23. Yamada Y, Kaneko F (2010) On the Derivation of $CATS_{thr}$ within the Framework of IMO environmental FSA studies. 5th International Conference on Collision and Grounding of Ships, June 14th–16th, Espoo, Finland