Evaluating Risk Control Options
Very Preliminary Draft

Jack Devanney
Center for Tankship Excellence, USA, djw1@c4tx.org

IMO has concocted and is still in the process of concocting a methodology for evaluating proposed regulations to attempt to determine whether a particular measure — known as a Risk Control Option (RCO) in IMO jargon — would be beneficial to society as a whole. The IMO methodology (Formal Safety Assessment (FSA), Risk Evaluation Criteria, Safety Level Approach, etc) has gotten so far off track the CTX does not intend to spend much time discussing it. Rather we put forward our own proposed system. Anyone who wishes can contrast it with the IMO methodology. The footnotes are a start in that direction.

We must always keep in mind that in devising regulatory systems, the goal is not optimality. ***The goal is detecting clearly beneficial regulation and most importantly avoiding stupid regulation. This means the system must be simple, transparent, understandable, objective, not impose its own values, and difficult to manipulate by special interest groups.***

Here is a brief outline of the methodology that CTX would use.

1. **Assume society is an expected value decision maker**.
   Given society's combined wealth and the small overall impact of all the RCO's that IMO is likely to consider on world wealth, world population, world whatever, this is not unreasonable. This assumption allows us to collapse all our random variables (uncertainties) down to the mean of the variable's distribution, ignoring the spread.[1] This is not very intelligent behavior for an individual faced with a really important decision, nor would it be for society as a whole if the downside consequences were dire enough (e.g nuclear war). But for the issues IMO addresses it is probably OK. In any event, without this assumption, quantitative analysis becomes totally intractable requiring bogus assumption after bogus assumption to do.
2. **Assume society wishes to minimize total killed, total oil spilled, total whatever.**
   In other words, treat every death, every ton of oil spilled the same. This requires that we go linear in people killed, spill volume, etc. This is an immense simplification methodologically. As far as deaths go, this assumption strikes the CTX as philosophically unassailable. Society in passing regulations should not favor one individual over another. The fact that the media makes a great deal more fuss over an airline crash that kills 200 Americans than it does over 5000 industrial accidents in China each of which kill one person is irrelevent. To adopt an evaluation system which accepts more deaths overall in order to protect against a smaller number of grouped deaths should be anathema.[2]
   As far as oil spillage is concerned, the argument for focusing on total volume is a bit weaker; but still the best, feasible way to go from the point of society as a whole. Yes, the cost of clean-up and somebody's non-disinterested guess at the "cost of environmental damage" is sharply

[1] The IMO system makes this assumption albeit in a not very explicit fashion.

[2] Most readers are probably wondering why we are belaboring such an obvious point. In the literature, a school of thought has arisen that argues that regulation should be biased toward "catastrophes", that somehow it is better to prevent a "catastrophe" which kills 100 people than prevent 101 "accidents" each of which kill one person. The IMO system has incorporated this strange thinking in its methodology.

location and volume dependent. But once again this is irrelevent. Pretend you are Mother Nature. Do you really want these stupid humans to come up with a system that says it is OK to put more oil into the water rather than less as long as you do it in gargantuan splotches?[3] Environmental regulation should be concerned with the environment, not somebody's beach. In any event, going after anything other than total volume not only puts more oil in the water, but also forces the analysis to require a whole range of far worse assumptions to go forward; and all such assumptions (always bogus, often hidden) become the focus of charlatans and special interest groups, usually called experts.

3. **Full transparency.**
   All the raw data on which the analysis is based must be public. All the assumptions must be clearly highlighted. Anyone should be able to take the same data and the same assumptions and come up with the same output. Public regulation based on private data is an oxymoron. Once again this would seem obvious; but not to the IMO. The great bulk of Formal Safety Assessments to date, as well as most of the regulation surrounding double bottoms and double sides is based on data collected by the Classification Societies or on data bases such as that developed by the Lloyds Maritime Information Unit. In the former case, the confidentially of the data is a legal requirement of the contract between the Classification Society and the ship owner. In the latter case, the data cannot be made public per the terms and conditions under which LMIU and others sell access to their proprietary data. In many cases, even the IMO delegates can't see the data upon which the analyses are based.
   Confidentiality can generate strange distortions. Analysis of many RCO's requires damage location data. IMO uses data collected by the Classification Societies for this purpose. In order to "preserve confidentiality", the Classification Societies non-dimensionalized their penetration data. That is, instead of reporting the depth of penetration in meters they reported the depth as a percent of the struck ship's beam (for transverse penetration) or depth (vertical) or length (longitudinal). Regulators had no choice but to base regulations on these non-dimensionalized numbers. So now the standard IMO casualty scenario assumes that in the same collision a wider ship suffers more penetration than a narrow ship. In reality of course the depth of penetration depends on the size of the striking ship, not the size of the struck ship. The whole thing is horribly biased against bigger ships.
   Without third party scrutiny, errors in the data cannot be caught and corrected. The LMIU data has lots of errors. This is particularly important in ship casualties. Much of the so-called data actually involves a judgement on the part of the data coder. These judgements are subject to all kinds of biases.[4] And we don't know how the data is censored. Does the Class damage extent data include ships that were totally lost? We don't know. Probably not, since it is based on surveyors' repair reports. Yet almost all tanker deaths and almost all tanker spillage are in casualties in which the ship was lost and never repaired.
   To be fair, the Guidelines for Formal Safety Assessment require transparency. In practice this requirement is ignored. One would think that this would invalidate such an FSA; but apparently it does not.

[3] Almost all the oil that is spilled by tankers is spilled in a handful of extremely large spills. This is a product of the size range of spills. Tanker spills can be as small as 1 liter and as large as 300 million liters. In terms of numbers, almost all the spills are at the very lower end of this range. But the size range is so large, that one spill at the top end of the range spills far more than tens of thousands of spills at the lower end of the range. A society which wishes to minimize total spill volume should focus on decreasing the probability of a really large spill. A society that attempts to minimize a decreasing function of spill volume will probably end up doing nothing about those few spills which spill almost all the oil. For example, the US Transportation Research Board after all sorts of poorly supported calculations concluded that the cost of an oil spill goes as the 0.4 power of volume. This means that each ton of oil spilled by the Exxon Valdez cost society 1/600 as much as if that ton had been spilled by itself. If we really believe this, we would be crazy to worry about big spills. In fact, we should adopt measures that reduce the number of small spills even if they increase the probability of an extremely large spill in a manner that on average will put far, far more oil on the water. With double bottoms that's probably what we've done.

[4] Still worse, even the design of these data bases is fundamentally unsound. Most of them treat grounding, collisions, and fire as if they were causes. These are not causes but consequences of whatever caused the casualty. The real cause, which should be the focus of regulation, is lost.

4. **Don't turn opinions into numbers**.
   Use no expert polls, no probabilities by committee. Generate your densities from whatever data you have via Bayes Rule starting with zero information priors. Any "analytical" system whose outcome depends on which group of experts you pick is not analysis; it's politicking.[5] Sometimes you may have no useful data at all on an important variable; for example, Mean Time to Repair (MTTR) in a cost benefit analysis of twin screw. In some cases, you will be able to assume a clearly advantageous (to the RCO) value and the cost benefit will still come out negative. In other cases, the reverse will be true. In some cases, you will be forced to keep the number a variable, and then at the end of the analysis display the range of (say) MTTR's for which the RCO has a positive net benefit. Obviously, you must keep the number of such variables to a minimum for the analysis to stay tractable. The important point is whatever you do with such variables it be both transparent and highlighted. Burying the uncertainty about such variables beneath "expert guesses" is neither.
5. **Do the analysis**.
   Work out the expected cost of the RCO, work out the mean total lives saved, work out the expected impact on total spill volume. This will almost always be a very big job, but the fact that we only have to work with means and we are assuming linearity will usually make the job doable. For one thing, we are free to interchange the order in which we apply probabilities to the various numbers.
6. **Keep your axes separate**.
   Put off valuing deaths versus spills versus dollars until the very end. Going expected value and focusing only on totals (linearity) allows us to do this. This is impossible under any other set of assumptions.
7. **Dump the dominated RCO's**.
   Eliminate all the inefficient RCO's, those for which other mutually exclusive alternatives have both better cost and better deaths/less spilled. This will get rid of a lot of the really stupid ideas.
8. **Set up the net societal benefit equation**.
   Finally set up your expected discounted benefit less expected discounted cost equation with the $/death, and $/m3 spilled coefficients **regarded as variables**. Eliminate RCO's which require patently unrealistic $/death and/or $/vol to generate a positive net benefit. More bad ideas gone. But be sure and properly document all the eliminations.[6]
9. **Present the results in the form of acceptance regions**.
   For the remaining RCO's figure out the range of $/death and $/m3 for which the RCO has a positive net benefit.[7] If an RCO is truly valuable, it will generate a positive net benefit for

[5] Turning unsupported opinions into hard numbers is at the heart of the IMO methodology.

[6] At the end of the day, the IMO system goes linear in both deaths and volume, although this is inconsistent with much of the preceding bureaucratese. The slope of these two lines is called CAF (Cost [sic] of Averting a Fatality) and CATS (Cost of Averting a Ton Spilled). But unfortunately, using some extremely uncompelling arguments, they attempt to come up with numbers for these two coefficients. The last time the CTX looked, these numbers were $3,000,000 per death and $60,000 per ton. Based on the way things are going, these highly debatable guesses will become fixed into law. And the only thing we will know about a particular RCO is a bogus net dollar amount.

[7] In doing so, it is extremely important to avoid *ratio tests*. Ratio tests are usually concocted by dividing the discounted benefit by the discounted cost — the max bang per buck mentality. This does not work. A possible tanker risk control option would be a 1% increase in design loads which would generate a slightly stronger ship. Another possible RCO is a 5% increase in design loads. Still another possible RCO is a 20% increase in design loads. The 1% increase will generate more benefit per dollar, than the 5% increase which in turn will generate more benefit per dollar than the 20%. This does not necessarily mean that society should stop at a 1% increase. Society should increase the design loads to the point where the marginal increase in lives saved and spillage avoided is equal to the marginal cost of the extra steel.

This is elementary economics. Despite this the IMO methodology is rife with ratio tests.

In the case of FSA, benefit and cost are inverted to come up with a cost per unit benefit. This not only has the same problem with scale as our benefit per cost ratio, but allows for all sorts of manipulation. Often it is a matter of judgement whether a portion of the benefits is monetized. By moving a portion of the demonimator to the numerator or vice versa, one can flip the ratios all over the place. Finally, such ratios cannot handle multi-dimensional problems. Almost any really useful tanker RCO will reduce both deaths and oil spillage. In this case, it is not at all clear what clearly conservative combinations of \$/death and \$/vol. If an RCO can't do this, then society should probably hold off mandating it.

The CTX methodology does not require somebody to tell it what the value of a death is nor what the cost of a ton spilled is. What the CTX system does, after tossing out the clearly bad ideas, is say to the body politic:

- here's the expected cost,
- here's the expected lives saved,
- here's the expected reduction in spillage,

you decide whether the cost is worth it.[8]

The CTX methodology is objective, the assumptions are clear, and the overall approach is simple enough so that it can actually be done with the data we normally have available.

is the denominator. The answer of course is there should be no denominator. Renounce all ratio tests.

[8] Leaving the evaluation to the body politic, allows us to address some of the locational issues we decided to ignore in step 2. Suppose a certain RCO applies only to tankers which will trade all their lives in the Baltic. And suppose society believes that spillage in the Baltic is worse than spillage elsewhere. Then it would be rational for society to have a larger acceptance region for this RCO, than an RCO aimed at worldwide trading.

If spillage location really is important, the analysis can attempt to break down the spillage reduction by region, and present those regional numbers to the body politic.