

Uses and Abuses of Ship Casualty Data

Jack Devanney

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

Abstract

Ship casualty data has been collected at least since the inception of marine underwriting. From time to time, attempts have been made to analyze this data to derive ship design and regulatory recommendations. Unfortunately, these efforts have been severely handicapped, if not totally undermined, by problems with the data itself. The casualty data is censored, subjective, and almost always unauditible. Problems with the data are compounded by the design of just about all maritime casualty databases which attempt to divide casualties into overlapping categories, confuse cause and effect, and usually fail to record any real causal information. This makes it easy to generate often meaningless correlations which can be manipulated to produce just about any result desired by any special interest group.

This paper outlines these problems, recommends that a consortium of port states set up their own fully public, freely available ship casualty data base, and offers an example of what that database should look like.

Keywords

Maritime Regulation; Ship Casualties

1 Introduction

Ship casualty data has been collected at least since the inception of marine underwriting. From time to time, attempts have been made to analyze this data to derive ship design and regulatory recommendations. Recently, this process been enshrined in a bureaucratic procedure known as Formal Safety Assessment, which promotes overall statistical analyses over detailed study of individual casualties.[3] Unfortunately, these efforts have been severely handicapped, if not totally undermined, by problems with the data itself. Ship casualty data is drastically censored, highly subjective, and, due to proprietary restrictions, almost always unauditible.

Problems with the data are compounded by the design of just about all maritime casualty databases which attempt to divide casualties into overlapping categories, confuse cause and effect, and usually fail to record any real causal information. This makes it easy to generate often meaningless correlations which can be manipulated to produce just about any result desired by any special interest group.

This paper outlines these problems and suggests some remedies.

2 The Problems

2.1 Censored data

The first problem with maritime casualty data is that it is largely missing. Crews don't report problems to their owner/manager unless they absolutely must. They know any problem will be regarded as a black mark. A problem free ship is a good ship. They also know that most owners don't want to hear about problems and certainly don't want a paper trail.¹ Almost all minor casualties go unreported.

Owners in turn don't report problems to Class and their underwriters unless they are forced to or decide to make a claim. Even rather major casualties can be covered up at this level. We had one class of ULCC's, built in 2002-2003, which were prone to fuel oil piping leaks. These ships had at least 38 fuel oil piping failures in about 10 ship-years. On one ship, the Hellepont Tara, the fuel oil piping began leaking on her maiden laden voyage while headed down the east coast of Africa. The ship had so many of these leaks, that the Chief quickly went through his entire stock of spares. He was forced to shut down one cylinder. More pipes began leaking. We were forced to go to the engine manufacturer, Sulzer, and ask if the engine could be operated with two cylinders shut down. The reply "Maybe, if you go slow enough". The Chief managed to make a temporary repair of some of the pipes and the Tara, with 420,000 tons of cargo, limped down the coast to East London, where we were able to helicopter out some spares.

These problems were never reported to Class. Nor was a catastrophic cylinder liner failure which immobilized a ULCC in the Gulf of Mexico for 12 hours, nor two badly cracked turbo-charger diffusers, nor at least six main bearing failures. All in about ten ship-years of operation.

And that was nothing compared to the State. In 1986, we bought a British-built VLCC and called her the State. Her Class records were perfectly clean. But we quickly found her engine room was an operational disaster. Breakdown after breakdown. The generator control system was so erratic and prone to blackouts, that we ended up gagging the governors, an extremely dangerous practice. The State's worst habit was snapping main maneuvering valve stems, rendering the ship powerless for 10 to 12 hours. This happened four times in the less than two years we owned her. None of these problems were reported to Class. When we sold her in 1988, the State's Class records were clean as a whistle.

And when Class or the insurers do learn about a casualty that data is by contract confidential. Here's Lloyds Register's wording:

LR will keep confidential and not use or disclose to any third party any technical information or operating data derived from the Client in connection with the Services except as may be required by law or as may be requested by the Client. This obligation will survive termination of the Contract.

Many structural problems get buried at this level. Class usually hears about major structural failures. But that is as far as it goes. On 2002-12-31, one of our ULCC's the Hellepont Embassy was southbound in the Suez Canal in ballast. The pilot called for hard port, and then hard starboard. When the helmsmen executed the second order, the steering failed to respond. The ship struck the bank hard in way of 1 Port. The impact resulted in a very large hull indent and some of the internals were buckled. The ship had to drydock and the repair required some 50 tons of steel. Class attended. There was an insurance claim, and plenty of paperwork. Yet this casualty has never made it into any of the public databases.²

¹ The ISM Code says "the Safety Management System should include procedures ensuring that non-conformities, accidents, and hazardous situations are reported to the Company." The main effect this has had is to create a time-consuming paperwork game between the owners and their auditors, in which the paperwork rarely matches what is really going on on-board.

² A particularly bizarre result of Class confidentiality is *non-dimensionalization* of hull penetration data. The only entities that have a reasonably complete database on hull penetration are the Classification Societies collectively. But when IMO turned to the Classes for that information, they ran up against the confidentiality clauses in the contracts between Class and individual owner. To get around this, Class non-dimensionalized the data by ratioing it to the

Even some very public casualties don't make it into the LMIU and other databases. In August, 1988, one of our ULCC's, the Orpheum, had a major boiler fire off Fujairah when re-lighting her boiler after five day shutdown. We were very lucky to contain the fire to the engine room. The entire area above the boiler was a mess. The massive economizer was melted down. The ship had to be towed to Dubai Drydocks for a six week, multi-million dollar repair. The casualty was well-known on the Dubai waterfront. Yet it does not show up in the LMIU database.

According to IMO resolutions, the flag states are supposed to investigate Serious and Very Serious Casualties and submit a report to IMO But the Flags of Convenience rarely do.³ In February, 2001, the loaded 30,000 ton tanker KRISTAL suddenly broke in two in bad weather off northwest Spain. 11 crew members were killed. But since her cargo was molasses, there was no public outcry. According to IMO records, no investigation report had been received from the flag state, Malta.[12] In the few cases where a flag state investigation is done, the main purpose is to exonerate the flag state, usually by blaming the crew. Finally, these investigations are rarely made public.⁴ The successful FOC's have no interest in embarrassing their ship owner clients, which would result in their transferring to a more understanding flag.⁵

We have produced a regulatory system which hides rather than provides casualty data. The end result is that the vast majority of ship casualties go unreported or are buried in inaccessible Class/underwriter records.

There are a few exceptions that prove the rule. About 2001/2002, the US Coast Guard convinced American pilots it was in their interest to report problems while they were on-board. These reports end up in CGMIX database. When this happened, the number of reported casualties for ships under pilotage in US waters, jumped several orders of magnitude. The CTX did a comparison of the 2003 US CGMIX and LMIU databases and found that the incidence of machinery failures was as least one hundred fold higher in the USCG data than the LMIU.[6]

In international waters, unless a casualty leads to a grounding, collision, a major fire, or spill, or the ship is so desperate she is willing to sign a Open Form, it will not become public. A rough rule of thumb is that 1% of all marine casualties in international waters are reported. Almost all statistical analyses of ship casualties ignore this inconvenient truth.

2.2 Hidden, Unauditable Data

The second problem with maritime casualty data is that what data we do have is unauditable. Just about all world-wide ship casualty databases are proprietary. Some countries maintain databases of the casualties that occur in their own waters, and make this data available to the public. These include Australia, the UK and the USA. However, only private for profit groups have stepped in to combine the data worldwide, including the Lloyds Marine Intelligence Unit (LMIU), Aspen Publishers, and Environmental Research Corp.

To access these private databases, one must

- pay a substantial amount of money,

size of the ship. For example, in the case of side damage, they divided the depth of penetration by the beam of the damaged ship. This effectively hid the ship's identity, and along the way made it impossible to check the data for accuracy, completeness, etc. Even the IMO delegates who write the regulation do not see the real penetration data. Much worse, in using the non-dimensionalized data in evaluating new designs, IMO had no choice but to assume that the penetration is proportional to the size of the struck ship. For example, in the same collision, IMO assumes that a narrow ship suffers less penetration than a wide ship. This of course is total nonsense. The depth of penetration depends on the size of the striker, not the size of the struck. The whole system is not only totally opaque, but ridiculously biased against bigger ships. Non-dimensionalized penetration data is useless.

³ A search of GISIS reveals that Greece has submitted no investigation reports to IMO, nor has Cyprus. Marshall Islands has filed 1, Malta 2, and Liberia 3. Panama is a recent exception. She has filed some 40 reports, almost all covering post-2006 casualties. These mostly one-page Spanish summaries usually contain nothing useful.

⁴ The IMO resolutions require a report, but not a public report. In fact, these reports are kept in a special room at the IMO, separate from the IMO library, available only to a select few.

⁵ In 2006, Panama had an apparent change of heart; and, since then has put reports of most major casualties involving Panamanian flag ships on the web. However, these reports are window dressing. They are just a one or two page restatement of press reports in Spanish. No evidence of any real investigation. In most cases, nobody has bothered to translate these documents into English.

- much worse, accept restrictions on disclosure.

No legitimate researcher can accept restrictions on disclosure. This violates one of the most basic principles of science. If anyone publishes some sort of summary or analysis of the casualty data, then anyone else should be able to go to the data and reproduce the same summary. If he can't, we don't have science; we have advertising.

This is not a theoretical issue. In 2003-2005, the POP&C and Safedor projects spent millions of European taxpayer Euros developing, among other things, a proprietary tanker casualty database, largely based on the LMIU data.[8][7] One of the major results is that structural failure is not a particularly important cause of tanker deaths or spillage.[7] One of the few public worldwide tanker casualty databases is that maintained by the Center for Tankship Excellence (CTX).⁶ The CTX database claims that structural failure is by far the single most important cause of both tanker deaths and spillage. ***Why this fundamental, critically important difference?*** Anyone can examine the CTX database and see if they agree with the cause assignment casualty by casualty. No third party can do the same for the POP&C/Safedor data.⁷

One of the POP&C project's most surprising findings is that 66% of structural failures are due to "excessive loading".[8] This is completely contrary to the CTX database in which only one structural failure out of 119 is due to excessive loading.⁸ It is also completely contrary to my 25 years as a tanker operator. Tanker crews never intentionally over-load a ship, because they know they will be caught by the multitude of cargo surveys. And when they inadvertently over-load a ship, it is almost always by such a small margin that the over-loading has nil effect on stresses. If the POP&C result were true, it would be a bombshell; pointing to the need for new regulations, new training, new paperwork, etc. Strangely the project makes little of this astonishing finding. We need to know why and how they came to such a surprising result.

There are many other conclusions in the POP&C study that bear third party examination. For example, it would be interesting to see how they split the structural failures between "structural degradation" and "poor design/construction". And blaming 83% of powered groundings on the "squat effect" seems strange to me. Are they saying that the grounding would not have occurred if the ship had been going slower? The CTX CDB groundings do not support this conclusion.

An important ancillary benefit of openness is improvement in the data. If the casualty data is publicly available anyone can question it, correct it, and add to it; and the quality of the data, which in the private databases is often execrable, will be improved. Many seafarers, salvors, spill responders, etc know a great deal about a handful of casualties; but there is no easy way for them to contribute that knowledge. And it certainly makes little sense to contribute if the recipient then takes this contribution and turns it into his/her private property.

2.3 Subjectivity

The process of converting incomplete and sometimes conflicting casualty descriptions to a bunch of computer codes, necessarily involves subjective judgement, especially when it comes to assigning causes. This is unavoidable; but what is important is that these judgements be transparent and reviewable. Reviewable here means that not only must the individual casualty data in the database be available; but also the sources and the textual descriptions upon which the coding was based.

Consider the AEGEAN SEA. Most sources (including the Spanish courts) blame the Aegean Sea grounding on navigational/seamanship errors on the part of the Master/Pilot.

What happened was that the fully loaded, single screw Aframax was at anchor north of La Coruna when she was ordered into port. Weather was 20 to 30 knot squalls from the west. The ship needed to make a nearly 180 degree turn to enter the harbor. The ship was proceeding at very low speed both because she had just picked up her anchor, and immediately after making the turn,

⁶ The CTX database can be accessed at www.c4tx.org/ctx/job/cdb/prod/search.html.

⁷ The CTX guesses that the POP&C/Safedor database calls many structural failures, fires or explosions. Many, if not most, major tanker structural failures result in a fire or explosion. But since we cannot examine the POP&C/Safedor data, this is only a guess.

⁸ See the ENERGY CONCENTRATION. An exhausted Chief Officer was given an extremely difficult and prolonged discharge sequence by his chartering department.

she needed to pick up the pilot. After raising the anchor, the weather suddenly deteriorated as a squall with winds over 60 knots came through. The Master was turning to port. On a single screw ship, this means the rudder must push the stern to starboard. But on a fully loaded tanker, all the windage is aft. At the low forward speed, rudder forces were unable to push the stern to starboard, even though the Master correctly went full ahead. The ship went aground well before completing the turn. The ship was a double hull OBO. Oil leaking into the double bottom caught fire, the ship was rocked by a series of explosions, and she ended up being destroyed, and losing essentially all of her 77,000 tons of cargo.

The CTX database calls machinery failure the primary cause on the grounds that the casualty was caused by nil low speed maneuverability. The ship was not designed to do what she needed to do. Twin screw would almost certainly have prevented this loss and massive spill. Others argue that the Master should have realized his nearly non-existent maneuverability in these conditions and either refused to enter the port, or proceeded farther west/northwest before beginning the turn.⁹ Still others argue that the pilot station was located too deep within the harbor channel.¹⁰

The point is that there is validity in all these positions. No single cause code can fully capture what happened. The CTX is comfortable with its choice of primary cause, but other choices are defensible.

What is crucial is that the user of the data not only be able to examine the coder's choice of cause for each casualty, but that he be able to review the reasoning behind this choice. Blindly burying these critically important nuances in an unauditable statistical analysis is not the road to intelligent, effective design and regulation.

2.4 Confusing Cause and Effect

The LMIU database has a field called Initial Cause which divides casualties into a number of categories such as collision, grounding, fire/explosion, etc. Most other databases follow a similar scheme.¹¹ For example, the POP&C data base divides all casualties into six major categories: collision, contact, grounding, fire, explosion, and the strangely named Non-Accidental Structural Failure (NASF).

There are two basic problems with this approach.

1. In the real world, casualties don't partition themselves into neat categories. Many casualties involve combinations of structural failure, fire, grounding, etc. The sequence: collision, fire, grounding is not uncommon. The NASSIA casualty in the Bosphorus which killed 42 includes this sequence.¹² Is the NASSIA a collision, or a fire, or a grounding? The correct answer is "all of the above". And the correct answer for cause is "none of the above". The cause was a black out on the bulk carrier BC Shipbroker. Without electrical power, the Shipbroker had no steering, and turned into the Nassia.

Moreover, some casualties involve multiple occurrences of the same event. Some such as the HYDE PARK involve multiple collisions (and allisions) and some involve multiple groundings. In both the massive URQUIOLA and SEA EMPRESS spills, the great bulk of the spillage occurred in the secondary groundings. The SEA EMPRESS grounded at least four separate times. Roughly 2500 tons of cargo were spilled in the initial grounding. Over 69,000 additional tons were lost in the subsequent groundings.

2. Far more importantly, **collision, grounding, fire/explosion are consequences, not causes**. Something always happened first. The key question in any casualty is not whether it involved a collision, or a grounding, or fire but what caused these events. Treating something

⁹ This would have required him to predict the occurrence of a freak squall which lasted only about 15 minutes. CTX calls this a necessary cause.

¹⁰ This was clearly a contributory factor. But the location of all pilot stations is a compromise between the risk to the pilot and the risk to the ship. Thus many pilot stations are located in areas where there is very limited room to maneuver. See the OCEAN EAGLE and IRON BARON strandings. This must be recognized in the design of these ships.

¹¹ There is also a field called Secondary Cause but is it almost always blank.

¹² If you are reading this manual on-line, and a ship name occurs in small caps LIKE THIS, then by clicking on the name, your browser should display the CDB Precis file (See Section 3.7) for that casualty.

like grounding as a cause is doubly irrational. It shifts regulatory focus away from preventing the grounding in the first place and towards mandating grounding-proof ships, an impossibility. Blaming a spill on grounding is like blaming the earth for an airplane crash. **To be a cause, an event must be an error or defect or failure.**

In 2008, project Safedor presented another paper based on their database.[7] According to their data, fire and explosion are very important “causes” both for deaths and spillage. But since we don’t know what caused these fires/explosions, we can say nothing about how to prevent them. The CTX database claims that the most important cause of fire/explosion is structural failure. Could it be that some of the missing NASF failures are listed as fire/explosion? Until we see the data, we simply don’t know.

On page 18 of this paper, collision is a surprisingly important “cause” of deaths. This might lead one to focus on improving bridge management and the like. In fact, on page 29, the authors do just that. But most of the collision deaths on page 18 are from the NASSIA-BC Shipbroker (42 killed) and the NAGASAKI SPIRIT-Ocean Blessing (51 killed) collisions. In the former, the cause was a machinery failure on the BC Shipbroker. In the latter, the cause was piracy. The Ocean Blessing had been boarded by pirates. She was not under control. Her crew had either locked themselves in to avoid the pirates, or been locked in by the pirates. Either way they all died in the fire, which occurred when she T-boned the Nagasaki Spirit. There is strong evidence that the crew of the Spirit were murdered for their watches and wallets after abandoning ship. In any event, only two of her crew members lived. We need to go beyond simple categorization. And that means examining individual casualties.

It is important to note that the critique of the POP&C/Safedor analyses made in this and other sections could not be made from the data provided by these projects. If the CTX database had not existed, all the POP&C/Safedor claims, including the misleading ones, could not have been challenged. Once again this is the opposite of scientific inquiry. If you make a claim, you must provide the data on which this claim is made.

This is recognized by the IMO guidelines for Formal Safety Assessment (FSA) which says in part[3, Section 9.2.1]

Those submitting the results of an FSA process should provide timely and open access to relevant supporting documents ..

I am aware of only one major casualty data based FSA that attempts to abide by this all-important rule.[1]

2.5 Multiple Causation

Multiple causation is common place in casualties, often taking the form of the casualty would not have occurred unless A and B and C all happened. As we have seen, the Aegean Sea is a case in point. In 1978, Psaraftis et al undertook a detailed study of 75 casualties from the Greek Ministry of Merchant Marine database.[9]¹³ In only one case, were they able to assign a single cause code. At the other extreme, in one case, they required 14 cause codes. Of course, we are not allowed to examine these casualties to see why this was the case; but for now the point is multiple causation is a real phenomenon, which is ignored in just about all statistical analyses of ship casualties.

¹³ Before allowing these authors access to this data, the Ministry blanked out the names of the ships, owners, etc. and required the authors not to reveal individual casualty details. Like many flag states, Greece is more interested in preserving the confidentiality of owners whose ships have gotten into trouble, than in understanding why and how casualties occur.

Still the Greek attitude is better than some flag states. At least, they do the investigation. As of 1998, the Ministry database included some 432 detailed investigations of serious casualties involving Greek flagged ships over 1000 GRT, including the mysterious and catastrophic Irenes Serenade explosion. But all that valuable information is hidden and essentially useless.

2.6 Irrational Ratios

Ship casualty analyses are replete with ratios. The procedure is to take deaths or spill volume or whatever you are interested in and divide it by an “exposure” variable, such as ship years. The resulting ratio is often called a frequency.

The main theoretical problem with this exercise is that it assumes the variable you are interested in is linear in the exposure variable. An argument as to why this is true is almost never presented. In fact, in most cases, the assumption is not even made explicit. Is tanker oil spillage linear in ship-years or port-calls or loaded ton-miles? I suspect the answer is none of the above; but I don’t know, and neither does anybody else.

The main practical problem with these ratios is that they hide scale. Is a ratio big because the numerator is big or the denominator is small? Should society be more concerned about one death in one ship-year or one thousand deaths in ten thousand ship years? When I said that the collision deaths in [7] were mostly from the Nassia and Nagasaki Spirit casualties, I cheated. My statement was based on page 18 of [7] which is reproduced below as Figure 4. All we are actually told is that the Safedor database contains 0.00491 collision deaths per ship-year. We have no idea how many deaths are involved. Could be one; could be a thousand. My assertion is really based on the CTX database which has 106 deaths in pure 60,000 ton plus tanker collisions between 1990 and 2007, of which 93 are from the Nassia and Nagasaki Spirit.

The Safedor ratio may include the 142 killed in the Agip Abruzzo/Moby Prince collision. The tanker Agip Abruzzo was anchored off Livorno when she was hit by the ferry Moby Prince. All the deaths were on the ferry. Some say the ferry crew was distracted by a football game on TV. If you are really interested in tanker collision/allision deaths, you will obtain far more insight from examining these three casualties, then by even the most protracted study of the number 0.00491.

Looking at Figure 4 should we be more concerned about *collisions* or *explosion-operational phase* whatever that is. Obviously, to make this determination, we need to know the number of deaths involved. By themselves, ratios such as these are at best meaningless and can easily be misleading.

2.7 Statistical Significance

Statistical significance can be a real problem in analyzing marine casualty data. This is particularly true of oil spill volumes.

The problem with oil spills from a statistical point of view is that one spill can change just about any total drastically, sometimes by a factor of a thousand or more. Spills have been recorded down to one-third of a liter. The largest tanker spill so far is over 300 million liters, The spill size range is a factor of a billion. Almost all tanker spills are at the very lower end of this range. But almost all the oil that is spilled is spilled in a small handful of extremely large spills.¹⁴

This generates a big trap for the unwary. In 2005, the POP&C project published an analysis of spillage by Aframax tankers between 1978 and 2003.[8] In so doing, POP&C threw out combination carriers on the grounds that these “tanker subtypes have special features which are not representative of the whole class of tankers”. The main special feature that these ships have that older pure tankers didn’t have is a double hull. In excluding these ships, they eliminated almost all older double hulls from their study. In particular, they excluded the Aegean Sea 75,000 ton spill and the loss of the Algarrobo which killed 26 crew.

We are told the LMIU derived POP&C database contains 60 Aframax tanker spills, totalling 366,294 tons. LMIU is very poor at picking up minor spills. We can be sure that between 1978 and 2003, Aframax tankers had far more than 60 spills; but we can also be sure that all the missing spills were very small and would have had nil effect on the total volume. Of this total, three spills (Braer (88,214), Prestige (77,000), and Irenes Serenade (80,000) represent 67% of the total.¹⁵

¹⁴ See [5], Chapter 3 for a much more detailed discussion of this issue.

¹⁵ If the project had not arbitrarily thrown out the Aegean Sea (75,000 tons then four spills would have represented 73%. A great deal of mystery surrounds the Irenes Serenade. Ship was 105,000 tonner and CTX has the spill at 87,000 tons. CTX believes the most likely cause of this killer casualty was corrosion in the forepeak ballast tank, which allowed cargo vapor access to the focsle or foredeck. These vapors were ignited by sparks from the anchor chain

This dependence on a single event means that almost any statement about oil spill volumes is statistically insignificant. This unfortunate fact has not prevented the project from making some strong statements about oil spillage. These statements are largely embodied in Table 3 of [8], a portion of which I have reproduced below.

	pre-Marpol	Marpol SH	DB	DS	DH
Total volume	351,186	1,517	4,377	8,851	363

In the the POP&C database, almost all the oil (351,000 tons) was spilled by pre-Marpol single hulls. Conversely, Marpol single hulls (1,517 tons) and double hulls (363 tons) spilled very little oil.¹⁶ The authors congratulate themselves and then reach some strong conclusions.

To the authors' knowledge the data of Table 3 provide for the first time a statistical analysis of the accidental pollution performance of the different tanker hull configurations in operation. As expected, DS [Double Sided] and DB [Double Bottom] are seen to perform better than the SH-non SBT/PL [pre-Marpol] tankers and worse than the DH [Double Hull] tankers. In this respect, their performance justifies the exemptions which were included in the 2003 amendments to MARPOL. What is however surprising is the nearly excellent performance of the SH-SBT/PL [Marpol single hulls] tankers, which incidentally are seen here performing better than the DS and DB fleets. It is worth pondering here whether the regulators would have been so keen to legislate the accelerated phase-out of the SH-SBT/PL tankers had they been aware of these findings.[8]

Given the statistical insignificance of the volume numbers, these conclusions are unwarranted. One way of seeing this is by changing the sample slightly. If the project had not arbitrarily thrown out the Aegean Sea, a perfectly good Aframax double hull, then the double hulls would have spilled 75,000 tons in Table 3. And if the project had not arbitrarily focussed on 80,000 to 120,000 tons, then the database would have included the Sea Empress (72,000 tons), the Exxon Valdez (40,000 tons) — not to mention a great deal of other spillage by all types — both of which are Marpol single hulls.

The point is that, when your numbers are dependent on a very small sample of extremely large spills, you must make that sample as large as you can; and then still be very careful what you say. We could have a 300,000 ton spill tomorrow, which would completely change the decadal totals.

Somewhat surprisingly, the same thing is true of deaths. When this was written, the CTX CDB contained 1,305 casualties with 4,839 total known killed. The single worst casualty in the database in terms of deaths is the Moby Prince/Agip Abruzzo collision with 140 killed. All the other casualties have fewer than 70 deaths. At first glance, the killed numbers seem to be less prone to the single big casualty problem than spillage. But if the CTX database were enlarged slightly, we would see that this is not true. There have been a number of thousand or more deaths ship casualties. For example, the CTX CDB does not contain the Dona Paz/Vector collision on the dubious grounds that the Vector, the tanker involved, was much smaller than 5000 dwt, the CTX lower limit. The official death toll in this casualty is 1,749. However, credible sources argue the real number was higher. Most sources agree that the ferry Dona Paz was badly over-loaded with non-manifest passengers, headed to Manila for Christmas. Several sources put the death toll at over 4000.

2.8 Statistics vs Engineering

Once you have a poorly designed database with enough casualties, it is easy to substitute numbers for common sense. One can dumbly attempt to correlate anything to anything. The favorite target of this exercise is hull type. The POP&C project is one example. The authors spend a great deal

being let out. But at this stage this is conjecture.

¹⁶ I assume that almost all the Marpol single hull spillage is the Nagasaki Spirit which CTX has at 1,200 tons. But since we can't look at the data, who knows?

of effort calculating spill volume by tanker hull type, **implicitly assuming that type of hull causes spillage**.

But as we have seen, the Aegean Sea did not run aground because she was a double hull. The Braer did not drift aground because she was a pre-Marpol single hull.¹⁷ The Prestige did not break up because she was a pre-Marpol single hull.¹⁸ The Irenes Serenade did not explode because she was a pre-Marpol single hull.¹⁹ The same thing is true of the Exxon Valdez and Sea Empress, and just about every really big spill that you can name. In short, type of hull has almost nothing to do with big spill volume.

But in the POP&C/Safedor database the real cause is missing or ignored. This, coupled with a selective sample and simplistic statistics, leads the POP&C authors to the conclusion that Marpol single hulls are an effective means of reducing tanker oil spillage. In fact, a Marpol single hull will on average spill more oil than either a **similarly damaged** double hull or pre-Marpol single hull.[5][page 328] This conclusion is based not on statistics; but hydrostatics and engineering analysis.

3 The Solution

3.1 Censored data

There is not to much we can do about the data being censored at the ship and owner level. The ISM code already forbids this, but it is unenforceable.²⁰

However, there is something the port states can do at the Class level. If a ship is involved in a major casualty, where *major* should at a minimum include any casualty in which any of the following are true:

- some one was killed or seriously injured;
- oil was spilled, either cargo or bunkers;
- loss of power requiring tugs to be mobilized;
- the casualty involved a fire, collision or grounding;
- the casualty involved a structural failure requiring tugs to be mobilized, or cargo transferred, or the ship diverted;
- the casualty resulted in the ship being lost, scrapped or declared a Constructive Total Loss;

then Class must make public their complete record on this ship including all survey reports, thickness measurements, all correspondence, (as well as drawings and plans if requested). Surely, if a ship capable of killing 30 or more people, and in some cases oiling 100's of km of coast, is involved in a major casualty, the owner loses whatever bogus claim he has to privacy.

Enforcement is simple. The ships of any Class that fails to comply will be banned from the Port States' ports.

3.2 Hidden, Unauditable Data

The proprietary data problem is easily solved. Just about every casualty that comes to the attention of LMIU is reported by a number of sources. These include local newspapers, coastal state organizations, and industry newspapers and magazines. In addition, some coastal states have ca-

¹⁷ Poor seamanship allowed loose pipes to roll around on the aft deck, destroying the BFO tank vents, and letting sea water into the bunkers. This error combined with lack of twin screw redundancy caused the spill.

¹⁸ The ship was converted to SBT without properly protecting the new ballast tanks. Lousy maintenance and lousy oversight by Class resulted in a corroded structure which could not withstand a rather normal Bay of Biscay storm.

¹⁹ Probably corrosion in the forepeak ballast tank. The forepeak ballast tank is the same in all types of hulls.

²⁰ Actually, there are technological solutions that should be investigated. We have Voyage Data Recorders(VDR), Alarm Loggers, Oil Discharge Monitors. It would not be technically difficult to set up a system that extracted certain kinds of casualty data automatically and transmitted that data to a regulatory body. For now, if a ship is involved in a casualty, there should be massive fines and more importantly jail terms for the owner, if the crew fails to hit the save button on the VDR. And all the VDR data should be made public and recorded in a permanent, public data base. Currently, the VDR data almost always simply disappears.

casualty reporting systems which contain many casualties that LMIU does not. Much of this data is available on the web.

The volume of casualties we are talking about is no more than a couple a day. All the port states need to do is to set up a one person office to collect this data and enter it into a properly designed, publicly available database. Suggested additions or corrections to that database could be submitted by anybody to this office. This effort could build on the existing public databases.

In a perfect world, this would be done by the IMO. But this is not a perfect world; and in this case it is essential that this be done by the port states. IMO is a consortium of flag states, mainly FOC's. The flag states pay IMO's bills, and select IMO's top management.

The big flag states have no interest in a objective, comprehensive casualty database which could be embarrassing both to themselves and their owner clients. Nor are their partners, the Classification Societies, interested in anybody prying into the design and condition of ships that they have approved and surveyed, which have then gotten into trouble. IMO wishes to project an image that all is well; and, if there are any problems, they are the fault of the crews. In the words of ex-Secretary General O'Neill.

The IMO wishes to concentrate on the human and operational problems which contribute to tanker casualties, not the technical.

In fact IMO has a casualty database called Global Integrated Shipping Informations System (GISIS). Portions of this database are available to the public at gisis.imo.org and for really major, post-2000 casualties can contain a great deal of factual information, sometimes backed up with insightful coastal state reports. But GISIS has several major failings:

1. Many casualties are missing, and for most of the others GISIS consists largely of blanks. Unless the casualty has resulted in a coastal state report, you are unlikely to find much useful in GISIS. There is clearly no push at IMO to maintain this database.
2. There is no machine readable way to represent the casualty sequence.
3. The causal fields concentrate almost entirely on human factors, mainly crew screw ups. GISIS takes a psychologist's view of the world. In particular, there is no machine readable way to describe structural and machinery failures. Nor structural condition, nor damage location. No technical analyses could be based on GISIS data if it did exist.
4. At least for the public, the database can only be searched casualty by casualty. It is not possible to do any combined casualty analysis. For the public GISIS is not machine readable.
5. Legally, the database is not really public at all. It is subject to a very restrictive Terms of Use which, if read literally, would prevent anybody from doing anything useful with the data.

Overall GISIS demonstrates why the IMO cannot be relied on for the database we need.

3.3 Subjectivity

Given that casualty descriptions are almost always incomplete and sometime conflicting, and given the fact that exactly what caused the casualty is often a matter of debate, subjectivity in coding is unavoidable. What is important is that these judgements be transparent and reviewable.

The database itself must not only be public but also the sources upon which the coding judgements were made must be available. This can be accomplished by either including at least a synopsis of each source in the text accompanying each casualty; or, where available, links to the source. If an important source is not on the net, one of the office's jobs would be to computerize it and put it on a web site. In this manner, the office would create a linkable library of sources.

3.4 Database Design

An easily accessible, properly designed data base which avoids the major mistakes of the current databases is essential. This database must avoid the false causality of the LMIU-derived databases. The raw data must be available to anyone without specialized software. It must be operating system agnostic. It must allow a detailed description of a casualty; but at the same time recognize that for

many if not most casualties, we will have very little data. It must be expandable in the sense that new fields can be added without affecting existing software. It must be accessible through a web interface; but also allow any one who is interested to download the whole thing. It must recognize multiple causality.

A database which meets these specifications has been developed by the Center for Tankship Excellence. It currently contains some 1400 tanker and bulk carrier casualties. The database is available on the web at www.c4tx.org/ctx/job/cdb/dev/search.html. The format is self-identifying XML. Figure 1 shows an example.²¹

The XML format is both machine readable and human readable. Without consulting the manual nor even knowing anything about XML, you can make a pretty good guess as to what most of the fields mean. It requires no specialized software. If you ask any browser to display http://www.c4tx.org/ctx/job/cdb/dev/xml/ctx_core.xml and search on Exxon Valdez, you will see something that looks much like Figure 1.

This particular casualty shown in Figure 1 is obviously the Exxon Valdez stranding. It is about as detailed a casualty description as you are likely to see. Almost all the casualties in the CTX database include far less data. But when we have reasonably complete casualty data such as this, combined with detailed data on the ship, all sorts of analyses are possible, including flooding, spillage, and residual strength. Hopefully, a database capable of accepting all this data will encourage its collection.

The heart of the CTX Casualty Database is the event sequence, the lines starting with `<event` in Figure 1. For the CTX CDB, a casualty is simply a sequence of events.²² A casualty is not a collision, although it may contain one or more collision events. A casualty is not a grounding, although it may contain one or more grounding events. A casualty is not a fire/explosion although it may contain one or more conflagration events. And so on. A casualty may contain any combination of structural failure/damage events, machinery failure/damage events, attacks, bridge events (such as course, speed alterations), response events (such as coastal stage provision/denial of refuge), and many more including collisions, groundings, and fires.

In the CTX CDB, a casualty is **NOT** bound to a ship. Rather ships are bound to events. Each event may have one or more *actors* associated with it. In almost all cases, the actor is a ship, but an actor may also be a terminal, an offshore platform, or any other entity that plays an important role in the event.

Causality is handled at the event level. Each event may fall into exactly one of the following cause categories:

Necessary Cause Each casualty must have one or more *necessary* causes. A *necessary* cause is an error or defect or failure which, if had not happened, then with high probability the casualty would have been averted. Exactly one of the *necessary* causes is designated the *primary* cause. If we have little or no information on cause, the primary cause is set to **Unknown**.

Possible Cause A *possible* cause is an error or defect or failure which, if had not happened, then the casualty might not have occurred. A casualty may have zero or more possible causes.

Secondary Cause A secondary cause is an error or defect or failure which, if it had not happened, then with high probability the casualty would have been mitigated. Poor or non-existent inerting is a common example. A casualty may have zero or more secondary causes.

Non-causal The event is a consequence, not a cause. Collisions, groundings, and fires among others are consequences not causes.

The treatment of causality does three things:

1. It allows multiple causality, a commonplace in most casualties.
2. It separates the necessarily subjective cause assignment from the more factual data. The CTX CDB can be used by people who disagree with the coder's cause assignments.

²¹ A full description can be found at www.c4tx.org/ctx/pub/cdb_man_v4.pdf.

²² This is not a new idea. It goes back at least to 1982,[11] and was used again by Psaraftis et al in 1998.[10]

```

<casualty
  id="19890324_001" date="19890324" edu="8" site="Prince William Sound "
  locale="R"          tod="0009"          acc="B" lat=" 60.850" long="-146.867"
  coastal="US"        area="VZ"          weather=" 3"          vis=" 3"
  note="master not on bridge, nav error, stranded leaving Valdez      ">
  <event ec="GY" sid="1" date="19890323" tod="2353" cause="N" sure="5"
    drugs="M"
    note="VLCC with 20 man crew, leaving Valdez after loading.
          Master leaves nav to tired mate to work around ice"/>
  <event ec="NA" sid="1" date="19800324" tod="0001" cause="P" sure="5"
    tired="Y"          crew_small="Y"      gps="P"          owner_care="Y"
    note="Mate starts stbd turn too late"/>
  <event ec="WS" sid="1" tod="0009" nuc="N" str="Y" days="12" spd="12"
    hop="4.6"          lop="160.0"        cur_spd=""       cur_dir=""
    depth="15.0"       chart=""          swell="0"
    note="Hits Bligh reef 2 hrs before high tide at 11-12 kts">
    <tide datum="MLLW" next_hi_tod="0155" next_hi="+3.81"
      at_hit="+2.99" next_lo_tod="0811" next_lo="+0.0"/>
  </event>
  <event ec="HL" sid="1" vol="41000000" mat="C" hbl="Y" ds="Y" db="M"
    steel_wt="3500" ref="ntsb91, fig 8, height numbers look optimistic"
    note="160m long damage, 8 of 11 cargo tanks holed, 41000m3 spill">
    <tank code="1C" size="" perim=""
      note="ntsb way low in this tank, damage extended past cl">
      <hi xs="236.1" ys="-12.1" zs="1.0"/>
      <lo xs="274.5" ys="0.0" zs="0.0"/>
    </tank>

    .... lots more compartments ...

    <tank code="FP" size="" perim=""
      note="jack guess, nstb said nothing about FP, but it was a mess">
      <hi xs="274.6" ys="-15.0" zs="3.0"/>
      <lo xs="292.0" ys="0.0" zs="0.0"/>
    </tank>
  </event>
  <event ec="DL" sid="1" date="19890325" tod="0736"
    note="lightering to Exxon Baton Rouge begins"/>
  <event ec="DL" sid="1" date="19890330" tod="1518"
    note="lightering to Exxon San Francisco begins"/>
  <event ec="DL" sid="1" date="19890402" tod="2200"
    note="lightering to Exxon Baytown begins"/>
  <event ec="DR" sid="1" date="19890405" tod=""
    note=""/>
  <event ec="Sr" sid="1"
    note="3500 tons of steel, renamed S/R Mediterranean"/>
  <ship sid="1" imo="8414520" class="AB" name="exxon valdez"
    st="TC" dwt="214853" yob="1986" flag="US" tnks="11"
    ht="SM" grt=" 94999" status="L" cgo="PC" sts="N"
    ns="1" crew="20" ig="Y" pob="N">
  <load ref="djw1"
    draft_fp="17.07" draft_ap="17.07" heel="0.0"
    cargo_wt="174391" ballast_wt="0" fuel_wt="1336"
    note="djw1 guess to give 56 ft draft, even keel, min fuel to LA">
    <tank code="1C" pct="83.24" sg="0.897" temp="35"/>

    .... lots more tanks ...

    <tank code="F0_P_F" pct="15.00" sg="0.980"/>
  </load>
</ship>
</casualty>

```

Figure 1: Sample Navigation Error

3. Most importantly, it avoids confusing cause with effect.

Figure 2 is a different kind of casualty: a machinery failure resulting in an allision. This casualty demonstrates the need for breaking major events down into sub-events. In this case, the primary cause was loss of main propulsion power. But there were at least three machinery component failures that need to be recorded. It is interesting that the maneuvering characteristics of this ship were so bad, that even though she was making about 10 knots through the water, and never lost her steering, the momentary loss of the propeller wake rendered her uncontrollable.

```

<casualty
  id="19961214_001" date="19961214" edu="7" site="Riverwalk, New Orleans"
  locale="H" tod="1410" acc="C" lat=" 29.950" long=" -90.060"
  coastal="US" area="MR" weather="GD" vis="GD"
  note="lost power, then steerage, downbound Miss, hit Riverwalk, 62 hurt">
  <event ec="MY" sid="1" cause="P" sure="5"
    owner_care="Y" maint="Y"
    note="mn engine tripped on low LO pressure, horrible maintenance">
  <component sfi="713004" name="No 1 lube oil pump"
    failure="low pressure tripped mn engine"
    note="lo pump in bad condition">
  </component>
  <component sfi="713006" name="No 1 LO pump filter"
    failure="filter partially clogged"
    note="lube oil contaminated with fuel">
  </component>
  <component sfi="797010" name="LO pump automation"
    failure="standby pump auto-start failed"
    note="relay had high resistance">
  </component>
  </event>
  <event ec="DS" sid="1" cause="N" sure="5" design="Y" lop_hrs="0.03" ts="Y"
    note="still had rudder, 10 kts, but lost steerage anyway"/>
  <event ec="Ca" sid="1" cid="1" nuc="Y" spd="10" impact="BP" manouv="Y"
    enc="L"
    note="Turned port, full stbd did not work, hit Poydras St Wharf"/>
  <event ec="Ca" sid="2" cid="1" dop="+17" angle="45"
    note="Penetrated 50-60 ft into wharf at 40-45 deg angle."/>
  <event ec="HL" sid="1" dead="-3" hurt="-3"
    note="Forward port tanks holed. No spill reported.">
  <tank code="FP" flood="Y"/>
  <tank code="H1" flood="Y"/>
  </event>
  <event ec="HL" sid="2" dead="-3" hurt="+62"
    note=" At least 62 injured on shore, 20 mm dollar damage"/>
  <ship sid="1" imo="8715302" class="NV" name="bright field"
    st="BC" dwt=" 68200" yob="1988" flag="LR" tnks=" 7"
    ht=" " grt=" 36120" status="L" cgo="G_" sts="N"
    ns="1" crew=" " ig=" " pob="Y">
  <load draft_fp="11.96" draft_ap="12.06" sag=""
    cargo_wt="56397" ballast_wt="" fuel_wt=""
    note="cargo corn, 7 holds, river at high stage">
  </load>
  </ship>
  <ship sid="2" st="NV" name="Poydras St. Wharf">
  </ship>
</casualty>

```

Figure 2: Sample Machinery Failure

The Exxon Valdez and Bright Field are extremely unusual in that we know quite a bit about these two casualties. Figure 3 is an example of a far more common situation, a casualty for which we have almost no information. About the only thing we know about this sinking is that 24 crew were killed. The point of this sad little example is that the XML format of the CTX CDB is

flexible enough to handle everything from almost no data to a rather complete story. This allows us to start out with little more than a date and a ship name, and add information as it becomes available.

```
<casualty
  id="19900315_001" date="19900315" edu="2" site="off West Australia  "
  locale="0"      tod="  "      acc="  " lat="  " long="  "
  coastal="AU"    area="IN"      weather="GD"      vis="  "
  note="putrid ex-OBO lost in good weather, no distress signal, cause?  ">
<event ec="H_" sid="1" cause="P" sure="2" btm="M"/>
<event ec="SK" sid="1" dead="24" hurt="-3" vol="-1" mat="B" igs="M"/>
<ship sid="1" imo="6803222" class="KR" name="alexandre p"
  st="CB" dwt=" 94532" yob="  " flag="PA" tnks="  " pob="Y"
  ht="DH" grt="  " status="L" fate="SK" cgo="OI" sts="N"
  ns="1" crew="24" ig="N">
</ship>
</casualty>
```

Figure 3: Sample Casualty with Nil Data

3.5 Irrational Ratios

The solution to misleading and non-informative ratios is simple. Don't use them. They are not necessary. If you must divide, deaths or spill volume, or NOX emissions by ship-years or whatever, always accompany this number with the numerator and the denominator. After you do that, the best thing to do with the ratio is throw it away.

Sometimes people create ratios (frequencies) and then confuse these small sample frequencies with probabilities. If you really need a probability based on small sample data, go Bayesian starting with a non-informative prior. See [4] for one way to do this.

3.6 Statistical Significance

While oil spill volume is an example of a random variable with an extraordinarily wide range, the issue of statistical significance is hardly unique to ship casualties. The methods for handling this problem are well-known. We simply have to use them. And then only make statements that are consistent with the level of significance.

Any analysis of ship casualty data that publishes derived statistics without displaying at least sample size and confidence intervals or the like should be ignored.

3.7 Statistics vs Understanding

Even a well-designed database will not prevent us from producing meaningless statistical correlations if we are so minded. I am convinced that there is more to learn from a detailed individual study of those casualties for which we have a reasonably complete story, than from overall statistical analyses. If you want to understand collisions, you will learn far more from Cahill than any FSA.[2]

To this end, each casualty in the CTX database is represented by a *precis* file. This file contains text descriptions of the casualty from whatever sources we have been able to locate — including sources CTX may disagree with — and/or links to other sites that have descriptions of the casualty, the most important of which are usually the coastal state investigation reports. These *precis* files are supported by a *pics* folder for each casualty which contains whatever photos, drawing and charts we have been able to obtain relevant to the casualty. The port state database I am suggesting should have something similar. In fact, any casualty database which is not supported in this manner is incomplete and next to useless, except for those who wish to manipulate numbers in the furtherance of some special interest.

To aid in using the *precis* files, the CTX core database contains a field, cryptically called `edu=`, which grades each casualty on how complete and instructive our information is on this casualty.

Those casualties with a high grades should be studied carefully by any one who is sincerely interested in marine safety.

4 Conclusion

The current situation with respect to ship casualty data is unacceptable, and a major hindrance to improving ship design and intelligent regulation. To remedy this, a consortium of port states should set up their own ship casualty database. This database:

1. Must be fully public and freely available both through a web interface and complete download.
2. Should be both human readable and machine readable XML.
3. Must avoid the false causality and major design errors of the LMIU-derived databases.
4. Must be flexible enough to handle nil data but at the same time be able to record very detailed data.
5. Must include full Class data on at least all ships involved in major casualties.
6. Must be backed up by readily available source material documenting coding decisions.
7. Must encourage corrections and additions to the data from anybody.

References

- [1] T. Arima. Fsa study on bulk carrier safety. Technical report, Formal Safety Assessment International Conference, 2002. MSC 75/5/2.
- [2] R. Cahill. *Strandings and their Causes*. The Nautical Institute, 2002.
- [3] Maritime Safety Committee. Guidelines for formal safety assessment (fsa). Technical report, IMO, May 2007. MSC 83/INF.2, 14 May 2007.
- [4] J. Devanney and R. Stewart. Bayesian analysis of oil spill statistics. *Marine Technology*, pages 365–382, October 1974.
- [5] Jack. Devanney. *The Tankship Tromedy*. Center for Tankship Excellence, 2006.
- [6] Jack Devanney. The argument for twin screw tankers. *WMU Journal of Maritime Affairs*, pages 353–380, April 2008.
- [7] E. Eliopoulou, A. Papanikolaou, and R. Hamann. Risk analysis of large tankers. Technical report, 2th Int. Workshop on Risk Based Approaches in the Maritime Industry, March 2008.
- [8] A. Papanikolaou, E. Eliopoulou, and A. Alissafaki. Critical review of aframax tanker incidents. Technical report, ENSUS Conference, 2005.
- [9] H. Psaraftis, P. Caridis, G. Panagakos, N. Desypris, and N. Ventikos. The human element as a factor in marine accidents. Technical report, IMLA-10 Conference, 1998.
- [10] H. Psaraftis, G. Panagakos, N. Desypris, and N. Ventikos. An analysis of maritime risk factors. Technical report, ISOPE-1998 Conference, 1998.
- [11] Lloyds Casualty Information Service. Tanker casualties 1960 to 1980. Technical report, IMO, 1982.
- [12] IMO staff. Casualty statistics and investigations, fsi.3/circ.4. Technical report, International Maritime Organization, 2004.

