effects that would necessarily be borne unevenly over the fleet. This would lead to a heterogeneity within the population which was not observed.

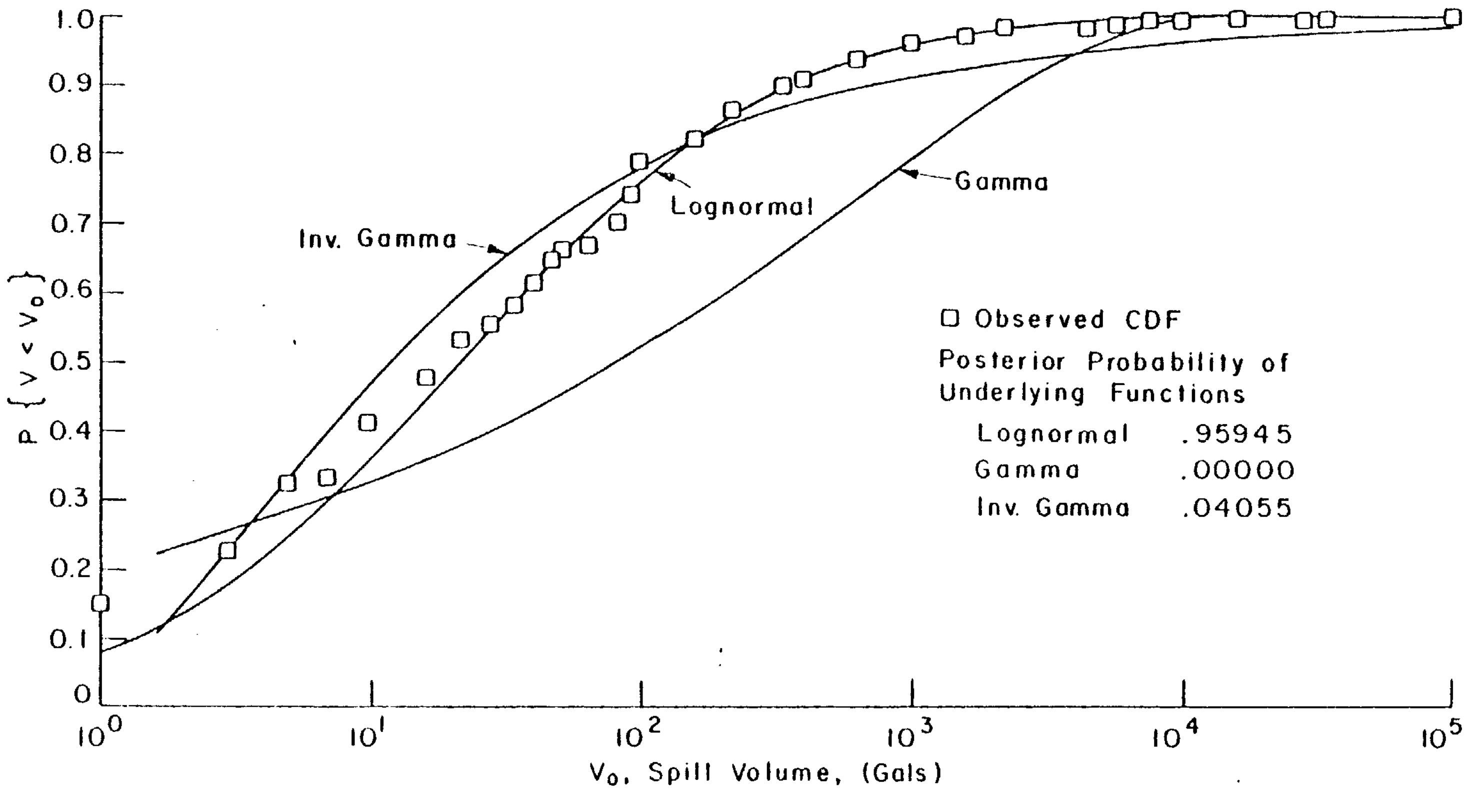
In any event, it seems reasonable to us to use the 1975 data and the Poisson model as a basis for predicting the incidence of oil spills from U.S. tankers, at least into the near future. This approach is compatible with the Bayesian techniques presently used by the Risk Analysis Group of the Department of the Interior to predict the number of oil spills from tankers. This algorithm requires historical exposure and spill number data and an estimate of the exposure accompanying the hypothetical development. For the case of predicting the total number of spills for tankers, the historical exposure is thus 235 ship-years, and the spill number is 91. The development scenario must be constructed so as to include total production and optimal ship size and route. These can be translated into ship years with a few assumptions regarding vessel speed (15 knots is reasonable), port time (2 days to offload), and time spent loading at the development or its shoreside terminal (e.g. Valdez for North Slope oil). The optimal ship size will be determined largely by draft limitations on the crude carrier route.

U.S. tanker spill volume model

The spill incidence model of the preceding section pertains to the aggregated tanker spill population. Such a model makes no distinctions based on cause or other spill descriptors. We now show that spill volume predictions can be somewhat improved if we separate the spills by the PIRS causal factor. We begin with the cumulative histogram for the aggregated spill population. This is shown in Figure 2. Also shown on this figure are the three posterior cumulative distributions and their posterior probabilities based on the sufficient statistics of the sample. As can be seen, the posterior probabilities of the lognormal, inverse-gamma and gamma were .96, .04, and .00 respectively. The posterior lognormal CDF fits the observed histogram very closely in the range above 20 gallons. Above 2,000 gallons and below 20 gallons, the observed values fall between the inverse gamma and the lognormal, except at 1 gallon. While both the lognormal and the inverse gamma fit the histogram reasonably well everywhere except in the vicinity of 1 gallon, the gamma CDF is obviously much too low in the range 10 to 3,000 gallons. This range is important because 58% of the observations fall within it. Very attractive features of both the lognormal and the inverse-gamma are their behavior in the upper tail region. The poor fits in the lower tail may be associated with data censoring, interpretive roundoff, or errors in reporting spill volume, none of which is accounted for in the theory underlying the Bayesian posterior distributions used in this study.

The .96/.04/.00 split in posterior probabilities might seem to be a rather strong statement that the lognormal is indeed the underlying distribution. However, in numerical

FIGURE 2
POSTERIOR CUMULATIVE DENSITY FUNCTION
ALL U.S. TANKER SPILLS



experiments using lognormally distributed populations, we found that the posterior probability statistic usually assumed values in excess of .99 for samples of 100, provided the variance of the $\log_e$(VARIATE) was order 1 or larger. In the case at hand, the variance of the $\log_e$(VARIATE) was 4.74. and the number of samples was 370. We inferred from this that the underlying distribution of the aggregated population was not a simple lognormal. This led us to seek ways of sorting the data to see if we could identify subpopulations whose volume distributions were at least coherent when viewed from the standpoint of yielding posterior probabilities that were close to unity. Our reasoning here was that aggregating a collection of subpopulations, each having a distinct spill volume distribution, would be one way of generating a population PDF that was not a simple lognormal.

The shape of the cumulative histogram of Figure 2 is like a flattened S, with the exception of the bumps at 5, 10, 20, and 100 gallons, the the depression between 60 and 90 gallons, which we attribute to interpretive roundoff. This suggests that if the number of subpopulations is small and if each is unimodal (both are necessary conditions for further analysis), then all subpopulations must have similar modes. A cumulative histogram constructed from a small number of subpopulations that did not satisfy this requirement would most likely lead to a curve with many points of inflection, i.e. a compound S.

Because the underlying subpopulations are liable to have similar modes provided they are few in number and unimodal, we determined that sorting based on spill volume would not be helpful. This might seem obvious, but if we had a collection of apples and watermelon, a very simple screening would be made by putting all items under 2 pounds in Basket A and all the rest in Basket B. Indeed, Danenberger (1976) attempted to sort spill data from federal OCS production activities on the basis of a 50 BBL cutoff. His idea was that the 50 BBL cutoff separated minor, or unimportant, spills from the important ones. Unfortunately, it is not clear from reviewing the properties of his big and little spills that this sorting results in identifying subpopulations that are distinct in any way except in their spill volume and, presumably, "importance".

After experimenting with several parameters, we found that a simple and useful sorting of the aggregated population was possible using the cause code. Those codes beginning with an "A" corresponded to hull ruptures, while the others corresponded to faulty equipment or faulty procedures. The median spill volume for the hull rupture was 10 gallons. The median value for the remainder was 22 gallons. However, the mean volume for the hull rupture group was 4,411 gallons, while the mean for the remainder was only 198 gallons. There were 53 hull rupture incidents and 317 non-hull-ruptures in the 1973-1975 period. Figures 3 and 4 show the cumulative histograms for the two groups.

FIGURE 3
POSTERIOR CUMULATIVE DENSITY FUNCTION
U.S. TANKER SPILLS NOT INVOLVING A HULL RUPTURE

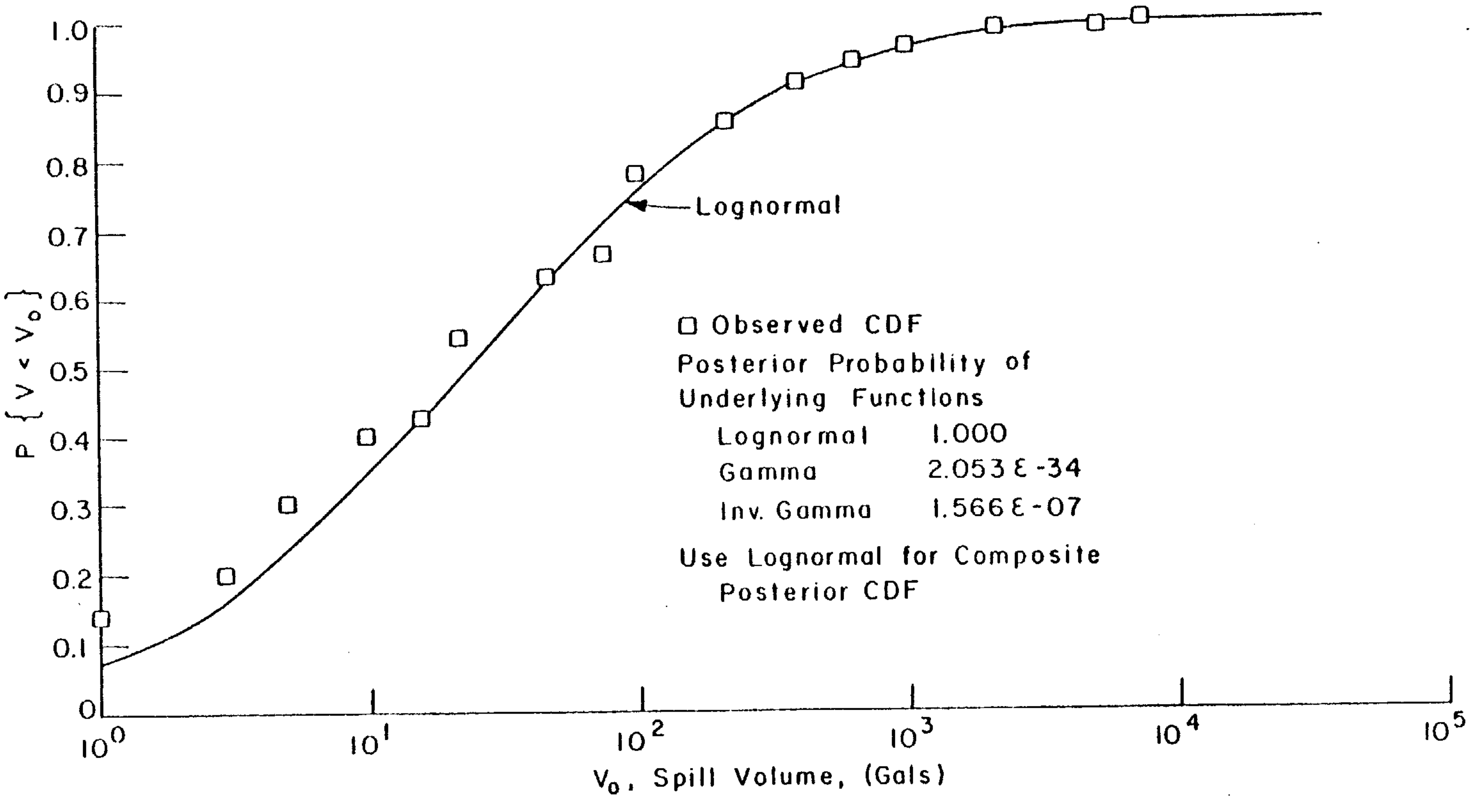
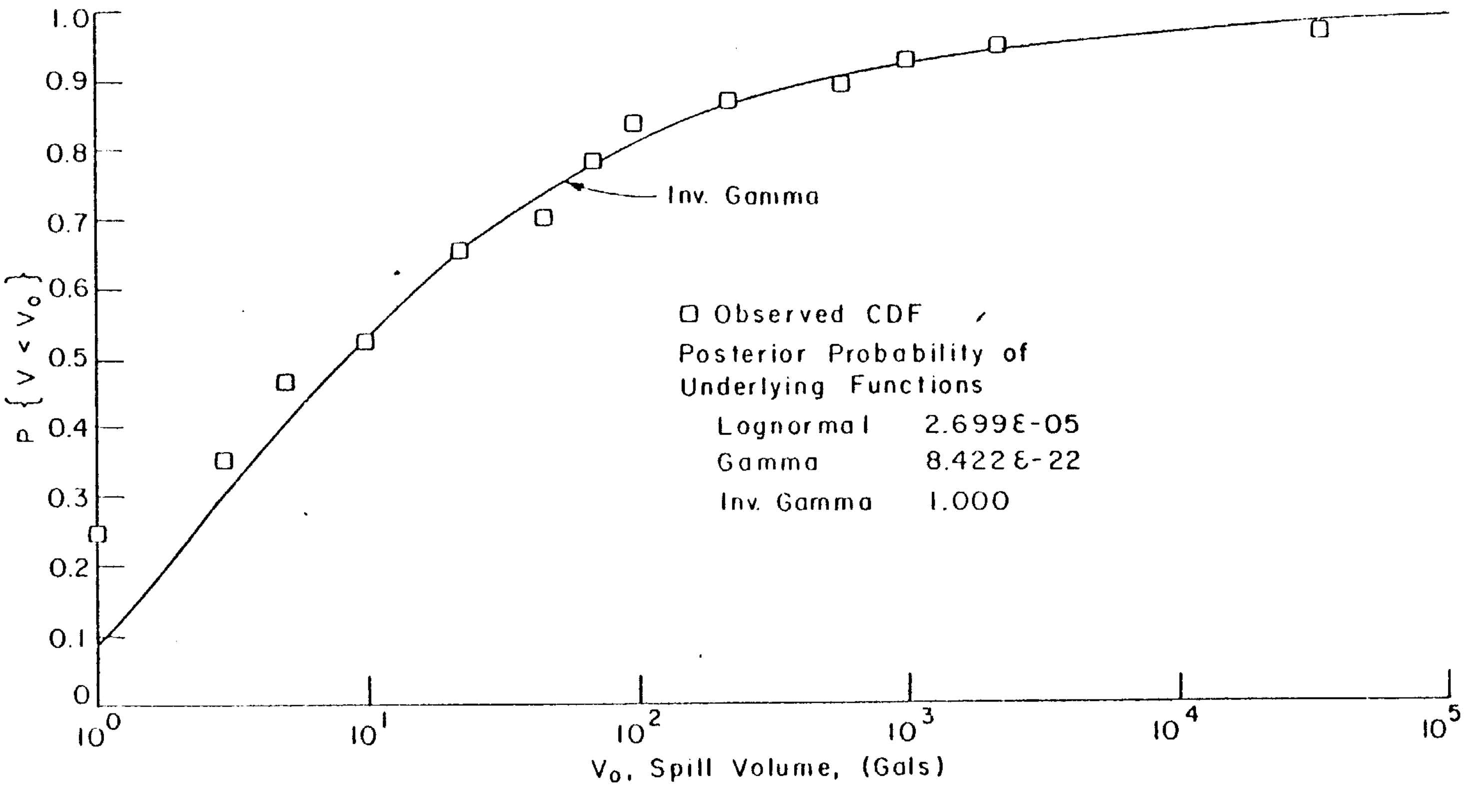


FIGURE 4
POSTERIOR CUMULATIVE DENSITY FUNCTION
U.S. TANKER SPILLS INVOLVING HULL RUPTURES



As suggested by the similarity of the median values coupled with the vastly different means, the two populations do not share the same underlying distribution. The posterior probabilities for the hull rupture spills indicated that they preferred the inverse-gamma distribution (Figure 4), while the other group showed a preference for the lognormal (Figure 3). Both posterior probabilities are so close to unity (see Figure 3 and 4) that they satisfied our initial requirements for "coherence". On the basis of the large posterior probabilities we might think there is no reason to consider further screenings. However, this may not be the case, as we explain below.

The idea that spills resulting from a hull rupture would exhibit a distinct PDF is intuitively plausible. Obviously, mending a ruptured hull is a much more difficult undertaking than closing a valve on a leaking hose. However, not all hull ruptures are alike. There are eight causes of hull rupture that accounted for one or more U.S. tanker spills in 1973 to 1975. These causes, the number of spills, and the minimum and maximum spill volumes are listed in Table 8. It is readily seen in Table 8 that the larger spills are associated with collisions, groundings, and adverse weather. Material faults, while accounting for a large number of spills, caused only small spills. The other or unknown category may be composed of small spills simply because interest in small spills is slight, and so they may not warrant the added trouble of determining their

TABLE 8
CAUSES OF HULL RUPTURES, 1973-1975

	Number of Incidents	Minimum Volume (Gals)	Maximum Volume (Gals)
Collision	5	2	2,000
Grounding	5	10	196,100
Adverse weather	2	150	1,000
Minor damage	1		1
Other Casualty	2	1	42
Material fault	15	1	100
Corrosion	1		50
Other or unknown	22	1	840

exact cause. Thus, this category might include a few groundings, collisions, and adverse weather spills.

The most interesting group of the eight is the grounding group. In addition to the 196,100 gallon and 10 gallons spills, there was also one spill of 32,110 gallons, and two spills of 20 and 40 gallons. Both of the big spills occurred while the vessel was underway; both occurred during the evening; and, both occurred in March, although the largest was in 1973, while the other was in 1975. The largest spill was caused by the tanker Hillyer Brown of 17,710 GRT, while the other large spill was caused by the tanker Colorado of 30,590 GRT. Cleanup efforts were partially successful in the Hillyer Brown spill (24% recovery) and nearly completely successful in the Colorado spill (93% recovery).

The fact that material faults and groundings form rather distinct categories within the hull rupture group suggests that a further separation of the population is desirable. This might, for example, lead to two lognormally distributed subpopulations of the hull rupture group as opposed to the present Inverse Gamma. Such a possibility is not ruled out by the Bayesian methodology, since the theory is capable of drawing comparisons only between distributions explicitly included by the analyst. A distribution formed from the sum of two distinct Lognormal distributions might under some easily imagined circumstances look more like a simple Inverse Gamma than a simple Lognormal.

Unfortunately, we have not examined the result of such a further screening, as interesting as it would be. Our reasons are twofold. First, as a practical matter, the posterior distribution of Figure 4 is very close to the cumulative histogram, and we feel it forms a reasonable analytic basis for spill volume prediction. Secondly, time limitations must at some point be heeded, and as readily as the proposed screening could be examined, we have no way of knowing that this screening would resolve the problem, or whether another iteration would be required.

There is one very important point that can be made here based on the difficulties discussed above about the Bayesian hypothesis test, in addition to emphasizing that it deals only with those distributions explicitly included in its formulation. This point is that the test is not uniform in its treatment of the three underlying distributions when the input data may consist of aggregations of distinct subpopulations. In particular, the interpretation of a posterior probability close to unity differs depending on which distribution it applies to. Because the inverse gamma has a stronger tail than the lognormal, it can play a role in an aggregation of lognormally distributed subpopulations. The reverse is not true. The lognormal, in turn, can play a role in an aggregation of gamma-distributed subpopulations. Thus, the posterior probability of nearly unity found for the non-hull-rupture group for the lognormal distribution implies more about the homogeneity of this group than does the near-unity value found for the hull-rupture group. These are rather subtle points, but since we want to use the posterior probability statistic for drawing inferences regarding the aggregated data, it is necessary that we understand this.

There remains the question of predicting the number of hull-rupture and non-hull-rupture spills. Table 9 shows the number of ships having 0, 1, 2, etc. spills for 1973-1975 from hull-rupture and non-hull-rupture causes. The Chi Square test and the Fisher dispersion test indicate that the fit of a Poisson model based on ship years isn't quite as good using this technique for categorizing the data as it was for the aggregated population (Table 3). In both 1973 and 1974, the Fisher dispersion test indicated strong heterogeneity in the population of hull-rupture spills. In 1973, this was caused by three spills from the <u>Poling Brothers No. 9</u> (1934, 1242 GRT); and in 1974, this was caused by three hull rupture spills from the <u>Santa Clara</u> previously mentioned.

TABLE 9
U.S. TANKER SPILLS BY PRIMARY CAUSAL FACTOR, 1973-1975

		Number of Spills								Fisher Test of Dispersion		
Year	Cause	0	1	2	3	4	5	$\hat{\lambda}$	χ^2	χ^2	DF	%
1973	Hull rupture	213	11	0	1	0	0	.062	$.19^{1df}$	307.4	224	.01
	Non-hull rupture	139	61	22	3	0	0	.507	1.82^{2df}	233.4	224	32
1974	Hull rupture	212	15	2	1	0	0	.096	5.39^{2df}	312.5	229	.01
	Non-hull rupture	148	64	14	2	2	0	.461	$.93^{3df}$	262.9	229	6.1
1975	Hull rupture	221	13	1	0	0	0	.064	$.72^{1df}$	251.3	234	21
	Non-hull rupture	175	46	13	0	1	0	.323	3.27^{2df}	276.5	234	2.9

However, we were already aware of the heterogeneity in the population associated with age. It is of little value to us to pursue these matters since the Chi Square goodness of fit test nowhere indicates a significant departure from the model. In dealing with the aggregated U.S. fleet, it is this goodness of fit test which is most important, since it is sensitive to the overall pattern of the fit and insensitive to small errors.

Again, the existing Offshore Risk Analysis Group spill frequency model can be used to generate spill number predictions. The historical exposure of 235 ship-years is the same for both hull-rupture and non-hull-rupture groups (the 1975 value). The number of spills is 15 for the hull-rupture group and 76 for the non-hull-rupture group (Table 9). Again, the hypothesized scenario must include sufficient information to allow calculation of the number of tanker-years required for crude transport.

A variety of interesting statistics can be estimated to characterize spill volume once we have spill number probabilities from the frequency model above. These include the median largest spill, modal largest spill, average largest spill (if an average exists), and average total spillage. The techniques for calculating these statistics require more accurate estimates of the posterior CDF than are available from Figures 3 and 4. The required offsets are provided in Appendix B. Since the calculations required for these statistics are similar for tankers, offshore production platforms, and offshore pipelines, we defer discussion of these methods to the section following the offshore pipeline section.

Intentional discharge

The above sections have dealt solely with accidental oil spills from U.S. tankers. The discussion, furthermore, was limited to those aspects of the problem that could be argued on quantitative bases with relatively few assumptions. Several aspects of the tanker spillage problem were not covered above, because they were not sufficiently well understood to allow simple quantitative arguments. These issues are, nevertheless, important to the impact assessment problem. The only question is how they may best be addressed in an impact analysis.

The problem of the intentional (and legal) discharge of oil beyond the 12-mile limit is the most important of these issues. The U.S. presently has no law which prohibits the discharge of nonpersistant oils (gasoline or kerosene), or fuel or bilge oil, beyond 12 miles, or the discharge of persistent cargo oil when it is done for the safety of the ship or the crew. If such discharges are transported to within the 12-mile limit by the action of wind and current, then the vessel responsible for the discharge may be subject to legal action. Fear of this possibility is the only constraint on intentional discharges at present.*

Intentional discharges are held to be quite common, and there is presently considerable speculation that the bulk of the oil introduced into the ocean is mainly from this source. These discharges are not solely caused by tankers, but it is suspected that tank cleaning and deballasting by tankers are the major contributors. A 1973 IMCO convention (MARPOL) which addresses this question has been proposed. It would limit total discharges to less than 1/15,000 of the vessel's DWT (1/30,000 for new vessels) per voyage. At present, it is estimated that about 1/1,000 of the DWT is discharged by vessels not employing Load On Top (LOT) procedures. Exxon, for example, uses LOT and has found that it recovers about .3% to .4% of the vessel's DWT in oil and water residues (Gray, Carven, and Becker, 1977).

The amount of oil intentionally discharged off the U.S. coast is not known. The 1973-1975 PIRS data lists only a few such incidents, and none are of any substantial size. In a few recent cases, large spills of mysterious origin have blown ashore (the Florida Keys incident in 1976, for example),

---

*It is interesting to note in this regard that the captain of the Argo Merchant was not required by any law to seek permission to discharge his cargo once he went aground. It was prudent of him to seek permission, only because the discharged oil might have carried to within the 12-mile limit. The 1969 IMCO International Convention which was invoked by the U.S. when it finally assumed control of the Argo Merchant salvage merely allows unilateral intervention after it is determined that the situation threatens the state's coastline.

but there is no simple pattern, nor is there any basis for a quantitative assessment of the magnitude of the problem.

If we accept the 1/1,000 or 1/15,000 figures, it is easy to calculate the volume of oil discharged by a tanker on her return voyage. At 300 gallons (approximately) per long ton, the average U.S. tanker (around 40,000 DWT) would intentionally (and legally) discharge 12,000 gallons and 800 gallons per voyage respectively under the two constraints. To put this in perspective, consider that one accidental spill in 20 exceeds 800 gallons, and one spill in 200 exceeds 12,000 gallons (Figure 5). Further, the average U.S. tanker will have only .4 accidental spills per year, based on the 1975 data. The same tanker will make ten to twenty voyages per year.

This indicates that the intentional discharge problem is of some importance. However, certain features distinguish intentional discharges from accidental spills. An intentional discharge is dispersed over the tanker's route, and it may only amount to a few gallons per mile. Second, the intentional discharge will occur on the open ocean. Because our understanding of the complicated oceanic ecology is insufficient to determine whether such dilute but chronic dosages are environmentally harmful, these differences make it impossible to compare this kind of pollution with that associated with a large, nearly instantaneous spill in nearshore waters, particularly on a gallon-for-gallon basis.