A Practitioner's Approach to Marine Liability Pricing Using Generalised Linear Models

Brian Gedalla, Msc, CStat, FSS
D. Jackson, BSc(Hons), FSS, and
D. E. A. Sanders, FIA, ASA, MAAA, FSS

A PRACTITIONER'S APPROACH TO MARINE LIABILITY PRICING USING GENERALISED LINEAR MODELS

By B GEDALLA MSc CStat FSS, D JACKSON BSc(Hons) FSS, and D E A SANDERS FIA ASA MAAA FSS

ABSTRACT

Marine Liability underwriters – notably those at the Protection and Indemnity (P&I) Clubs – have traditionally used empirical approaches based on individual risk experiences to arrive at their pricing. But P&I is a direct class of insurance and the underwriters have at their disposal significant data volumes. This means that it is more than possible to apply the kind of modelling techniques to P&I (and, for that matter, to other classes in the marine sector) that have become commonplace elsewhere in the General Insurance (Property & Casualty) world. In this paper we note the traditional methods, the data available and indicate how the Generalised Linear Modelling technique can be used to derive rating models that apply to Marine Liability business.

1. INTRODUCTION AND BACKGROUND

1.1.1 As much as 90% of the world's merchant shipping obtains its Marine Liability insurance via the loose network of Protection and Indemnity Associations (the "P&I Clubs") that are members of the International Group of P&I Clubs. Both shipowners and charterers enjoy the benefits of membership of a system that has survived since the middle of the nineteenth century and which itself grew out of the marine hull market based around Lloyd's of London. Traditionally, underwriting methods have been empirical in nature and tend to rely heavily on simple loss ratio statements based on simplistic experience models. Frequently, they will not even allow for IBNR/IBNER. Risks are usually underwritten as part of a fleet assessment with the historic experience of the vessels being the principal factor taken into account by the underwriter for renewals. Subjective assessments, such as the quality of the fleet's management will often influence the rating decisions.

1.1.2 Larger risks or groups of risks have historically been insured through the P&I Clubs, and the smaller risks, notably smaller vessels requiring limited liability cover or cover on a fixed premium basis, have been insured by the company/Lloyd's market, where special facilities to cater for their needs have developed. However, these facilities have not always proved profitable and few have maintained a consistent place in the market.

1.1.3 The total premium for the P&I club market is of the order of $1.8 billion premium (2002). This figure represents the total expected premium receipts, including reinsurance premiums. It is based on a total insured gross tonnage of nearly 700 million tons.

1.1.4 P&I Clubs (at least, those that are Members of the International Group) are pure mutuals and are owned by their insured members. A typical P&I Club will have two groups of Directors – the first being the Club's main Board who will be elected from amongst the shipowning membership. However, for day-to-day matters, the shipowners are usually

content to cede control to the insurance professionals who make up the management companies that run the Clubs. The second group is the management company who will have its own senior management and Directors, who are sometimes, but not always, subject to the formal approval of the Shipowners Board

1 1 5 Shipowners or charterers insure their vessels by "entering" the risks with one or more of the Clubs The shipowner will agree with the underwriter a premium rate per ton entered for each vessel. This is usually known as the "Expected Total Call" (ETC) (the term varies from Club to Club). The Club will expect an agreed proportion of this ETC to be paid up front as an "Advance Call". In recent years, many Clubs have increased the proportion of the ETC called in advance and one Club has recently announced that for 2004/05, the Advance Call will be 100% of the ETC. Where the Advance Call is less than 100% of the ETC, the remainder will be called by the Club at a later stage, possibly a year or eighteen months after the start of the Policy Year A policy year traditionally begins on the 20th February.

1.1.6 Most Clubs review their expected ultimate losses at regular intervals with a view to closing the Policy Year three years after its start Being mutuals, they reserve the right to ask their shipowners for additional premium at any time up to the date of closure Their record of collecting these additional premiums is good, with members' bad debt normally running at less than 5% of total premium Members seeking to leave the Club before the policy year is closed can usually expect to pay a "Release Call", which would normally be set at the Club's highest level of probable future Calls on that policy year.

1.1.7 The Clubs in the International Group operate a Claims Pooling agreement where large claims are shared between them on an equitable basis derived from their entered tonnages, Called Premium and aggregate claims experience over some twenty years. This Pooling agreement has operated since 1993 in two layers, currently between the Pool Retention of $5m and the Upper Pool Limit of $30m.

1.1.8 Beyond $30m, the International Group jointly purchase Excess of Loss Reinsurance in the London Market, using a multi-layered programme. For some years, the Clubs themselves have co-insured up to 25% of the working layer of this programme.

1.1 9 The Group reinsurance currently runs to losses of $2.03bn. Beyond that point, a claim, should it occur, would revert to the Clubs. Funding for such a loss would come from a variety of sources, including overspill reinsurances taken out by some Clubs, calls on Club reserves, and ultimately (as the Clubs are Mutuals) by direct Calls on the members.

1 1.10 Set out in Figure 1 below is a pictorial representation of the 2002 International Group reinsurance programme

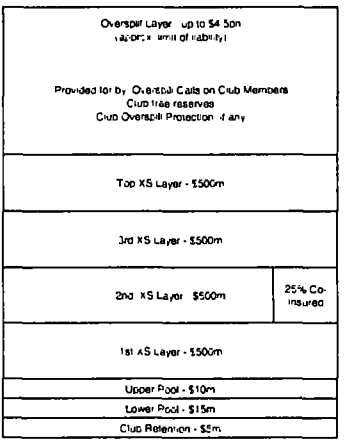

Figure 1 - The 2002 International Group Reinsurance Programme.

1.1.11 Until the late 1990s, standard P&I cover was unlimited, so theoretically a major catastrophe could result in financially crippling calls that could threaten the entire system. Because of this danger, the Group has for some years limited oil pollution risks to first $400m and more recently $500m. For the last few years, the Group has also imposed a liability limit for non-oil pollution claims, using a tonnage based formula derived from the 1976 Athens Convention non-cargo liability limits. This limit effectively establishes a cap on liability claims of around $4.5bn.

1.1.12 Most P&I claims are actually quite small, with only a dozen or so breaching the $5m Pool Retention each year. The largest ever P&I claim was the Exxon Valdez loss in 1989, believed to have cost around $8bn in total. However, as this was oil pollution, the Group loss was limited to the then limit of $400m. The largest non-oil pollution case remains the Betelgeuse loss in 1978 (an explosion off the Irish coast that resulted in several crew deaths), which cost approximately $118m.

1.1.13 It should be noted that there have been incidents in the past which could easily have generated much larger claims. Perhaps the most well known of these was the Texas City explosion in the late 1940s. The cost of that loss at today's prices would run to several billion US dollars - and that loss occurred before US courts started imposing punitive damages on top of other claims.

1.1.14 Today, the P&I risk, while limiting oil pollution losses, still leaves the Clubs exposed to some potential large losses, such as Liquid Petroleum Gas (LPG) tanker explosions and the potentially catastrophic impact of a major passenger cruise liner loss.

2. CURRENT RATING METHODS AND UNDERWRITING MODELS

2.1.1 At some stage during the months leading up to Renewal, every ship owning member or his Broker will have been presented with evidence of his loss record (going back over a period of years). The Club Underwriters will have discussed with the shipowner the Club's overall financial position together with the general level of increase that the Club's Shipowning Board of Directors will have agreed early in the Season should be applied to all Members' rates at the start of negotiations. The Shipowner will argue, perhaps, that they are a special case - they have implemented new stringent levels of ship management and loss prevention; they have replaced ageing elements of their fleet with new state-of-the-art vessels; they no longer carry dangerous cargoes; they no longer sail into potentially litigious US ports and so on The Shipowner will offer to increase the deductibles operating on their Policy and there will be a healthy discussion as to the effect such an increase might have on the loss ratios.

2.1.2 Clubs use different techniques to aid their arguments Some will rely on fairly simple gross loss ratio calculations, while others will present rather more sophisticated pricing models to support the discussion In the end, however, a deal will be done and the business duly renewed.

2.1.3 It is a testimony to the stability of the International Group system that surprisingly little tonnage moves between Clubs at the 20th February Renewal During any year, mergers and acquisitions between Shipowners result in vessels being moved from Club to Club, but a feature of the renewals process in recent years has been that the vast majority of Shipowners stay with their Clubs. Increasingly, larger (and not so large) Shipowners choose to belong to more than one Club, entering some vessels with one Club, some with another, or occasionally splitting their entry pro-rata between Clubs, so that each Club has, for example, 50% of each vessel in a group of vessels. Such Shipowners may vary their distribution of vessels between their Clubs at renewal, but again, few will make radical changes.

2 1.4 Against this background, rates have fallen in the 1990s. Underwriters always talk of insurance cycles and certainly a soft rate cycle afflicted Lloyd's in those years It is undoubtedly true that Hull rates fell in the London Market and this generated pressure from Shipowners and Brokers for P&I Underwriters to follow suit. Counter-arguments that Hull and P&I insurance are completely different have tended to fall on deaf ears and the perceived threat from the entry into the market of fixed premium writers reinforced the pressure. The Clubs, it is pointed out, are pure mutuals and their substantial assets are ultimately the property of the Shipowning members. These Shipowning members feel that it is not unreasonable to expect the Clubs to release free reserves in the soft years - reserves that have been built up in the harder years of the cycle, when higher premiums were collected.

2 1.5 Since 2000, a new realism has gripped the market and Rates have substantially increased in the last 2-3 years and are continuing to rise. Most Clubs still believe their rates are too low and that they are continuing to draw down on their reserves Accordingly, typical general increases sought by the Clubs for the 2004 renewal are still in excess of 15%

2.2 Original Rating Process

2.2.1 Most P&I Club rating procedures currently in force are based on a simple model, with premium rates based on tonnage. Typically, there are several deductions from the gross premium to derive the retained premium, and these are assessed against historic experience on a judgemental basis to reach an acceptable ratio of retained to gross premium. This type of process is typical for many types of risk in the London market.

The deductions may include the following:

- Excess of Loss Reinsurance
- A premium in respect of the upper Pool ($20m - $30m) of losses, possibly based on the reinsurance premium
- A premium in respect of the lower Pool ($5m -$20m). Again possibly based on the reinsurance premium or on some function of the Club's contribution level to lower pool claims.
- An abatement layer below the Pool to smooth out the effect of large losses. Depending on the size of the Club, this might be set at any point between, $100,000 and $2 million and cover the layer from the abatement point to the Pool retention of $5 million.
- Alternatively, some models may not make any allowance for an abatement layer, but may cap claims at the Pool retention point, currently $5 million

2.2.2 The remaining net premium is used to assess the retained loss ratio = premium net of deductions/gross premium. The retained premium for the insured vessel is compared against the corresponding losses If a shortfall arises the rate is adjusted upwards

2.3 Underwriting Models

2.3.1 Larger fleets may be broken down into roughly homogeneous groups of vessels (crude oil tankers, for example, may be assessed together), but it is unusual for the assessment to be any more detailed The simplest underwriting models may do little more than calculate the historic gross loss ratios by underwriting year, with no adjustment for unexpired risk, IBNR or unallocated expenses. These simple calculations will be used to judge whether the rating group is profitable From this judgement, a loading will be applied in addition to the overall increase previously agreed by the Club's board.

2.3.2 There are more sophisticated models in the market P&I Clubs in the International Group pool their losses above $5 million and collectively purchase Excess of Loss reinsurance above $30 million, one variation on the basic loss ratio model, is to cap claims at the $5 million retention point and apply an overall loading to account for the Club's share of Pool and reinsurance claims. A variation on this theme is to recognise that $5 million is far too high a point to share large claims without seriously distorting the loss ratio model for those fleets with a large claim. Therefore, the abatement layers described above are introduced to smooth out the distortions

2.3 3 Other models do exist, with adjustments for IBNR, IBNER, expenses and so on. Some models attempt to relate premium to the risk by developing a simple burning cost model, based on losses per entered ton.

2 3.4 The common factor in all these models is that they are essentially one-dimensional or at best two-dimensional, and make no real statistical use of the wealth of data held on the underwriting systems, and the interaction between the various factors that drive the claims experience

2.4 A Different Approach

2 4.1 The original approach outlined above is simplistic in that it does not fully reflect all of the factors underlying an insured's experience. Using a multifactor approach should give rise to consistent internal premium rates, with the need to increase rates only to reflect the overall market condition, or individual risks that perform badly as a result of poor risk management. A Generalised Linear Model (GLM) approach gives a more scientific basis for estimating rates.

2.4.2 A GLM creates a multi-dimensional representation of the data that enables the inter-dependent relationships in the data to be visualised in a way quite impossible by inspection alone. Such relationships are obvious when there are only two rating factors and can be identified by simple one and two-way tables Even with three factors, and a fair amount of patience, the various combinations of tabular analyses can be explored. But once the number of variables starts to climb, this quickly becomes impossible GLMs explore the data using powerful statistical software and establish the relativities present, as well as evaluating the statistical errors associated with the models derived. In this way, the actuary or statistician can evaluate the possible solutions indicated by the modelling process and select the models that best explain the variation in the data.

2.4 3 GLMs also give an equitable approach to rating between the various fleets or Club members. The rating could be readily extended to the higher layers to allow for the abatements and reinsurance premiums

3. DATA SELECTION

3 1.1 The key to carrying out the GLM modelling process successfully is to obtain as much data from internal underwriting and claims systems as possible. It is important to capture both sides of the data store as valuable descriptive information will often only be reliably held on the underwriting system while the detailed claims cost information will usually only be held on the claims system

3.1 2 The data should be extracted from these systems on an individual risk basis together with measures of exposure period. If a policy has an adjustment mid-term, resulting in a change to information we would wish to use as a rating factor, there should be a single record entry representing each of the rating factors applicable to each portion of the policy. This should

not present too much of a problem with P&I business as the incidence of mid-term adjustments to policies is very rare. If an exposure period cannot be calculated within the underwriting system, then enough date information should be extracted for each record produced to allow the accurate calculation of exposure periods

3.1.3 The data extracted for the exercise should be of a recent nature and of sufficient volume to ensure the models fitted accurately reflect the expected claims experience going forward. It is normally expected to cover around four or five policy year's worth of data, although more is acceptable if available Care needs to be taken though as including older data may result in the model no longer reflecting current experience A rough guide is to use a minimum exposure of approximately 15,000 vessel years

3.1.4 The data needed will most likely come as two sources. 1) an underwriting file consisting of single records for each exposure unit – probably a "vessel-year" – representing a unique combination of rating factor details for each period of risk, and 2) a detailed claims file with information relating to every claim incurred by the exposure units on the underwriting file

3.2 Rating Factors

3.2.1 Most existing pricing models analyse actual experience by underwriting year, maybe split according to some approximate vessel classification within a fleet There is a wide range of classifying factors about each vessel routinely captured by the underwriting systems and several of these can be used to analyse the risks. The levels to be modelled for each rating factor are generally easily determined by the nature of a particular rating factor. However, for rating factors with a large number of levels it is more practical to group together levels with similar properties so that more stable parameter estimates are produced within the GLM model.

3.2.2 The common rating factors used in Marine Liability pricing are:

- Type of vessel,
- Age of vessel,
- Classification society (Lloyd's Register, the American Bureau and so on),
- Vessel flag,
- Nationality,
- Tonnage either in terms of gross tonnage or entered tonnage,
- Various types of deductibles.

Other factors can be identified from the existing data such as those vessels with limited liability or those extending their standard P&I Cover to include 4/4ths of vessel collision claims, otherwise known as Running Down Costs (RDC), which by maritime tradition are normally split between P&I and Hull insurances

3.2.3 Up to 150 different types of vessel exist, but for rating purposes these should be aggregated into 10 or so categories at the most. It is practical at this stage to identify vessels that carry dry cargo or tankers carrying clean cargo and rate these as separate factor levels, as different

reinsurance arrangements will apply to these vessels at a later stage. The age of vessel factor is most sensibly grouped into bands of 5 years. The classification society factor is usually grouped into 10 or 11 levels representing the major societies plus an 'other' category containing the smaller societies and those vessels where the classification society cannot be identified or is not recorded. The vessel flag factor can usually be split into levels representing 12 or so of the major flag nations plus one level representing the smaller flag nations combined experience

3.2.4 The nationality factor represents the vessel owner's country of origin. This is normally best grouped together by geographical region with the major countries such as Greece, USA, Russia, and China identified separately. Deductibles are best grouped into 5 or 6 bands representing the amount of deductible taken. The types of deductible being taken are identifiable from the Rule codes they are attached to and are therefore easily classified into types consistent with the types of claims being analysed. Typical types of deductible are.

- Collision with other vessels,
- Collision with fixed and floating objects,
- Pollution,
- Cargo,
- Personal injury,
- Other

Personal injury deductible types can further be split into crew, passenger and other personal injury if desired. Some vessels can be subject to an all-claims deductible, rather than picking up one or more of these individual deductibles separately.

3.3 Claims Data

3.3.1 Different insurers hold different levels of detail on their claims systems. Some may be able to provide little beyond a total amount paid and a total outstanding estimate for each claim. For modelling purposes however, much more detailed information is required

3.3.2 A P&I Claim can include claims of various types, such as collision damage both in terms of collisions with other vessels and collision with fixed and floating objects, pollution, cargo, personal injury and others. The personal injury element of the claim could also be split further into crew passenger injury, passenger personal injury, stevedore injuries and other injury types if so desired. The type of claim can normally be determined relatively easily, as different aspects of the claims transaction information are normally assigned to Rule codes External fees relating to each claim should be included in the claims amounts to be modelled and are most easily analysed when they are assigned directly to the relevant Rule code for the claim that they apply to

3.3.3 The claims file should be provided on a full transactional basis, allowing full analysis of all claims incurred. In order that the model is fitted to data representing a stable and settled claims position, each incurred claims amount should be increased to take account of any

IBNER The IBNER factors are usually derived from other reserving work carried out upon the same book of business and are applied on a policy year and type of claim basis. The claims transaction file is then summarised on an individual incident per policy per vessel basis.

3.3.4 Each of the individual claims should be capped at an appropriate level to remove the effect of large claims The level at which the claims are capped may be predetermined by the choice of a particular P&I Club's abatement level In other cases this level is too high and an arbitrary figure of $100,000 is chosen. It is also useful to cap individual claims amounts at the abatement level used and at any retention amounts that apply, so that appropriate loadings can be evaluated and applied at a later stage

3.4 Preparation of Modelling Data

3.4.1 The task then is to merge the files, eliminate errors and aggregate claims costs, so as to end up with a manageable data file containing one record for each exposure unit with summarised claims information appended Inevitably at this stage there will be some degree of mismatch when linking the claims data to the underwriting data. Care needs to be taken here to ensure the mismatched claims are investigated. If the mismatched claims are for business to be included in the model, then the mismatch amount needs to be evaluated and a loading for this should be applied at a later stage in the modelling process

3.4.2 Once a single manageable data file has been produced, the data must be examined and any records representing business not required in the model should be removed For instance, a particular Club may want to fit a model to owned-only vessels and not chartered vessels, or they may wish to exclude those vessels insured under consortium arrangements and price this business separately

4. THE GENERALISED LINEAR MODELLING APPROACH

4.1 Modelling

4.1.1 Having generated the database, the most important stage of our work is the modelling process itself For some time now, actuaries and statisticians have been applying a class of mathematical models known as GLMs to mass-volume insurance data to identify relationships between risks and establish relativities between different levels of rating factors.

4.1.2 The underlying assumption in rating Marine Liability business is that the risks are similar in many aspects to those found in personal lines insurance, in particular those found in motor insurance. P&I club risks covered are usually single vessels, each of which is considered to be comparable to a private motor policy Large fleets of vessels on cover are considered to be comparable to a motor fleet policy

4.1.3 There have been a large number of papers written which make use of the GLM techniques to rate motor business. The underlying theory we have used to form the basis of this paper can be found in Brockman & Wright (1992).

4.1.4 We have fitted a basic frequency-severity model to the risk premium per ton calculated using the capped incurred claims cost, using a poisson error structure with a log link, weighted by exposure measured in terms of entered tonnage The log link results in a multiplicative model being fitted, which is preferred as negative fitted values cannot be obtained from the model, unlike in an additive model. Also, a greater level of accuracy can be obtained by fitting multiplicative models as opposed to additive models, where more terms would need to be included in the model to achieve the same outcome. We use a poisson error structure as the incidence of P&I claims being modelled are measured in terms of claims cost per ton over a fixed time period

4.1 5 Another approach is to fit separate models to frequency and severity and examine the results of the two separately. The frequency model would be fitted to the number of observed claims per ton using a poisson error structure with a log link, just like the model used above The severity model is different in that it uses a gamma error structure rather than the poisson error structure. A detailed outline of the theory for severity models can be found in Brockman and Wright (1992).

4.2 Time Dependency

4.2.1 In choosing the data to be included in the model, care must be taken to ensure the exposure periods chosen are suitably recent so that the claims experience being predicted by the model can be expected to be of a similar nature to the historical experience and that the volume of the data being used is large enough to reduce random variation in the parameter estimates. It is common practice to select data that covers the most recent four or five-year period, to ensure that both of these criteria are met. Arguably, very recent claims data should not be used in the model due to its undeveloped nature. This is easily overcome by ensuring that the claims amounts being used are incurred amounts, including both paid amounts and all outstanding estimates, together with an appropriate development for an element of IBNER.

4 2.2 When fitting the models, a time factor should be allowed for as an explanatory variable. This is to ensure the trend in the size of claims due to inflation is identified This way there is no need to remove inflation by making prior adjustments to the claims data. This claims inflation should not be assumed to be the same as the RPI inflation, or the claims inflation experienced in other lines of business. Another reason for fitting a time factor in the model is to remove the effects of any changes in portfolio mix over time as this could result in the parameter estimates being distorted.

4 2.3 To check the stability of parameter estimates over time for a particular rating factor, the selected model should be re-fitted containing an additional interaction term. This interaction term includes both time and the rating factor to be tested. Separate models should be fitted for each of the main rating factors in turn It usual to plot the results of this fit on the same graph as the results of the fit from the selected model. The graph in Figure 2 below shows

the results of fitting an additional term for the interaction of vessel type and time. We can see from this graph that when compared with the fit obtained from a main effects model fit, the same general trend across the vessel types is observed in each of the policy years under analysis.

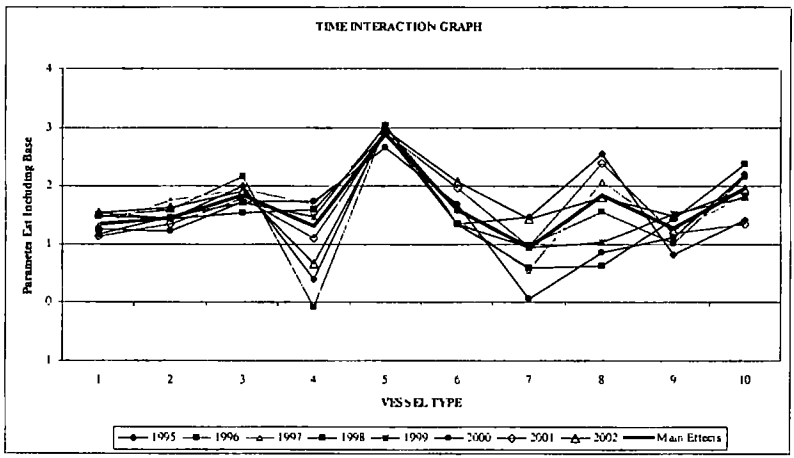

Figure 2 - Time Interaction Graph

4.2.4 If data is available for multiple claim types as mentioned earlier in section 3.3.2, separate models can be fitted to each claim type separately. This will give a much deeper insight into the factors driving the claims experience. Additionally including the time factor in the model provides the ability to estimate the claims inflation for each of the claim types separately, and also to identify trends in the data applicable to individual claim types, without being affected by changes in the portfolio mix over time.

4.2.5 The standard model assumptions of constant variance should be checked by producing a plot of standardised residuals against the fitted values, and also by producing plots of the standardised residuals against the levels of each rating factor in turn. The graphs in Figure 3 and 4 below show the plots of standardised residuals against fitted claims costs per entered ton, and standardised residuals by vessel type.

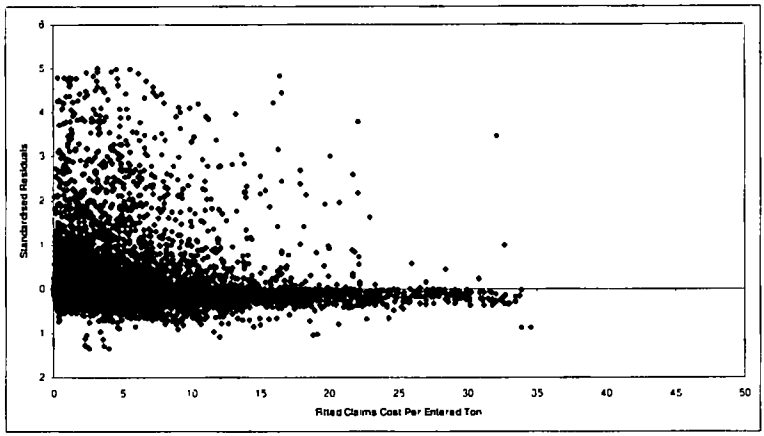

Figure 3 – Plot of Standardised Residuals versus Fitted Values

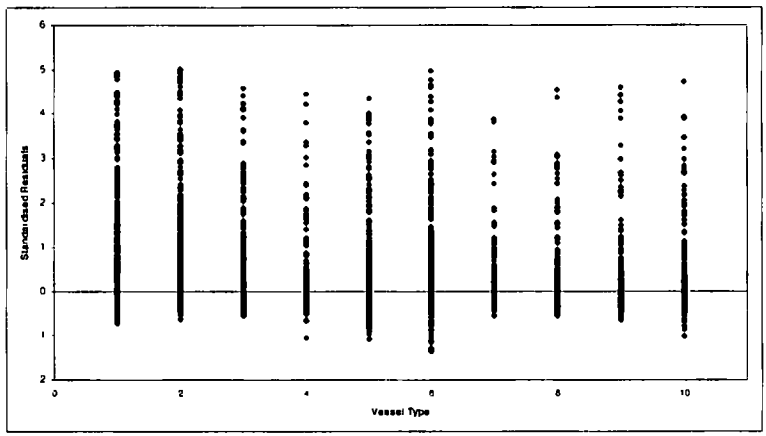

Figure 4 – Plot of Standardised Residuals By Vessel Type

4.3 Validation

4.3.1 The final stage of the modelling process is to turn the GLM output into a set of relativities together with a base rate. This base rate is taken from the model base and will need to be increased to take into account additional costs These will include a loading for any mismatched claims during the data preparation stage, loadings for the capping of claims between the chosen capping level and the abatement level (if the two are different), abated claims, expenses, the club's share of pooled claims, future inflation and reinsurance costs

4 3.2 Having derived the model, it is then applied to the underwriting information to compare the indicated premium for each risk with the actual premium charged The total indicated premium can be examined to ensure that it is sufficient to cover the historic losses.

5. CONCLUSION

5 1.1 In this paper, we have described the background to existing Marine Liability pricing models.

5 1.2 We have gone on to describe the application of powerful new modelling techniques based on Generalised Linear Models to the available data.

5.1 3 The end result is an easy to apply multiplicative rating model that can be used to derive a statistically valid premium for each vessel. Nothing in this work deprives the Underwriter of his or her ability to negotiate a different rate from that indicated by the model. However, with an appropriate modelling technique added to the toolkit of methods, the Underwriter is better placed to conduct a meaningful negotiation with the Shipowner armed with the results of a formal analysis of the past experience.

6. REFERENCES

BROCKMAN, M & WRIGHT, T (1992) *Statistical Motor Rating: Making Effective Use of Your Data* Journal of the Institute of Actuaries, **119,** 457-543

GEDALLA, B (2003) *Marine Liability Pricing* Published in Viewpoint (Autumn 2003) by Milliman UK

GEDALLA, B (2001) *Are P&I Rates Too Soft?* Published in Viewpoint (Summer 2001) by Milliman UK

MALDE, S et al (1994) *Marine Insurance.* General Insurance Study Group, 1994 General Insurance Convention,467-530