A GENERALISED PROPERTY EXPOSURE RATING FRAMEWORK THAT INCORPORATES SCALE-INDEPENDENT LOSSES AND MAXIMUM POSSIBLE LOSS UNCERTAINTY

BY

PIETRO PARODI

ABSTRACT

A generalised property exposure rating framework is presented here to address two issues arising in the standard approach to exposure rating, especially in the context of direct insurance and facultative reinsurance (D&F) property pricing:

(a) What to do when the main assumption of exposure rating, scalability – that is, that the probability of a given damage ratio does not depend on the maximum possible loss (MPL) but only on the type of property – breaks down.
(b) How to take account of the uncertainty around the MPL, that is, the fact that the MPL (unlike the insured value, IV) is an informed estimate rather than a contractual feature.

The first difficulty is addressed by making the exposure rating framework more flexible by introducing exposure curves that are a mixture of scale-dependent and scale-independent losses, and where the weight given to the two components is a function of the MPL. This allows to model the impact on the expected losses of changes in the underlying deductibles, a classic problem in D&F pricing normally solved with the help of deductible impact tables. The second difficulty is addressed by working out the mathematical implications of having a finite probability of exceeding the MPL and extending the exposure curve up to the maximum of MPL and IV. A practical application of this framework – in which the scale-independent and scale-dependent losses are identified with attritional and large losses respectively – is described. A discussion of how this generalised framework can be calibrated based on actual data is included, and implementation code for the framework is made available.

KEYWORDS

Exposure curves, scale-independent losses, maximum possible loss, insured value, impact of deductible changes.

Astin Bulletin 50(2), 513-553. doi:10.1017/asb.2020.11

1. Introduction

Property exposure rating (see references in Section 1.1) is predicated on the assumption that property losses can be best described as a proportion of the maximum possible loss (MPL), in the sense that for a given type of property (e.g., residential homes) the probability that the damage ratio X/M (where X is the loss in the monetary amount and M is the MPL) is less than or equal to a given number x is independent of M. As an example, under this assumption the probability that a house of value £180,000 suffers a loss of at most £60,000 is the same as the probability that a house of value £900,000 and similar features suffers a loss of £300,000. This allows us to build *normalised* severity curves, that is, severity curves that depend only on the damage ratio and not on the absolute monetary amount of the loss. From these severity curves we can build *exposure curves*. An exposure curve is a re-engineered version of the severity curve that allows to calculate easily the amount of losses that are retained by the (re)insured as a function of the deductible. Exposure curves are also typically normalised and are expressed in terms of the ratio between the deductible and the MPL.

The approximation that allows us to use the same normalised severity and exposure curves for the same type of property is of course only true to the extent that the properties are indeed similar. The fact that each property is unique and that two properties never share exactly the same exposure curve, is, however, only the most obvious problem with exposure rating. Another, subtler one is that some of the losses may not actually be scale-dependent, that is, they may remain the same (or scale non-proportionally) when "transferred" to a property of different value. Possible examples of this are content of fixed value or contained electrical damage.

This paper investigates how the exposure rating framework should be modified to take the presence of scale-independent losses into account. Note that although scale-independent losses will tend to be smaller, the difference between scale-dependent and scale-independent losses is not in principle a distinction between attritional and large losses: it is possible to have small scale-dependent losses and relatively large scale-independent losses. In practice, the difficulty of categorising losses into scale-dependent and scale-independent may make it necessary to equate attritional and scale-independent losses as a first approximation.

The relevance of this extended framework is that many of the exposure rating curves available in the market for large properties were derived based on losses above the underlying deductible (or "local deductible" (LD), in the sense that it applies to a specific property, and may be different for different properties in a portfolio). If the LD remains more or less constant in time (at least in real terms) and is consistent across properties (i.e., equal to a "standard" market deductible), the curves remain relevant. However, when market pressure causes deductibles to move downward (soft market) or upwards (hard market), the traditional exposure curves need to be adjusted accordingly. The traditional way of addressing this (see Parodi (2014b)) in direct insurance and facultative reinsurance (D&F) is to

- assume that the exposure curve describes only losses above the underlying deductible, that is, that it gives the percentage retained above $y = (X - LD)/(M - LD)$, where LD is the local deductible;
- have a rule that describes by what percentage the premium rate (or, better, the expected loss rate) on value should be increased if the LD is changed from a local deductible LD to a new deductible LD'. (The **premium rate** on value (ROV) is the premium divided by the insured value (IV). The **expected loss** ROV is the expected loss divided by the IV.)

This problem can be eliminated if one can somehow estimate the from-the-ground-up loss curve. However, the effect of a change in deductible is often larger than what can be obtained by a single MPL-independent exposure curve, especially with large industrial properties where MPLs in excess of \$100 m or even \$1 bn are not unusual. While part of the difference might be of a commercial nature, many underwriters argue that the main reason is that the LD allows to get rid of attritional losses that are not scale-dependent. A possible approach is to model attritional, scale-independent losses separately from large losses. Another approach – which we will attempt in this paper – is to build generalised exposure curves that somehow deal with both scale-dependent and scale-independent losses.

Another practical problem that we address here is that the MPL is not a contractual feature but an underwriter's estimate (perhaps based on an engineering assessment), and therefore it is subject to uncertainty. This means that one needs to allow for the possibility that a loss between MPL and IV (assuming that the latter is larger) occurs.

This second practical problem (MPL uncertainty) is not directly related to the first one (scale-independent losses) but a generalised framework should be able to address both problems at the same time.

1.1. Overview of related literature

There is of course a wealth of literature on exposure rating (see, e.g., Salzmann, 1963; Ludwig, 1991; Bernegger, 1997; Guggisberg, 2004; Riegel, 2010; Parodi, 2014a; Dutang *et al.*, 2019). Most of it focuses on treaty reinsurance and is not directly relevant to this paper except of course for providing the necessary foundational material on exposure curves.

To the author's knowledge, there is no literature on including both scale-dependent and scale-independent losses in a single exposure curve. However, the idea of splitting exposure curves into an attritional and a large-loss region has been adopted by some companies. The methods the author has come across, however, lacked mathematical consistency (see discussion in Section 2.1 about the requirements of well-constructed exposure curves), and did not attempt to derive exposure curves from first principles. This paper attempts to build on the original inspiration behind these heuristic approaches and develop an actuarially sound model based on the idea of including scale-independent losses in the framework.

2. Exposure rating in D&F pricing

The theory behind exposure rating is detailed in several sources (see references in Section 1.1). The context in most of these treatments is that of treaty reinsurance, for which the methodology was first developed. Here, and especially in Section 2.3, we follow mostly the treatment in Parodi (2014b), which is the most relevant to D&F pricing, and to which the reader is referred for more details.

2.1. Property exposure rating basics

Exposure curves are re-engineered versions of severity curves that allow to assess the effect on the expected losses of changing the retention level or allocate the expected losses to layers of (re)insurance. The main definitions are below.

Assume that X is the random variable representing a loss to a particular property, and that the MPL for that property is M. Since the loss is assumed not to exceed M, exposure rating normally describes any loss occurring to that property as a percentage of its maximum value M, $[X]=X/M$. $[X]$ will be called here **damage ratio**, but is also variously referred to in the literature as normalised loss, loss degree, degree of loss and fractional loss. Also, we can describe the deductible d as a percentage of M, $u=d/M$.

Let $[F](y)$ be the *normalised* cumulative distribution function (which will also be referred to as the severity curve), that is, the probability that the damage ratio is less or equal to y, and $[S](y)=1-[F](y)$ be the survival function. (In the following, we will speak about [F], [S] and [G] to denote the version of the CDF F, the survival function S and the exposure curve G respectively, normalised along the x axis, that is, with [0,1] as the co-domain.

The *normalised* exposure curve is defined by the following function $[G](u)$, which gives the percentage of the expected losses retained by the (re)insured if a deductible of value $u \times M$ is imposed.

$$[G](u) := \frac{\int_0^u (1-[F](y))dy}{\int_0^1 (1-[F](y))dy} = \frac{\int_0^u [S](y)dy}{\int_0^1 [S](y)dy}. \tag{2.1}$$

The exposure curve $[G](u)$ is in 1:1 correspondence with $[F](y)$. In this sense it can be considered a re-engineered version of $[F](y)$, as mentioned in Section 1. Equation (2.1) shows how to obtain $[G]$ from $[F]$. Equation (2.2) shows how to obtain $[F]$ from $[G]$:

$$[F](y) = \begin{cases} 1 - \frac{[G]'(y)}{[G]'(0)} & 0 \le y < 1 \\ 1 & y = 1 \end{cases}. \tag{2.2}$$

Some important consequences of the definition of the exposure curve given in Equation (2.1) are that: (a) $[G](0) = 0$, (b) $[G](1) = 1$, (c) $[G](u)$ is concave, continuous and increasing. These are properties that need to be considered by whoever wants to create their own exposure curve from scratch rather than by calculating the integral of the survival distribution $[S](y)$. Ignoring this may easily lead to non-sensical results: for example, an occasionally convex $[G](u)$ leads to a decreasing $[F](y)$ and therefore to a negative probability density.

The following relationship between the derivative of G and the expected damage ratio will be useful in Sections 3 and 4:

$$[G]'(0) = \frac{1}{E([X])}. \tag{2.3}$$

2.1.1. *Maxwell-Boltzmann-Bose-Einstein-Fermi-Dirac (MBBEFD) curves*
The MBBEFD class of distributions – a popular model for normalised exposure and severity curves – is defined as follows:

$$[G_{b,g}](u) = \frac{\ln\left(\frac{(g-1)b + (1-bg)b^u}{1-b}\right)}{\ln(bg)}, \tag{2.4}$$

$$[F_{b,g}](y) = \begin{cases} \frac{b(g-1)(1-b^y)}{b(g-1) + (1-bg)b^y} & \text{if } 0 \le y < 1 \\ 1 & \text{if } y = 1 \end{cases}. \tag{2.5}$$

The corresponding probability density is given by:

$$[f_{b,g}](y) = \frac{b(g-1)(1-b^y)}{b(g-1) + (1-bg)b^y} \times \mathbb{I}_{[0,1)}(y) + \frac{1}{g} \times \delta(y-1). \tag{2.6}$$

Where $\mathbb{I}_{[0,1)}(y)$ is the step function, which is 1 whenever $0 \le y < 1$ and 0 elsewhere, and $\delta(y-1)$ is the so-called Dirac's delta, which is 0 for $y \ne 1$, it is undefined for $y = 1$ and whose integral over the real number line is 1. This is the way of expressing mathematically that $[f_{b,g}](y)$ is a probability density defined as a normal function between 0 and 1 and with a probability mass of $\frac{1}{g}$ concentrated at $y = 1$.

(We ignore for simplicity special cases such as $b = 1$ or $bg = 1$, which are dealt with exhaustively in Bernegger (1997) and Riegel (2010).)

The Swiss Re "c" curves are a very widely used special case of the MBBEFD class with a single parameter c, with b and g calculated as

$$b = e^{\alpha + \beta c(1+c)}, \quad g = e^{(\gamma + \delta c)c}. \tag{2.7}$$

The values of the four parameters $\alpha, \beta, \gamma, \delta$ were set by Bernegger (1997) to approximate the empirical exposure curves used by Swiss Re for the various categories. The values published are: $\alpha = 3.1, \beta = -0.15, \gamma = 0.78, \delta = 0.12$.

In case we use the c parameterisation, we will denote the functions in Equations (2.4), (2.4) and (2.6) as $[G_c]$, $[F_c]$ and $[f_c]$.

2.1.2. *Calibration of MBBEFD curves*

A popular method of calibrating an MBBEFD curve given a set of damage ratios $y_1, y_2 \dots y_N$ is to maximise the following log likelihood function in the presence of censored data:

$$LL = \ln\left(\prod_{j=1}^{N'} [f_{b,g}](y_j) \times \left(\frac{1}{g}\right)^{N-N'}\right) = \sum_{j=1}^{N'} \ln\,[f_{b,g}](y_j) - (N - N') \times \ln g, \quad (2.8)$$

where we have assumed (without loss of generality) that the first N' data points are less than 1 and the remaining $N - N'$ data points are equal to 1 (total losses). Equation (2.8) can be usefully rewritten in a more compact format and in a way in which the order of the data points does not matter as follows:

$$LL = \sum_{j=1}^{N} \left(\ln\left[f_{b,g}\right](y_j) \times \mathbb{I}_{[0,1)}(y_j) - \ln g \times \mathbb{I}_1\left(y_j\right)\right), \quad (2.9)$$

where $\mathbb{I}_1\left(y_j\right)$ is equal to 1 if $y_j = 1$, 0 otherwise.

The maximisation of LL can be performed with respect to b, g in the general case and with respect to c in the case of Swiss Re c curves. The problem does not have a closed form solution and optimisation techniques need to be used.

Equation (2.9) will be the basis of the MLE method for the calibration of the generalised exposure rating model in the general case in Section 3.3.

2.2. Non-normalised curves

In this paper, it will often be more convenient to refer to the non-normalised version of the severity, survival and exposure curves, that is, the version in which the codomain is given in a monetary amount rather than as a fraction of the MPL or IV. These will be referred to as $F(M; x)$, $S(M; x)$, $G(M; d)$ or simply as $F(x)$, $S(x)$, $G(d)$.

The reason for this unusual choice is that there are different "maximum" monetary amounts at play in this paper: the maximum amount for scale-independent losses, the estimated MPL and the IV (which is often higher than the MPL), and at different points in our treatment we would have to normalise with respect to these different monetary amounts, with the potential (or rather, the certainty) of creating confusion.

The relationship between the normalised and the non-normalised curves is straightforward:

$$\begin{aligned} F(M;x) &= [F](x/M) \\ S(M;x) &= 1 - F(M;x) = [S](x/M) \\ G(M;d) &= [G](d/M) \\ f(M;x) &= \frac{[f](x/M)}{M}. \end{aligned} \tag{2.10}$$

We can also define the exposure curve $G(M;d)$ directly from $F(M;x)$ or $S(M;x)$ rather than from $[G](d/M)$, and write:

$$G(M;d) = \frac{\int_0^d (1 - F(M;x))dx}{\int_0^M (1 - F(M;x))dx} = \frac{\int_0^d S(M;x)dx}{\int_0^M S(M;x)dx}. \tag{2.11}$$

The relationship between F and G is the same as that between [F] and [G]:

$$F(M;x) = \begin{cases} 1 - \frac{G'(M;x)}{G'(M;0)} & 0 \le x < M \\ 1 & x = M \end{cases}. \tag{2.12}$$

Equation (2.3) can be rewritten as:

$$G'(M;0) = \frac{1}{E(X)} = \frac{[G]'(0)}{M}. \tag{2.13}$$

2.3. The traditional approach to D&F property exposure rating

In treaty reinsurance, the ground-up expected losses to a particular location can be estimated based on the original premium and the expected loss ratio (ExpLoss = OrigPremium × E(LR)). These losses can then be allocated to a given layer of reinsurance L xs D by using the relevant exposure curve: $EL(L \text{ xs } D) = \text{ExpLoss} \times (G((L+D)/M) - G(D/M))$, where M is the MPL. The question of the underlying (local) deductibles, that is, the small deductible that are retained by the reinsured's clients, can normally be safely ignored. There are of course many other subtleties in treaty exposure rating, but they are not relevant to our discussion.

In direct insurance (and potentially in facultative reinsurance), things work slightly differently. In the following, we will look at the main ingredients and recipes typically used for pricing. We will limit ourselves to the pricing of what are traditionally referred to as FLEXA losses (fire/lightning/explosion/aircraft collision), ignoring the treatment of natural catastrophe losses (normally not addressed by exposure rating) and we will focus on property damage. Business interruption (BI) is conceptually very similar, but it will not be addressed here since the framework developed in this paper is not needed for BI losses (see Section 5.1 for a short discussion). For a more comprehensive overview of all the components, see Parodi (2014b).

Let us assume that the following information is available for each property/location in the property schedule (i.e., the list of properties in the client's portfolio):

- type of occupancy (e.g., refinery, paper mill, supermarket, etc.) and geographical location;
- IV – specified in the contract;
- MPL – typically estimated by the underwriter based on the information that they receive from the client or from risk engineers. For the purpose of this paper, we will assume that MPL is what the name suggests – the largest possible value of the loss under the worst circumstances;
- LD – the amount retained by the insured. It is "local" in the sense that it is in general different for each property.

2.3.1. *Base rates on value and their use for expected losses/premium calculations*
For each type of occupancy (and possibly, territory), the pricing model will typically specify a table of **base ROV**, which give the premium (or the expected losses) per unit of IV assuming a **standard deductible** (SD) (i.e., a reference level of LD), before the application of any layer on insurance (in other terms, for the **full value** of the property). For example, to insure the full value of a widget factory using a SD of \$500,000 a premium base rate of 1‰ of IV might be required.

The base rate can either be specified as a premium base rate or as a loss base rate. If BaseRate_P is a *premium* base rate, the premium to be charged for the full value of the property above the SD is

$$\text{Premium} = \text{BaseRate}_P \times IV. \tag{2.14}$$

Note that the premium is always proportional to the IV and not to the MPL. The implicit assumption is that MPL controls the loss amounts, while the frequency of losses also depends on IV. For example, compare the situation of (1) a plant made of a single building that can burn out completely and that has MPL = IV = \$50 m to that of (2) a plant with two well-separated buildings both identical to that of Case (a). In this case the MPL may still be \$50 m (only one of the buildings may be lost to a fire originating in one of the buildings), but the IV in the second case is \$100 m. The loss amounts will be distributed in the same way in the two cases, but the number of expected losses will be double for Case (b): hence the double premium.

If BaseRate_L is a *loss* base rate, Equation (2.14) is replaced by a similar equation for the expected losses:

$$\text{Expected losses} = \text{BaseRate}_L \times IV, \tag{2.15}$$

and the premium is obtained by adding the loadings for expenses, cost of capital and so on in the usual way.

In the following, we will always assume that the base rate is a loss rate, as this is the most relevant to actuaries and the most general case as it allows to calibrate the loadings contract by contract depending on the expected volatility.

2.3.2. *Impact of changes in the underlying* LD*s*

The pricing model also needs to specify **how the expected losses increase/reduce if the LD is lower/higher than the SD**. In general, what we need is something that can be formally described as a transformation function $\mathbb{T}(SD, LD, MPL)$ that gives the multiplication factor to go from the SD to a specific LD for a property with a given MPL (this transformation function might be different for different occupancies):

$$\text{RateOnValue}_L(LD) = \text{BaseRate}_L \times \mathbb{T}(SD, LD, MPL). \tag{2.16}$$

The transformation function can be anything from a simple table such as:

Local deductible	**Multiplication factor**
50% of standard deductible	×1.5
100% of standard deductible	×1.0
150% of standard deductible	×0.8
200% of standard deductible	×0.6

to a more complex algebraic relation, such as this example in Lee and Frees (2016):

$$\mathbb{T}(SD, LD, MPL) = \left(\frac{\theta + \frac{LD}{MPL}}{\theta + \frac{SD}{MPL}} \right)^{-\alpha+1}. \tag{2.17}$$

The expected losses for a given LD then become:

$$\text{Expected losses} = \text{RateOnValue}_L(LD) \times IV. \tag{2.18}$$

Equation (2.18) gives the expected losses for the full value of the property – that is, up to the IV. In order to determine how much of these expected losses should be allocated to a specific layer of insurance L xs D, the pricing model typically includes **exposure curves** for each type of occupancy.

2.3.3. *Allocation of losses to layers*

Since the curves are normally assumed to be in excess of any LD, and the exposure curve $[G](d)$ is normally defined in terms of the MPL, the expected losses to a layer L xs D under a local deductible equal to LD can in this case be written as:

$$E(S(D, L); LD) = \text{BaseRate}_L \times \mathbb{T}(SD, LD, MPL) \times IV \times \left([G]\left(\frac{D+L-LD}{MPL-LD}\right) - [G]\left(\frac{D-LD}{MPL-LD}\right)\right). \tag{2.19}$$

(In Equation (2.19), we have assumed that the MPL is defined from the ground up and not in excess of the deductible.)

The pricing model will also include rules on how much premium or expected losses should be allocated between MPL and IV, for example, by using a minimum rate on line (i.e., premium divided by limit) for layers of insurance above MPL.

2.3.4. *Loss simulation*

The pricing model may include a simulation engine to calculate not only the expected losses to a given layer but also the full aggregate loss distribution, gross or net of the policy structure. This allows one, in turn, to calculate the expected losses for complex structure with aggregate features and to measure volatility, for example, in the context of calculating risk-based margins.

2.3.4.1 *Contracts with a single property*. Let us start with the simple case of a contract with a single property. The number of losses for a specific scenario will be sampled from a discrete distribution (e.g., Poisson), and the loss amounts will be sampled from the MBBEFD distribution. The expected number of losses in excess of the LD (and therefore, the mean of the frequency model) will be given by:

$$E(N) = \frac{(S; LD)}{E(X)} = \frac{\text{BaseRate}_L \times \mathbb{T}(SD, LD, MPL) \times IV}{(MPL - LD)/[G]'(0)}. \tag{2.20}$$

The loss amounts in excess of the LD can then be sampled from the MBBEFD distribution. This is easy because the normalised severity distribution described by Equation (2.5) is easily invertible, and therefore we can use **inverse transform sampling** (Devroye, 1986), that is, generate a pseudo-random number uniformly between 0 and 1 and then calculate the corresponding normalised loss as $y = [F]^{-1}_{b,g}(u)$. To be more specific, to produce a random loss between 0 and MPL we can use the following simple algorithm:

1. Inputs: parameters b, g of the MBBEFD curve; MPL, LD
2. Generate a pseudo-random number u uniformly between 0 and 1
3. If $u < 1 - 1/g$, set $y = \ln\left(\frac{b(g-1)(1-u)}{b(g-1)+(1-bg)u}\right)/\ln(b)$; else, set $y = 1$
4. Return from-the-ground-up loss: $\text{Loss} = LD + y \times (MPL - LD)$

Algorithm 1. *Generation of random variates $[F]^{-1}_{b,g}(u)$ from a (normalised) MBBEFD distribution, and the related from-the-ground-up losses. Note that this algorithm produces a loss equal to MPL with finite (i.e., non-zero) probability.*

The outputs from the frequency and severity model can then be used to create a large number of different aggregate loss scenarios and calculate all necessary statistics for the aggregate loss distribution on a gross and net (i.e., after the policy structure has been applied) basis.

2.3.4.2 *Contracts with multiple properties.* In most cases, contracts will include multiple properties. If we can assume that the losses from the different properties are independent (a safe assumption for non-catastrophe losses), the loss S for the whole portfolio in a given scenario is the sum of the losses of the individual properties: $S = S_1 + S_2 + \ldots S_K$, and the loss distribution for S is the convolution of the loss distribution of the individual properties.

The annual number of losses for the whole contract can be modelled as a Poisson distribution with mean equal to the sum of the mean for each of the K individual properties:

$$E(N) = \sum_{k=1}^{K} E(N_k) = \sum_{k=1}^{K} \frac{\text{BaseRate}_{L,k} \times \mathbb{T}(SD, LD, MPL_k) \times IV_k}{\frac{MPL_k - LD_k}{\left[G_{b_k, g_k}\right]'(0)}}. \tag{2.21}$$

(We have assumed here that the parameters of the exposure curve $[G](b_k, g_k; u)$ may vary from one property to the other.)

We already know how to simulate the loss amount for a specific property. All that remains to be decided is how to pick for each scenario the properties that are going to have a loss. This can be done proportionally to the expected number of losses for each property, $E(N_k)$.

The algorithm (in its simplest version) for the case where the frequency model is Poisson goes as follows. The algorithm can be adapted to the case where a negative binomial distribution is used.

1. For each scenario, sample the number of losses from a Poisson distribution with rate $\lambda = E(N)$
2. Assign each of the losses to one of the K properties in proportion to $E(N_1), E(N_2) \ldots E(N_k)$, calculated for each property as per Equation (2.20)
3. For each loss: if property k is selected, sample a loss amount from 0 to MPL_k using Algorithm 1.

Algorithm 2. *Generation of random loss variates for a contract. These random variates can then feed into various calculations, for example, expected losses to a layer or VaR/xTVaR.*

This is the basic form of the simulation engine. In practice, plenty of bells and whistles need to be added in order for the model to deal with BI losses and their correlation with the property damage losses, the presence of machinery breakdown losses, natural catastrophe losses and other components as may be necessary.

These considerations, albeit important, are beyond the scope of this paper.

2.3.5. *Building the technical premium*

Finally, the pricing model will include rules for premium breakdown into expected losses, expenses, profit, etc. typically in a "Cost+" framework (Parodi, 2014a, 2014b). Since these rules depend critically on the financials of the company (including such information as the target return on capital) they may change regularly.

3. Model description

3.1. Scale-dependent and scale-independent losses

The proposed model is based on the assumption that the losses are a mixture of scale-independent losses (following the distribution $F_A(x)$) and scale-dependent losses (following the distribution $F_L(x)$). "A" stands for "attritional" while "L" stands for "large". This is to maintain notational consistence with Section 4, an important special approximation where scale-independent losses will be identified with attritional losses, and scale-dependent losses will be identified as large losses; at this stage, however, it is just an arbitrary notational choice.

In both cases, there is a maximum, and the maximum for scale-independent losses (M_A) will not depend on the size of the property and will typically (but not necessarily) be lower than the maximum of the scale-dependent losses (M_L). In the following, we will assume that this is always the case and that therefore the MPL is also equal to M_L: $M_A \leq M_L = \text{MPL}$.

Unless otherwise specified, all the losses are assumed to be from the ground up.

The CDF of the mixture distribution is

$$\begin{aligned} F(x) = \Pr(X \leq x) &= \Pr(A)\Pr(X \leq x \,|\, A) \\ &+ \Pr(L)\Pr(X \leq x \,|\, L) = \Pr(A)F_A(x) + \Pr(L)F_L(x). \end{aligned} \tag{3.1}$$

If λ_A and λ_L are the expected number of claims for scale-dependent and scale-independent losses respectively, $\Pr(A)$ and $\Pr(L)$ can be calculated as:

$$\Pr(A) = \frac{\lambda_A}{\lambda_L + \lambda_A}, \quad \Pr(L) = \frac{\lambda_L}{\lambda_L + \lambda_A}. \tag{3.2}$$

Therefore, the CDF can be rewritten as

$$F(x) = \frac{\lambda_A}{\lambda_L + \lambda_A} F_A(x) + \frac{\lambda_L}{\lambda_L + \lambda_A} F_L(x). \tag{3.3}$$

We can now calculate the exposure curve $G(d)$ for the mixture, using the formula

$$G(d) = \frac{\int_0^d (1 - F(x))dx}{\int_0^{\mathrm{MPL}} (1 - F(x))dx} = \frac{\int_0^d S(x)dx}{\int_0^{\mathrm{MPL}} S(x)dx}. \tag{3.4}$$

The result is as follows:

$$G(d) = w_A G_A(d) + (1 - w_A)G_L(d), \tag{3.5}$$

where d is the deductible, $G_A(d)$ is the exposure curve (expressed in monetary amount, not normalised) related to $F_A(x)$, $G_L(d)$ is the exposure curve related to $F_L(x)$ and

$$w_A = \frac{\lambda_A E(X_A)}{\lambda_A E(X_A) + \lambda_L E(X_L)}. \tag{3.6}$$

In the expression above, $E(X_L) = \int_0^{\mathrm{MPL}} S_L(x)dx$ and $E(X_A) = \int_0^{M_A} S_A(x)dx$ are the expected loss amounts for scale-dependent and scale-independent amounts respectively, given that a loss happened. The expected loss amounts $E(X_L)$ and $E(X_A)$ can be more usefully written in terms of the damage ratio $[X_L] = X_L/MPL$ and $[X_A] = X_A/M_A$: $E(X_L) = MPL \times E([X_L])$, $E(X_A) = M_A \times E([X_A])$. This is useful because $E([X])$, the expected damage ratio, is determined only by the shape of the exposure curve and not by the scale, that is, the maximum value. Based on this, we can rewrite w_A as:

$$w_A = \frac{1}{\rho_\lambda \rho_{DR} \rho_M + 1}, \tag{3.7}$$

where

- $\rho_\lambda = \lambda_L/\lambda_A$ is the ratio between the expected number of the scale-dependent and scale-independent losses. Typically, we expect this ratio to be much smaller than 1.
- $\rho_{DR} = E([X_L])/E([X_A])$ is the ratio between the expected damage ratio for scale-dependent and scale-independent losses.
- $\rho_M = MPL/M_A$ is the ratio between the maximum value of the scale-dependent losses and that of the scale-independent losses. This ratio is assumed to be ≥ 1.

The weight to be given to the two exposure curves will therefore vary depending on the relative values of ρ_λ, ρ_{DR} and ρ_M.

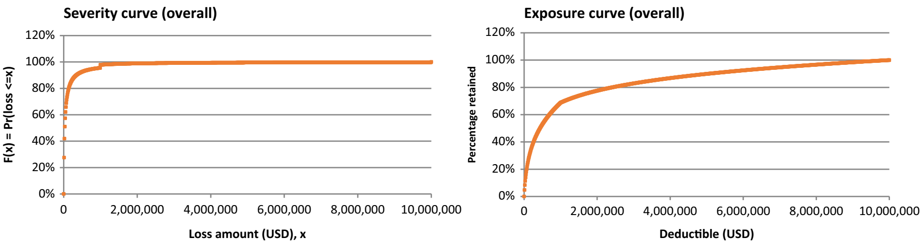

FIGURE 1: The severity and exposure curves for a blend of scale-dependent and scale-independent losses. The scale-dependent and scale-independent curves are modelled as MBBEFD with $c = 3.0$ and $c = 3.8$ respectively, with 70% proportion of attritional losses. The maximum attritional loss is \$1 m, while $MPL = IV =$ \$10 m. Note the step change at the maximum attritional loss, which causes a finite probability for a loss equal to that value – this is an unintended side effect of using a MBBEFD curve, which has little practical consequences but can be avoided by using a different cumulative distribution function such as the Beta distribution that does not have a mass probability point at the maximum value.

The important thing to notice here is that ρ_λ and ρ_{DR} depend on the general type of properties to be rated and are therefore portfolio-level parameters, whilst ρ_M is a property-specific ratio, because it depends on the MPL.

Since $0 \le w_A \le 1$, the overall exposure curve $G(d)$ is a convex combination of the exposure curves for the scale-dependent and scale-independent losses, suitably extended beyond their maximum value as needed. Figure 1 shows how the combined severity and exposure curves look like in practice.

3.1.1. *Limit behaviour*

It is instructive to analyse what happens under limit conditions.

- When $MPL/M_A \to \infty$ the overall exposure curve reduces to the large-loss exposure curve, and this is indeed the expected behaviour. This is indeed the case, on account of the fact that $\lim_{MPL/M_A \to \infty} \rho_M = \infty$ by definition and therefore $\lim_{MPL/M_A \to \infty} w_A = 0$. Consequently, $\lim_{MPL/M_A \to \infty} G(d) = G_L(d)$. Note that this does not mean that the underlying CDF will also converge to the large-loss CDF: $\lim_{MPL/M_A \to \infty} F(d) \neq F_L(d)$. This is not a contradiction and comes from the presence of a large number of very small losses that obviously affects the CDF but has little quantitative impact on the expected total losses (the effect is similar to that described in Appendix B of Parodi and Watson, 2019).
- When the large/attritional – loss frequency ratio ($\rho_\lambda = \lambda_L/\lambda_A$) goes to zero, the exposure curve converges to the attritional exposure curve, as expected: $\lim_{\lambda_L/\lambda_A \to \infty} G_{A/L}(d) = G_A(d)$. When on the other hand, the ratio goes to infinity, the exposure curve converges (again, as expected) to the large-loss exposure curve: $\lim_{\lambda_L/\lambda_A \to \infty} G_{A/L}(d) = G_L(d)$.
- In Section 4.3.2 we will show another discontinuity regarding the limit of $G_{A/L}(d)$ when $MPL \to M_A^+$, which is shown not to be equal to $G_A(d)$. This is

an interesting effect for the attritional/large model introduced in Section 4.1, where there are no large losses below M_A, but is a trivial fact for the general cases.

3.2. Calibration

In the special case where scale-dependent and scale-independent losses can be identified as such, calibration can proceed in the usual way for an MBBEFD curve (see Section 2.1). In most cases, however, it will not be possible or practical to identify which losses (especially smaller losses) are scale-dependent and which are scale-independent.

A general way of calibrating the model under these circumstances is by clustering techniques such as expectation-maximisation, which aims to maximise the likelihood of observed data in the presence of latent variables. In our case, the latent variable is the category (scale-independent vs. scale-dependent) that the data point belongs to.

Expectation-maximisation is a widespread statistical technique, which was proposed independently several times and most famously by Dempster *et al.* (1977). Our treatment follows that by Bishop (2006), to which the reader is referred for a general presentation of the method: here we will simply describe a possible application to our problem.

To show how the method works through an example, we make the following *assumptions*.

A.1 We have N pairs (x_j, M_j) where x_j is loss in monetary amount and M_j is the MPL of the affected property.

A.2 We assume that – as specified in Section 3.1 – losses can come from either a scale-independent distribution $F_A(M_A; x)$ whose maximum value is M_A or from a scale-dependent distribution $F_L(M_j; x)$ whose maximum value is M_j. We assume that the prior probability that the loss comes from the scale-independent distribution is $p_A = \Pr(A)$ (see Equation (3.1)), which is related to ρ_λ by the relationship $p_A = 1/(\rho_\lambda + 1)$.

A.3 For simplicity we will assume that $M_A < M_j$ for all j, which reflects the fact that $F_A(M_A; x)$ typically represents smaller losses and is an assumption that can be relaxed quite easily if needed.

A.4 Also for simplicity of illustration (and to keep the model simple so that calibration is feasible) we also assume that both $F_A(M_A; x)$ and $F_L(M_j; x)$ can be parameterised as Swiss Re c curves with $c = c_A$ and $c = c_L$ respectively.

The parameters we wish to determine are, in the most general case: c_A, c_L, M_A and p_A. Expectation-maximisation works according to the following iteration scheme.

Step 1 – Initialisation
Set initial values $c_A^{(\mathrm{old})}, c_L^{(\mathrm{old})}, M_A^{(\mathrm{old})}, p_A^{(\mathrm{old})}$ in an appropriate range (e.g., $c_A^{(\mathrm{old})}, c_L^{(\mathrm{old})}$ between 0 and 10, $M_A^{(\mathrm{old})} < \min_{j=1,\ldots N} M_j$, $0 < p_A^{(\mathrm{old})} < 1$).

Step 2 – Expectation
Calculate for each data point x_j the so-called "responsibility" $\gamma_{A,j}^{(\mathrm{old})}$ that $F_A(M_A; x)$ takes for x_j, that is, the posterior probability that x_j comes from $F_A(M_A; x)$, based on the current estimation of the parameters: $c_A^{(\mathrm{old})}, c_L^{(\mathrm{old})}$, $M_A^{(\mathrm{old})}, p_A^{(\mathrm{old})}$.

$$\begin{aligned}\gamma_{A,j}^{(\mathrm{old})} &= \Pr\left(A \,\middle|\, c_A^{(\mathrm{old})}, c_L^{(\mathrm{old})}, M_A^{(\mathrm{old})}, p_A^{(\mathrm{old})}; x_j\right) \\ &= \begin{cases} \dfrac{p_A^{(\mathrm{old})} f_{c_A^{(\mathrm{old})}}\left(M_A^{(\mathrm{old})}; x_j\right)}{p_A^{(\mathrm{old})} f_{c_A^{(\mathrm{old})}}\left(M_A^{(\mathrm{old})}; x_j\right) + \left(1-p_A^{(\mathrm{old})}\right) f_{c_L^{(\mathrm{old})}}\left(M_j; x_j\right)} & 0 \le x_j < M_A^{(\mathrm{old})} \\ 1 & x_j = M_A^{(\mathrm{old})} \\ 0 & M_A^{(\mathrm{old})} < x_j \le M_j \end{cases},\end{aligned} \tag{3.8}$$

where $f_{c_A^{(\mathrm{old})}}(M_A^{(\mathrm{old})}; x_j)$ and $f_{c_L^{(\mathrm{old})}}(M_j; x_j)$ are the density functions for $F_A(M_A; x)$ and $F_L(M_j; x)$ respectively, with the current estimation of the parameters. As we know from Section 2.2, these functions can also be written in terms of the normalised functions by setting $f_{c_A^{(\mathrm{old})}}(M_A; x_j) = [f_{c_A^{(\mathrm{old})}}](x_j/M_A)/M_A$ and $f_{c_L^{(\mathrm{old})}}(M_j; x_j) = [f_{c_L^{(\mathrm{old})}}](x_j/M_j)/M_j$.

Step 3 – Maximisation
Find the updated parameters c_A, c_L, M_A, p_A that maximise the expectation of the complete-data log likelihood calculated for generic values of c_A, c_L, M_A, p_A, that is, the following function:

$$\begin{aligned}LL(c_A, M_A, c_L, p_A) = \sum_{j=1}^{N} &\gamma_{A,j}^{(\mathrm{old})} \ln \Pr\left(x_j, A | c_A, c_L, M_A, p_A\right) \\ &+ \left(1 - \gamma_{A,j}^{(\mathrm{old})}\right) \ln \Pr\left(x_j, L | c_A, c_L, M_A, p_A\right).\end{aligned} \tag{3.9}$$

The interpretation of Equation (3.9) is that $LL(c_A, M_A, c_L, p_A)$ represents the weighted average of the *complete-data* log likelihood for the case where the data point comes from $F_A(M_A; x)$ and that where it comes from $F_L(M_j; x)$, given the value of the parameters c_A, c_L, M_A, p_A. The weights $\gamma_{A,j}^{(\mathrm{old})}$ and $1 - \gamma_{A,j}^{(\mathrm{old})}$ are given by the *posterior* probability of that data point coming from $F_A(M_A; x)$ and $F_L(M_j; x)$ respectively. We speak of "complete-data log likelihood" because $\Pr(x_j, A | c_A, c_L, M_A, p_A)$ (resp. $\Pr(x_j, L | c_A, c_L, M_A, p_A)$) is the joint probability that the loss amount is x_j *and* the loss comes from distribution $F_A(M_A; x)$ (resp. $F_L(M_j; x)$).

The parameters are then updated to the value that maximises $LL(c_A, M_A, c_L, p_A)$:

$$\left(c_A^{(\text{new})}, M_A^{(\text{new})}, c_L^{(\text{new})}, p_A^{(\text{new})}\right) = \underset{c_A, M_A, c_L, p_A}{\text{argmax}} \; LL(c_A, M_A, c_L, p_A)\,. \tag{3.10}$$

In our specific case (using Assumption A.4), Equation (3.9) can be expanded by writing $\Pr(x_j, A|c_A, c_L, M_A, p_A)$ and $\Pr(x_j, L|c_A, c_L, M_A, p_A)$ as:

$$\Pr\left(x_j, A|c_A, c_L, M_A, p_A\right) = \begin{cases} p_A \times \frac{[f_{c_A}](x_j/M_A)}{M_A} & \text{if } x_j < M_A \\ p_A \times \frac{1}{g_A} & \text{if } x_j = M_A \end{cases}, \tag{3.11}$$

$$\Pr\left(x_j, |c_A, c_L, M_A, p_A\right) = \begin{cases} (1-p_A) \times \frac{[f_{c_L}](x_j/M_j)}{M_j} & \text{if } x_j < M_j \\ (1-p_A) \times \frac{1}{g_L} & \text{if } x_j =_j \end{cases}. \tag{3.12}$$

Considering Equation (2.9), we can now rewrite Equation (3.9) as follows:

$$\begin{aligned} LL(c_A, M_A, c_L, p_A) = \sum_{j=1}^{N} \gamma_{A,j}^{(\text{old})} & (\ln p_A + (\ln [f_{c_A}](x_j/M_A) - \ln M_A) \\ & \times \mathbb{I}_{[0,M_A)}(x_j) - \ln g_A \times \mathbb{I}_{M_A}(x_j)) + (1-\gamma_{A,j}^{(\text{old})})(\ln(1-p_A) \\ & + (\ln [f_{c_L}](x_j/M_j) - \ln M_j) \times \mathbb{I}_{[0,M_j)}(x_j) - \ln g_L \times \mathbb{I}_{M_j}(x_j)) \end{aligned} \tag{3.13}$$

Note that the parameters g_A and g_L which appear in Equations (3.20)–(3.22) can be calculated from c_A, c_L through Equation (2.7).

The first thing to notice is that $LL(c_A, M_A, c_L, p_A)$ can be split into a part that depends on p_A only and another part that depends on c_A, M_A and c_L only:

$$LL(c_A, M_A, c_L, p_A) = LL(p_A) + LL(c_A, M_A, c_L)\,, \tag{3.14}$$

where:

$$LL(p_A) = \sum_{j=1}^{N} \gamma_{A,j}^{(\text{old})} \ln p_A + \left(1-\gamma_{A,j}^{(\text{old})}\right) \ln(1-p_A), \tag{3.15}$$

and:

$$\begin{aligned} LL(c_A, M_A, c_L) = \sum_{j=1}^{N} \Big(& \gamma_{A,j}^{(\text{old})} \left((\ln [f_{c_A}](x_j/M_A) - \ln M_A \right) \times \mathbb{I}_{[0,M_A)}\left(x_j\right) - \ln g_A \\ & \times \mathbb{I}_{M_A}\left(x_j\right)\Big) + \left(1-\gamma_{A,j}^{(\text{old})}\right)\Big(\left(\ln [f_{c_L}](x_j/M_j) - \ln M_j\right) \\ & \times \mathbb{I}_{[0,M_j)}\left(x_j\right) - \ln g_L \times \mathbb{I}_{M_j}\left(x_j\right)\Big)\Big)\,. \end{aligned} \tag{3.16}$$

Therefore the optimal choice for p_A is independent of the optimal choice for c_A, M_A and c_L:

$$p_A^{(\text{new})} = \underset{p_A}{\text{argmax}}\, LL(p_A)$$
$$\left(c_A^{(\text{new})}, M_A^{(\text{new})}, c_L^{(\text{new})}\right) = \underset{c_A, M_A, c_L}{\text{argmax}}\ LL(c_A, M_A, c_L)\,. \tag{3.17}$$

There is no closed formula for $\text{argmax}_{c_A,M_A,c_L}\ LL(c_A, M_A, c_L)$, so the values for c_A, M_A and c_L that maximise the likelihood at every step need to be found numerically. As for $p_A^{(\text{new})}$, this can be found simply by setting the partial derivative of $LL(c_A, M_A, c_L, p_A)$ to zero:

$$\frac{\partial LL(c_A, M_A, c_L, p_A)}{\partial p_A} = \frac{dLL(p_A)}{dp_A} = 0 \ \Rightarrow\ p_A^{(\text{new})} = \frac{\sum_{j=1}^N \gamma_{i,j}^A}{N}. \tag{3.18}$$

Note that $\frac{d^2 LL(p_A)}{dp_A^2} < 0$, so this is actually a maximum.

Step 4 – Iterate

Set $c_A^{(\text{old})} \leftarrow c_A^{(\text{new})}$, $M_A^{(\text{old})} \leftarrow M_A^{(\text{new})}$, $c_L^{(\text{old})} \leftarrow c_L^{(\text{new})}$, and $p_A^{(\text{old})} \leftarrow p_A^{(\text{new})}$. Go back to Step 2 and repeat until the improvement in the log-likelihood is below a desired threshold.

3.2.1. *Practical considerations*

The algorithm described above is a formal solution to the problem of calibrating a generalised exposure rating model, and can be easily verified to work if one uses artificial data from a known mixture and tries to estimate the parameters. This experiment also shows, however, that the landscape of solutions is rather flat and that getting trapped in local minima is easy if the initial choice of parameters is not reasonable (the EM method is not guaranteed to converge to a global maximum). With real-world data for which the double-MBBEFD curve is only an approximation, this exercise becomes even more questionable.

In practice, therefore, one is forced to adopt some heuristics, and incorporate as much actuarial/underwriter judgment as possible. Examples of such heuristics (which can also be used in combination) are:

- Estimate c_L separately by using MLE on all data points above M^* (using the conditional density) where M^* is an upper bound for M_A according to judgment. Then the problem of maximising $LL(c_A, M_A, c_L)$ at each step becomes a two-dimensional problem, which is more manageable.
- Restrict the range for c_A, M_A and p_A as much as possible based on judgment and prior information.

- If the number of scale-independent losses is much larger than the number of scale-dependent losses (as it is normally the case), it is possible to get an initial value of c_A for the optimisation procedure by assuming a value of M_A in the middle of the acceptable range and then assume that *all* data points below M_A are scale-independent losses and using MLE. The actual value of c_A found in this way depends of course on the value M_A, but this procedure proves to be quite robust for different choices of our guess for M_A.

In Section 4.4 we will see an even more radical way of avoiding the identifiability issue and the expectation-maximisation iteration by assuming a priori that all scale-independent losses are below M_A and all scale-dependent losses are above M_A (a broad approximation).

It should also be noted that in this section we have made the strong assumption that for every data point the information on both the loss amount x_j and the relevant MPL, M_j, was available. This is often not the case, especially for the smaller losses. The approach described in Section 4.4.2 will relax this assumption as well.

Finally, we should note that the simple model that we have explored in this section for calibration purposes – the "double Swiss Re c curve" model – has one obvious limitation: the scale-independent distribution $F_{c_A}(M_A; x)$ has a probability mass point at $x = M_A$ – in other terms, a finite proportion $1/g_A$ of losses exactly M_A. Whilst a finite probability for a total loss makes perfect sense for scale-dependent losses, there is no reason why there should be a finite proportion of losses that are exactly $= M_A$. This could be easily taken care of by using a rescaled version of the MBBEFD distribution such that $\lim_{x \to M_A^-} F_A(M_A; x) = 1$ but the difference in practice is not large (at least if the value of c is large enough), and priority has been given to simplifying the mathematical treatment.

3.3. Treatment of MPL uncertainty

Although MPL is theoretically (and according to the definition we have given in this paper) the largest possible loss value, in practice the MPL is an estimated figure and as such breaches are occasionally possible.

The cleanest way of taking this into account is arguably that of allocating a fixed probability p (say, 1%) of exceeding MPL, and then assuming a simple conditional distribution for that region which then translates into the correct behaviour for $G(d)$. Alternatively, we could change $G(x)$ directly. However, this is less intuitive and requires checking that $G(x)$ still has all the necessary properties (including concavity).

In the following, we are assuming that $MPL \leq IV$. In practice, this is not always the case (when, for example, the MPL for a property allows for the possibility of losing neighbouring properties), but we can safely ignore this case, as in that case we can assume that the severity curve goes all the way up to MPL, and IV is used for the calculation of the overall expected losses and the expected number of losses.

In the following, let $F_0(x)$ and $G_0(d)$ be the severity and exposure curves for losses up to MPL, and $F_1(x)$ and $G_1(d)$ the severity and exposure curves between MPL and IV.

In general, we can write the overall severity distribution as:

$$F_{\text{overall}}(IV;x) = \begin{cases} (1-p) \times F_0(MPL;x) & \text{if } 0 \le x < MPL \\ 1-p+p \times F_1(IV-MPL;x-MPL) & \text{if } MPL \le x \le IV \end{cases}, \tag{3.19}$$

which can also be written in terms of the survival function as:

$$S_{\text{overall}}(IV;x) = \begin{cases} p+(1-p) \times S_0(MPL;x) & \text{if } 0 \le x < MPL \\ p \times S_1(IV-MPL;x-MPL) & \text{if } MPL \le x \le IV \end{cases}. \tag{3.20}$$

We can then calculate G(x) in the standard way. If $G_0(MPL;d)$ (resp. $G_1(IV-MPL;d)$) is the exposure curve corresponding to severity curve $F_0(MPL;x)$ (resp. $F_1(IV-MPL;x)$), a little algebra will show that:

$$G_{\text{overall}}(IV;d) = \begin{cases} k\left(pd + \frac{(1-p)G_0(MPL;d)}{G_0'(MPL;0)}\right) & \text{if } 0 \le u \le MPL \\ k\left(pMPL + \frac{1-p}{G_0'(MPL;0)} + \frac{p\ G_1(IV-MPL;d-MPL)}{G_1'(IV-MPL;0)}\right) & \text{if } MPL < u \le IV \end{cases}, \tag{3.21}$$

where

$$k = \frac{1}{pMPL + \frac{1-p}{G_0'(MPL;0)} + \frac{p}{G_1'(IV-MPL;0)}}. \tag{3.22}$$

A simple choice for $F_1(x)$ is that of a linear CDF, that is, $F_1(IV-MPL;x) = (x-MPL)/(IV-MPL)$, corresponding to a uniform probability density (any loss above MPL is equally likely) above x and an exposure curve above MPL of the form $G_1(IV-MPL;d) = 1-(1-\frac{d-MPL}{IV-MPL})^2$ where d ranges from MPL to IV. Slightly more sophisticated probability density functions such as decreasing exponential may be used to ensure that it is more likely for the actual MPL to be near to the estimated value. Of course we could also assume that the actual MPL actually falls below the estimated MPL, and use, say, a Gaussian distribution centred around the estimated MPL, but the limited value of doing this does not justify the added nastiness of the relevant mathematics.

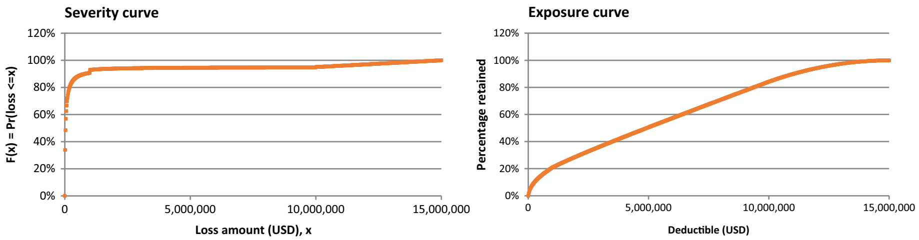

FIGURE 2: Illustration of the extension of a severity curve (left) and the corresponding exposure curve (right) beyond MPL. As in Figure 1, The scale-dependent and scale-independent curves are modelled as MBBEFD with $c = 3.0$ and $c = 3.8$ respectively, and the maximum attritional loss is \$1 m, with $MPL = \$10$ m. However, the curve has been extended to $IV = \$15$ m. A uniform loss distribution between MPL and IV is assumed (see text). The probability of exceeding MPL is set to 5% (an exaggerated value to emphasise the effect of the MPL uncertainty).

Notice that it is better to extend the CDF beyond MPL and calculate the relevant exposure curve rather than try to extend the exposure curve itself, as in the latter case we would need to add constraints to ensure that the concavity property for the overall exposure curve still holds after the extension.

Figure 2 below shows how this extension looks in practice. For consistency, we have chosen G_0 to be the same as the mix of MBBEFD curves shown in Figure 1, but assuming $MPL = \$10$ m, $IV = \$15$ m rather than $MPL = IV = \$10$ m.

3.3.1. *Determining the MPL exceedance probability*

If sufficient historical data is available, the calibration of the exceedance probability p (as defined in Section 3.2) is straightforward: p can be estimated as $\hat{p} =$(#losses breaching MPL)/(#losses). Corrections may be needed to take account of truncating and censoring of losses, and one may focus on losses above the attritional threshold, so the actual formula will vary.

As for the distribution of the amount of the MPL exceedance, this is of course more challenging but in the presence of sufficient data it will still be possible to form an idea of the general behaviour of the curve, that is, if it is closer to a uniform distribution, an exponentially decreasing distribution or some in-between behaviour.

Where historical data are insufficient, the probability p needs to be assessed using underwriting judgment. In this case, it would be difficult for underwriters to estimate the probability p that the MPL is exceeded directly: instead, underwriting management might prefer to set pragmatically a percentage of the expected losses to cater for the possibility of losses above MPL, $\%EL(>MPL) = EL(>MPL)/EL(\text{all})$. This information can then be translated into a corresponding value of exceedance probability. Note first of all that:

$$\%EL(> MPL) = 1 - G_{\text{overall}}(IV; MPL)\ , \tag{3.23}$$

where $(IV; PL)$ is given by

$$G_{\text{overall}}(IV; MPL) = \frac{pMPL + \frac{1-p}{G_0'(MPL; 0)}}{pMPL + \frac{1-p}{G_0'(MPL; 0)} + \frac{p}{G_1'(IV-ML; 0)}}, \tag{3.24}$$

which is of the simple form

$$G_{\text{overall}}(IV; MPL) = \frac{Ap + B}{Cp + B}, \tag{3.25}$$

where:

- $A = MPL - 1/G_0'(MPL; 0)$
- $B = 1/G_0'(MPL; 0)$
- $C = MPL - 1/G_0'(MPL; 0) + 1/G_1'(IV - MPL; 0)$

In the case where $F(x)$ is linear, $G_1'(IV - MPL; 0) = 2/(IV - MPL)$ and the expression for C in the formula above slightly simplifies to $C = \frac{IV+MPL}{2} - \frac{1}{G_0'(MPL; 0)} = \frac{IV+MPL}{2} - \frac{MPL}{[G_0]'(0)}$.

It is therefore straightforward to solve this with respect to p, yielding:

$$p = \frac{B \times (1 - G_{\text{overall}}(IV; MPL)\,)}{C \times G(IV; MPL) - A} = \frac{B\ \times\ \%EL(> MPL)}{C(1 - \%EL(> MPL)\,) - A}, \tag{3.26}$$

and use this p for our exposure rating calculations. In a slightly more sophisticated way, we can avoid the discontinuity at $x = MPL$ (which is rather artificial given that the total loss is not really a total loss and the MPL is subject to some uncertainty) and have the "tail" piece start at $F(MPL^-)$. We leave the calculation details for this case to the reader.

In practice, to avoid that a significant percentage of expected losses should be allocated to a relatively small amount if MPL is close to IV, we can set an upper bound for $\%EL(> MPL)$ and reduce it proportionally as MPL approaches IV:

$$\%EL(> MPL) = \begin{cases} 0 & \text{if } MPL = IV \\ \%EL_{\max} \times \left(1 - \frac{MPL}{IV}\right) & \text{if } MPL < IV \end{cases}. \tag{3.27}$$

For extra smoothness in the transition between the case $MPL < IV$ and the case $MPL = IV$, we can use $\left(1 - \frac{MPL}{IV}\right)^2$ instead of $1 - \frac{MPL}{IV}$.

4. A special case of practical relevance: The attritional/large model

We now look at a special case of the model described in Section 3 that is particularly relevant from a practical standpoint. Given the difficulty of separating the scale-independent and the scale-dependent losses mentioned in Section 3, and given that the scale-independent losses *tend* to be attritional while scale-dependent losses *tend* to be large, it is tempting to split the two by a rather crude approximation, assuming that all losses below an agreed threshold M_A are scale-independent and all losses above that threshold are scale-dependent (regardless of the property's MPL, which is assumed to be $\geq M_A$).

One advantage of this method is that the monetary amount by which the expected losses for a specific property are changed when the deductible changes depends only on the attritional loss curve, as long as the deductible remains below M_A. This is something that corresponds to the intuition of property underwriters.

Also, it should be noted that this model is not too dissimilar from the standard approach in which the exposure curve is assumed to be in excess of the underlying deductible, the main difference being that the exposure curve is in this case assumed to be above the same threshold M_A for all properties.

Throughout this section all the losses are assumed to be from the ground up, unless otherwise specified.

4.1. Description of the attritional/large model

The overall attritional/large CDF is a mixture of two components denoted as $F_A(x)$ and $F_L(x)$ as in Equation (3.3), which we repeat here with minor notation changes for convenience:

$$F_{A/L}(x) = \frac{\lambda_A}{\lambda_L + \lambda_A} F_A(x) + \frac{\lambda_L}{\lambda_L + \lambda_A} F_L(x). \tag{4.1}$$

In this special case, the attritional component can be written as:

$$F_A(x) = \begin{cases} F_A^*(M_A; x) & 0 \leq x \leq M_A \\ 1 & M_A < x \leq IV \end{cases}, \tag{4.2}$$

(very much as in the general case), where $F_A^*(M_A; x)$ could be any distribution with support $[0, M_A]$ such as a Beta distribution or perhaps a scaled MBBEFD curve (although this will cause a discontinuity around M_A, since there will be a finite probability that the loss is equal to M_A).

As for the large-loss component, this is zero up to M_A and (for convenience, although this can be changed) a normalised MBBEFD curve with parameters b_L, g_L from M_A to MPL.

$$F_L(x) = \begin{cases} 0 & 0 \leq x \leq M_A \\ \left[F_{b_L,g_L}\right]\left(\frac{x-M_A}{MPL-M_A}\right) & M_A < x \leq MPL \\ 1 & MPL < x \leq IV \end{cases}. \tag{4.3}$$

Note that we have extended both $F_A(x)$ and $F_L(x)$ to IV to prepare for the case where $IV \geq MPL$.

We can now rewrite Equation (4.1) in expanded format as:

$$F_{A/L}(x) = \begin{cases} \frac{\lambda_A}{\lambda_L + \lambda_A} F_A^*(M_A; x) & 0 \leq x \leq M_A \\ \frac{\lambda_A}{\lambda_L + \lambda_A} + \frac{\lambda_L}{\lambda_L + \lambda_A}\left[F_{b_L,g_L}\right]\left(\frac{x-M_A}{MPL-M_A}\right) & M_A < x \leq MPL \\ 1 & MPL < x \leq IV \end{cases}, \tag{4.4}$$

which makes it clear that $F_{A/L}(x)$ is obtained by piling $F_A^*(M_A; x)$ and $\left[F_{b_L,g_L}\right]\left(\frac{x-M_A}{MPL-M_A}\right)$ on top of each other, after scaling them down by $\frac{\lambda_A}{\lambda_L+\lambda_A}$ and $\frac{\lambda_L}{\lambda_L+\lambda_A}$ respectively.

As a consequence, the exposure curve can be written as:

$$G_{A/L}(d) = w_A G_A(d) + (1 - w_A) G_L(d), \tag{4.5}$$

where $G_A(d)$ can be written as:

$$G_A(d) = \begin{cases} G_A^*(M_A; d,) & 0 \leq d \leq M_A \\ 1 & M_A < d \leq IV \end{cases}, \tag{4.6}$$

and $G_L(d)$ will be made of two pieces, a linear section between 0 and M_A (corresponding to the region where the CDF is zero), followed by a translated MBBEFD exposure curve with parameters b_L, g_L from M_A to MPL, then extended to IV.

$$G_L(d) = \begin{cases} \frac{d}{M_A + E([X_L]) \times (MPL - M_A)} & 0 \leq d \leq M_A \\ \frac{M_A + \left[G_{b_L,g_L}\right]\left(\frac{d-M_A}{MPL-M_A}\right) \times E([X_L]) \times (MPL-M_A)}{M_A + E([X_L]) \times (MPL-M_A)} & M_A < d \leq MPL \\ 1 & MPL < d \leq IV \end{cases} \cdot \tag{4.7}$$

Note that $E([X_L])$, the expected damage ratio of the MBBEFD portion of the curve can be calculated as $E([X_L]) = 1/\left[G_{b_L,g_L}\right]'(0)$, but we use $E([X_L])$ to make the meaning of the denominator of $G_L(d)$ clearer: the denominator is the expected value of a large loss, which is M_A (the minimum large loss) plus a fraction of $MPL - M_A$, with that fraction depending on the shape of the exposure curve only and equal to the damage ratio.

The two curves $G_A(d)$ and $G_L(d)$ can then be combined using the convex combination above, with

$$w_A = \frac{1}{\rho_\lambda \rho_{EL} + 1}, \tag{4.8}$$

which is similar to Equation (3.7) where the parameters have this slightly amended (but fully consistent with Equation (3.7)) definition:

- $\rho_\lambda = \lambda_L / \lambda_A$ is the ratio between the expected number of the large and the attritional losses.
- $\rho_{EL} = (M_A + E([X_L]) \times (MPL - M_A)) / E(X_A)$ is the ratio between the expected value for large and attritional losses.

Finally, the whole thing is wrapped up to allow for some premium (or rather, expected losses) to be allocated between MPL and IV (in case IV is larger than MPL). Assuming that any loss between MPL and IV is equally likely, we can write the following adaptation of Equation (3.19):

$$F_{\text{overall}}(x) = \begin{cases} (1-p)\, F_{A/L}(x) & \text{if } 0 \le x \le MPL \\ 1 - p + p\frac{x - MPL}{IV - MPL} & \text{if } MPL < x \le IV \end{cases}, \tag{4.9}$$

which can be fully expanded as

$$F_{\text{overall}}(x) = \begin{cases} (1-p)\frac{\lambda_A}{\lambda_L + \lambda_A} F_A^*(M_A; x) & 0 \le x \le M_A \\ (1-p)\left(\frac{\lambda_A}{\lambda_L + \lambda_A} + \frac{\lambda_L}{\lambda_L + \lambda_A}\left[F_{b_L, g_L}\right]\left(\frac{x - M_A}{ML - M_A}\right)\right) & M_A < x \le MPL \\ 1 - p + p\frac{x - MPL}{IV - MPL} & MPL < x \le IV \end{cases}. \tag{4.10}$$

Again, this has a clear explanation in that it is made up of three components on top of each other, in this proportion: $(1-p)\,\frac{\lambda_A}{\lambda_L + \lambda_A} : (1-p)\,\frac{\lambda_L}{\lambda_L + \lambda_A} : p$. Figure 3 explains it graphically.

This corresponds to the following exposure curve (Figure 4):

$$G_{\text{overall}}(u) = \begin{cases} k\left(pu + \frac{(1-p)G_{A/L}(u)}{G'_{A/L}(0)}\right) & \text{if } 0 \le u \le MPL \\ k\left(p \times MPL + \frac{1-p}{G'_{A/L}(0)} + p\left(\frac{(IV - MPL)^2 - (IV - u)^2}{2 \times (IV - MPL)}\right)\right) & \text{if } MPL < u \le IV \end{cases}, \tag{4.11}$$

where

$$k = \frac{1}{p \times MPL + \frac{1-p}{G'_{A/L}(0)} + \frac{p \times (IV - MPL)}{2}} = \frac{1}{\frac{1-p}{G'_{A/L}(0)} + \frac{p \times (IV + MPL)}{2}}, \tag{4.12}$$

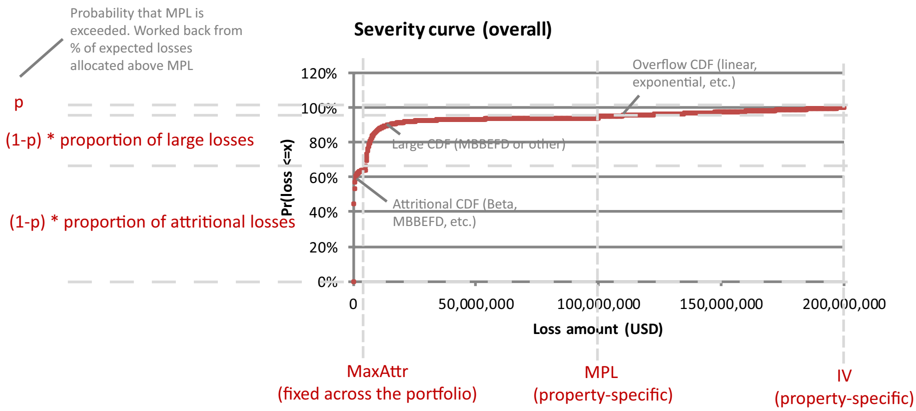

FIGURE 3: The CDF for the full model, where the losses are subdivided into attritional, large and overflow (i.e., above MPL). This chart was created with MBBEFD curves with $c = 3$ for both the attritional and the large loss component, and a linear CDF for the overflow part. The proportion of attritional losses is 70%. The probability of exceeding MPL is $p = 5\%$. Other parameters for this case: $M_A = \$5$ m, $MPL = \$100$ m, $IV = \$200$ m. Several of the parameters (and especially p) have been exaggerated for visualisation purposes.

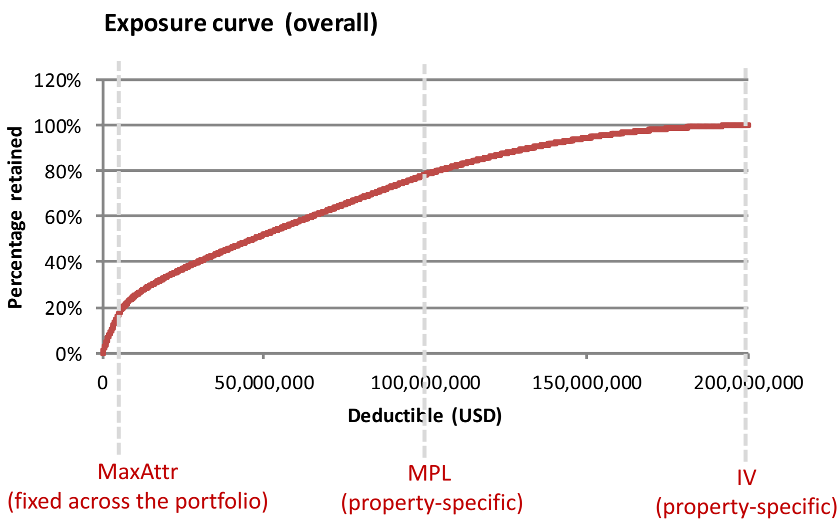

FIGURE 4: The exposure curve corresponding to the cumulative distribution function shown in Figure 3. Again, a reminder that the probability of exceeding the MPL is exaggerated for graphical clarity and that explains the significant percentage of the expected losses ($\sim 20\%$) allocated above \$100 m.

and

$$G'_{A/L}(0) = w_A \times G'_{A,1}(M_A; 0) + (1 - w_A) \times \frac{1}{M_A + (MPL - M_A) \times E([X_L])}. \tag{4.13}$$

Note that the denominator of Equation (4.12) has a straightforward interpretation as the expected value of the losses across the whole range, from 0 to IV, since $1/G'_{A/L}(0)$ is the expected loss below MPL and $(IV + MPL)/2$ is the expected loss above MPL, each multiplied by the relevant probability.

4.2. Alternative model

In the implementation described in Section 4.1, we have assumed that large losses can be modelled with an MBBEFD curve from M_A to MPL, shifted by M_A: specifically, we have modelled this section as an MBBEFD curve $[F_{b_L,g_L}]\left(\frac{x-M_A}{MPL-M_A}\right)$. This means that the large losses for any property will be modelled with the same curve between M_A and MPL, except from scaling the x axis.

Another option (which we have experimented with) is to use the same large-loss curve from 0 to MPL for all properties, and then replacing all losses between 0 and M_A with zeroes. This means that each curve will be slightly different depending on the ratio MPL/M_A (in case $MPL \le M_A$, there is of course no large-loss curve).

The jury is out on which model is more realistic – the amount of data available to us did not allow a data-driven decision on which option is better – but the model in Section 5.1 is obviously simpler to communicate and calibrate and should be chosen in the absence of evidence that it is not suitable.

In any case, we describe here the main features of the alternative model for the case where an MBBEFD model is used for the large losses. The attritional model will be identical to that described in Section 4.1. As for the large-loss severity curve, it can be described as follows:

$$F_L(x) = \begin{cases} 0 & 0 \le x \le M_A \\ \frac{[F_{b_L,g_L}]\left(\frac{x}{MPL}\right) - [F_{b_L,g}]\left(\frac{M_A}{MPL}\right)}{1-[F_{b_L,g_L}]\left(\frac{M_A}{MPL}\right)} & M_A < x \le 1 \end{cases}. \tag{4.14}$$

The survival function is given by:

$$S_L(x) = \begin{cases} 1 & 0 \le x \le M_A \\ \frac{[S_{b_L,g_L}]\left(\frac{x}{MPL}\right)}{[S_{b_L,g_L}]\left(\frac{M_A}{MPL}\right)} & M_A < x \le 1 \end{cases}. \tag{4.15}$$

As for the exposure curve $G_L(d)$, it will be made of two pieces: a linear section between 0 and M_{NS}/MPL, followed for example by another MBBEFD curve:

$$G_L(d) = \begin{cases} k \times d & 0 \le d \le \frac{M_A}{MPL} \\ k \times \left(M_A + \frac{[G_{b_L,g_L}](d/MPL) - [G_{b_L,g_L}](M_A/MPL)}{[G_{b_L,g_L}]'(M_A\ /\ MPL)} \times MPL \right) & \frac{M_A}{MPL} < d \le 1 \end{cases}, \tag{4.16}$$

where

$$k = \left(M_A + \frac{1 - [G_{b_L,gL}](M_A/MPL)}{[G_{b_L,g_L}]'(M_A\ /\ MPL)} \times MPL \right)^{-1} \tag{4.17}$$

is as in Section 4.1 the reciprocal of the expected value of the large losses.

The attritional/large model described above can then be extended to IV as in Section 4.1.

4.3. Other practical considerations for the implementation of the attritional/large model

We have so far described the basic elements of an implementation of the attritional/large model, for the standard cases where we have a combination of attritional losses (below M_A), large losses (from M_A to MPL) and a provision for the possibility of having mis-estimated the MPL and therefore having losses between the MPL and the IV.

In practice, however, one needs to take into account the possibility of more or less "pathological" cases and ensure that the pricing model can deal with all these different cases in a seamless way.

4.3.1. *Dealing with the case where $MPL < M_A$*

Occasionally, the MPL of a property may be smaller than the maximum attritional loss M_A as defined for a specific portfolio. In this case, we may use for that property the CDF $F_A^*(x) = F_A^*(M_A; x)/F_A^*(M_A; MPL)$, and the related exposure curve is $G(d) = G_A(M_A; d)/G_A(M_A; MPL)$. (For such small properties, we can safely assume that the MPL and the IV are equal.)

4.3.2. *Removing the discontinuity of behaviour at $MPL = M_A$*

An unintended consequence of the attritional/large model described in Section 4.1 and of the additional rule described in Section 4.2.1 for $MPL \leq M_A$ is that there is a discontinuity of behaviour between the case $MPL > M_A$ and the case $MPL \leq M_A$, because

$$\lim_{MPL \to M_A^+} G_{A/L}(d) = \frac{1}{\rho_\lambda \rho_{DR} + 1} G_A(d) + \frac{\rho_\lambda \rho_{DR}}{\rho_\lambda \rho_{DR} + 1} G_L(d) \neq G_A(d) \tag{4.18}$$

(remember that $G_{A/L}(d) = G_A(d)$ in case $MPL = M_A$). In order to avoid this effect (which is a logical consequence of the fact that the ratio of the frequency of large and attritional losses remains constant despite the fact that the MPL is approaching M_A), one can change the definition of the frequency ratio ρ^λ slightly as follows:

$$\rho_\lambda = \frac{\lambda_L}{\lambda_A} \times \frac{\max(0, MPL - M_A)}{MPL}. \tag{4.19}$$

This definition assumes that the large/attritional frequency ratio is progressively reduced as MPL approaches M_A, and approximates the usual definition when $MPL \gg M_A$. By adopting this definition, $\lim_{MPL \to M_A^+} \rho^\lambda = 0$ and Equation (4.18) becomes:

$$\lim_{MPL \to M_A^+} G_{A/L}(d) = G_A(d). \tag{4.20}$$

4.3.3. *Setting the exceedance probability above* MPL

As already discussed in Section 3.2.1, it is useful to define the probability p of exceeding MPL by first setting a maximum percentage of the expected losses that are allocated above MPL, and then setting the percentage of expected losses to the maximum amount times $1 - \frac{MPL}{IV}$ or $\left(1 - \frac{MPL}{IV}\right)^2$ to avoid a situation in which a substantial proportion of premium is allocated to a very small monetary amount.

4.3.4. *Case* MPL > IV

We have assumed that $MPL \le IV$. In some special circumstances, however, the MPL can be larger than the IV, if for example damage to a property can propagate to neighbouring properties, or when one insures a portion of a pipeline. In this case, from the point of view of the severity distribution one can (for the purpose of allocating the expected losses) ignore the IV and use an exposure curve that goes up to the MPL.

4.4. Calibration

The calibration of the attritional/large model may be quite different depending on the data available to the insurer and on the general data strategy. We describe below how this can be approached in the case where the MPL is known for all losses (Section 4.4.1) and when it is known only for the large losses (Section 4.4.2).

As for the calibration of the parameters related to MPL breaching, the discussion in Section 3.3.1 remains valid for the attritional/large model, with the additional provision that the estimate of the exceedance probability is simplified as it can be based on large losses only.

4.4.1. *Case where MPL is known for all losses*

In case the MPL is known for each loss, we can use the expectation-maximisation scheme described in Section 3.3. However, in this case the existence of a clear separation of attritional vs large losses allows some drastic simplifications, and the scheme reduces to a sequence of one-dimensional optimisation problems that can be solved straightforwardly.

We show below the EM scheme for the attritional/large model described in Section 4.1. The scheme can be easily adapted for use with the model described in Section 4.2.

The scheme will use the same basic assumptions (A.1–A.4) as in Section 3.3.

Step 1 – Initialisation

Set initial values $M_A^{(\text{old})}$ in an appropriate range. There is no need for initial values for p_A, c_A, c_L.

Step 2 – Expectation

The responsibilities are trivial in this case:

$$\gamma_{A,j}^{(\text{old})} = \Pr\left(A \,\middle|\, M_A^{(\text{old})}; x_j\right) = \begin{cases} 1 & 0 \le x_j \le M_A^{(\text{old})} \\ 0 & M_A^{(\text{old})} < x_j \le M_j \end{cases}. \tag{4.21}$$

Step 3 – Maximisation

The parameters $p_A^{(\text{new})}, c_A^{(\text{new})}, c_L^{(\text{new})}$ can be optimised separately like this:

$$\begin{aligned} p_A^{(\text{new})} &= \sum_{j=1}^{N} \gamma_{A,j}^{(\text{old})} = \frac{\#\text{ of losses} \le M_A^{(\text{old})}}{N} \\ c_A^{(\text{new})} &= \underset{c_A}{\text{argmax}} \sum_{\substack{\text{all } j \text{ s.t.} \\ x_j \le M_A^{(\text{old})}}} \left(\ln\left[f_{c_A}\right]\left(x_j / M_A^{(\text{old})}\right) - \ln M_A^{(\text{old})}\right) \\ &\quad \times \mathbb{I}_{\left[0, M_A^{(\text{old})}\right)}\left(x_j\right) - \ln g_A \times \mathbb{I}_{M_A^{(\text{old})}}\left(x_j\right) \\ c_L^{(\text{new})} &= \underset{c_L}{\text{argmax}} \sum_{\substack{\text{all } j \text{ s.t.} \\ x_j > M_A^{(\text{old})}}} \left(\ln\left[f_{c_L}\right]\left(\frac{x_j - M_A^{(\text{old})}}{M_j - M_A^{(\text{old})}}\right) - \ln M_j\right) \\ &\quad \times \mathbb{I}_{\left(M_A^{(\text{old})}, M_j\right)}\left(x_j\right) - \ln g_L \times \mathbb{I}_M\left(x_j\right). \end{aligned} \tag{4.22}$$

Note that g_A and g_L are calculated from c_A and c_L using Equation (2.7). As for M_A, it can now be updated by maximising with numerical methods the overall log likelihood:

$$M_A^{(\text{new})} = \underset{M_A}{\operatorname{argmax}} \left(\sum_{\substack{\text{all } j \text{ s.t.} \\ x_j \leq M_A}} (\ln [f_{c_A^{(\text{new})}}](x_j/M_A) - \ln M_A) \times \mathbb{I}_{[0,M_A)}(x_j) \right.$$
$$- \ln g_A^{(\text{new})} \times \mathbb{I}_{M_A}(x_j) + \sum_{\substack{\text{all } j \text{ s.t.} \\ x_j > M_A}} (\ln [f_{c_L^{(\text{new})}}](\frac{x_j - M_A}{M_j - M_A}) - \ln M_j)$$
$$\left. \times \mathbb{I}_{(M_A,M_j)}(x_j) - \ln g_L^{(\text{new})} \times \mathbb{I}_{M_j}(x_j) \right). \tag{4.23}$$

Step 4 – Iterate
Set $M_A^{(\text{old})} \leftarrow M_A^{(\text{new})}$. Go back to Step 2 and repeat until the improvement in the log likelihood is below the desired threshold. Note that the parameters $p_A^{(\text{new})}$, $c_A^{(\text{new})}$ and $c_L^{(\text{new})}$ are calculated anew at every iteration without reference to their previous value.

4.4.2. *Case where MPL is not known for attritional losses*
The calibration scheme described in Section 4.4.1 assumes that all loss data points have MPL information attached to them. In practice, this may not be the case, and MPL information is often available only for large losses (as more information is normally collected for large losses) but not for attritional losses, which may be recorded without reference to the originating location. Note that what counts as "large loss" in this case depends on the company business rules (e.g., the claims department may only record detailed information on losses above, say, £50,000) rather than on the nature of the loss.

In this situation a purely statistical estimation is formally possible but of no practical use because of the large uncertainties involved, and heuristic rules for the selection of parameters based on actuarial/underwriting judgment provide a more robust approach. Examples of these rules are described in the sections below.

4.4.2.1 *The choice of* $\mathrm{M_A}$. Since the identification of scale-independent and attritional losses is only a broad approximation, and there is additionally no obvious definition of an attritional loss, there is no obvious recipe for establishing what M_A should be. Considerations that should be made while making this selection are:

- M_A will in general depend on the type of occupancy, although one should strive to use as few values of M_A as possible for a specific portfolio.
- M_A should be higher than most deductibles that are used in practice, on the assumption that these deductibles are there to get rid of the attritional

losses. For example, if the deductible used for a given occupancy ranges between \$100,000 and \$1,000,000, selecting $M_A = \$1,000,000$ appears to be reasonable
- M_A should appear reasonable to the underwriters specialised in that particular sector.

4.4.2.2 *Calibration of the attritional curve*. If one accepts the admittedly brave assumption that all losses below M_A are attritional, the parameters of the attritional curve can be obtained by looking at individual large submissions with a large number of claims. Since attritional losses are assumed to be scale-independent, we do not need to worry about the fact that the MPL or the IV of the individual losses is often unavailable.

Ideally one should use submissions where the losses are given from the ground up. In practice, losses will be given above a deductible. If the deductible applied to each loss is known, it can be factored in the calibration of the curve by assuming a specific form for the CDF, for example, a Beta distribution or an MBBEFD, and by using MLE with censored data. If the deductible is not known and it can be assumed to be reasonably small on average, we can assume the average deductible to hold in all cases or ignore the deductible altogether, assuming (hoping?) that the shape of the curve will not be affected drastically.

4.4.2.3 *Calibration of the large-loss curve*. The parameter(s) for the large loss curve can be derived based on a claims database where all losses above M_A are reported. If an MBBEFD curve is used as suggested in Section 4.1, this is a straightforward problem of determining b, g with MLE after normalising all claims using the variable $u = (X - M_A)/(MPL - M_A)$.

Unlike in the case of the attritional curve, however, the information on MPL needs to be retrieved to estimate the parameters with any degree of accuracy. (Conceptually, it is still possible to derive an estimate even in the case where the MPL is not available – replaced perhaps by the knowledge of the statistical distribution of the MPLs – but the resulting parameter uncertainty is just too large for the results to be meaningful.)

In case not enough large loss information is available, the best strategy is to use a traditional value of c. Swiss Re's Guggisberg (2004) suggests for example a c-curve (a special case of the MBBEFD curve) with $c = 3.8$ for captives property damage – a relevant case to us because this curve is assumed to be in excess of an underlying deductible.

Note that depending on whether we use the model of Section 4.1 or Section 4.2, we need to remember that the probability of a "total" loss ($x \geq MPL$) is different, even though the same MBBEFD curve is used to represent large losses.

4.4.3. *Calibration of* p_A

This is perhaps the most crucial parameter when it comes to regulating the impact that a deductible has on the ROV. In principle, its value could be determined in a data-driven fashion by looking at the number of attritional and large losses per unit of IV for a large enough portfolio of clients and then calculating their ratio, according to the definition of ρ_λ ($\rho_\lambda = \lambda_L/\lambda_A$, and $p_A = 1/(\rho_\lambda + 1)$).

In practice, the results obtained in this way are beset by many uncertainties, including the unavailability of sufficient large loss data for some business sectors and the unavailability of LD information for attritional losses.

An alternative, safer approach (although not data-driven) is to use the underwriters' expert judgment on the impact of changes in deductible and translate it into constraints on the parameter p_A.

We are not going to delve into the details of how this translation exercise can be carried out. Suffice it to say that the underwriters can be surveyed by providing a number of statements regarding the most common situations in a particular business segment, such as:

- For a property with MPL = \$10 m, if the deductible goes from \$100 k to \$200 k, the premium should reduce by 10%
- For a property with MPL = \$10 m, if the deductible goes from \$100 k to \$50 k, the premium should increase by 5%
- For a property with MPL = \$50 m, if the deductible goes from \$100 k to \$200 k, the premium should reduce by 5%
- For a property with MPL = \$50 m, if the deductible goes from \$100 k to \$50 k, the premium should increase by 2.5%

and the value of p_A that best fits all these different constraints is selected for that business segment. This is a straightforward numerical optimisation problem with a single variable (p_A) that can be solved manually or by a simple spreadsheet function.

Figure 5 shows an example of how the deductible change rules listed above can be (approximately) achieved for different values of the MPL by setting the value of the percentage of attritional losses to $p_A = 91\%$ (in this particular case, the optimisation has been carried out manually). The calculations behind the graphs are available in Parodi (2019).

4.5. Using the model

In this section we look at how the attritional/large model can be used to answer the most typical questions that a D&F property pricing model is supposed to answer, that is: calculating the impact of deductible changes on the ROV (Section 4.5.1), calculating the expected losses to a layer (Section 4.5.2) and

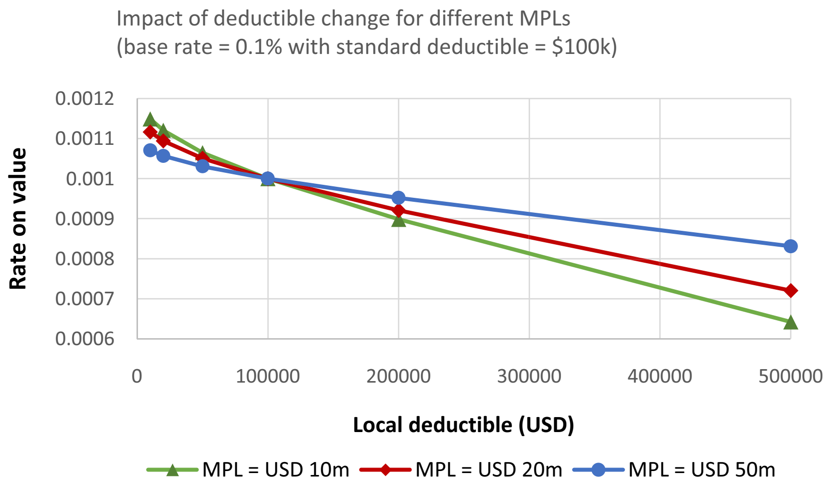

FIGURE 5: Impact of deductible change for different MPLs for the attritional/large model with the following parameters: attritional curve is an MBBEFD curve with $c = 4$; large-loss curve is an MBBEFD curve with $c = 3.8$; the maximum attritional loss is $M_A = \$1$ m. The insured value is equal to the MPL in all cases, and the probability of exceeding MPL is 0. The base rate is 0.1% at a standard deductible of \$100 k. The percentage of attritional losses is $p_A = 91\%$.

simulating losses via Monte Carlo simulation to produce the full aggregate loss distribution gross and net of the policy structure (Section 4.5.3).

4.5.1. *Calculating the impact of deductible changes on the ROV*

In the traditional framework, the relationship between the ROV at the SD and at an arbitrary deductible is given by Equation (2.16).

Using generalised exposure rating, if $\mathrm{BaseRate}_L$ is the base rate on an expected-loss basis, that is, the ROV when the underlying deductible is equal to the *SD*, the ROV at a different deductible *LD* is given by:

$$\mathrm{RateOnValue(LD)} = \mathrm{BaseRate}_L \times \frac{1 - G_{\mathrm{overall}}(LD)}{1 - G_{\mathrm{overall}}(SD)}. \tag{4.24}$$

Of course, a similar formula holds for the relationship between the ROV at *any* two different deductibles. Note that from-ground-up ROV is given by $\mathrm{RateOnValue}(0) = \mathrm{BaseRate}_L/(1 - G_{\mathrm{overall}}(SD))$.

Equation (4.21) then allows us to calculate the expected losses for a property whose *full value* (i.e., all losses above the underlying deductible, not just those to a specific layer) is insured:

$$E(S@\text{full value}) = \mathrm{RateOnValue}(LD) \times IV. \tag{4.25}$$

By setting *LD* to 0, one obtains the expected losses from the ground up: $E(S) = \mathrm{RateOnValue}(0) \times IV$.

4.5.2. *Calculating the expected losses to a layer* L xs D
Exposure curves' main job is to help actuaries and underwriters to calculate the expected losses to an insurance layer L xs D, where the deductible is defined in excess of a LD. In our generalised framework, this can be done by multiplying the expected losses from the ground up, $E(S) = \mathrm{BaseRate}_L \times IV/(1 - G_{\mathrm{overall}}(SD))$, by the difference between the value of the exposure curve calculated at $D + L + LD$ and at $D + LD$:

$$E(S(D, L)\,@LD) = \frac{\mathrm{BaseRate}_L}{1 - G_{\mathrm{overall}}(SD)} \times IV \times (G_{\mathrm{overall}}(D + L + LD) \\ - G_{\mathrm{overall}}(D + L))\,. \tag{4.26}$$

This is to be compared with Equation (2.19) for the traditional case.

4.5.3. *Simulating property losses*
In Section 2.3.3, we showed how the traditional property pricing model can be used to simulate losses in order to calculate the expected losses to a layer in the presence of complex features such as aggregate deductibles/limits, or to calculate non-trivial statistics such as value at risk or excess value at risk for a portfolio of properties (both of which could be used for capital cost calculations). In this section we show how to do loss simulation in the generalised exposure rating framework.

The discussion below applies directly to the attritional/large model of Section 4.1. We leave the reader to adapt it to the general case of Section 3 and to the alternative model of Section 4.2.

4.5.3.4 *Contracts with a single property*. Let us look first at the frequency model. Regardless of whether we choose a Poisson, a negative binomial distribution or some other count distribution, the expected number of losses for an individual property *from the ground up* can be written as:

$$E(N) = \frac{E(S)}{E(X)} = \frac{\mathrm{BaseRate}_L \times IV/(1 - G_{\mathrm{overall}}(SD))}{IV/G'_{\mathrm{overall}}(0)} = \frac{\mathrm{BaseRate}_L \times G'_{\mathrm{overall}}(0)}{1 - G_{\mathrm{overall}}(SD)}, \tag{4.27}$$

which replaces Equation (2.20), in which case however the expected number of losses were given in excess of the LD.

As for the simulation of loss amounts, Algorithm 1 in Loss simulation needs to be modified to take into account the fact that the normalised severity distribution is made of three pieces. A quite simple algorithm for the attritional/large model with MBBEFD distributions for both attritional and large losses is described below.

1. Inputs: parameters $b_A, g_A, b_L, g_L, M_A, p, \Pr(A)$; MPL, IV, LD
2. Generate a pseudo-random number u uniformly between 0 and 1
3. If u is between 0 and $(1-p) \times \Pr(A)$, then

$$\text{Loss} = [F]^{-1}_{b_A, g_A}\left(\frac{u}{(1-p)\, p_A}\right) \times M_A$$

4. If u is between $(1-p) \times \Pr(A)$ and $(1-p)$, then

$$\text{Loss} = M_A + [F]^{-1}_{b_L, g_L}\left(\frac{u-(1-p)\Pr(A)}{(1-p)\,(1-\Pr(A))}\right) \times (MPL - M_A)$$

5. If u is between $(1-p)$ and 1, then

$$\text{Loss} = MPL + p \times (IV - MPL)$$

6. Return Loss

Algorithm 3. *Generation of random variates from the attritional/large model. Note that* $\Pr(A) = \frac{\lambda_A}{\lambda_L + \lambda_A}$ *is the proportion of attritional losses and* p *is the percentage of losses above* MPL.

4.5.3.5 *Contract (with several properties)*. As in the case of Building the technical premium, the annual number of losses for the whole contract can be modelled as a Poisson with rate equal to the sum of the mean for each of the K individual properties. In our case, the expected number of losses from the ground up is given by:

$$E(N) = \sum_{k=1}^{K} E(N_k) = \sum_{k=1}^{K} \frac{\text{BaseRate}_{L,k} \times G'_{\text{overall}}\left(M_{A,k}, MPL_k, IV_k; 0\right)}{1 - G_{\text{overall}}\left(M_{A,k}, MPL_k, IV_k; SD\right)}. \quad (4.28)$$

As in the case of Equation (2.21), we have assumed that the parameters of the exposure curve, $G_{\text{overall}}\left(M_{A,k}, MPL_k, IV_k; u\right)$, vary from one property to the other, although we have not listed them all inside the function to avoid notation cluttering. Notice that we have allowed M_A to be different for different properties, although normally it will be the same for all properties of a specific category, for example, Energy Onshore: a client portfolio might, however, include disparate occupancies such as industrial plants and offices.

From here, things proceed much like in Building the technical premium: for each scenario the properties that are going to have a loss are picked proportionally to the expected number of losses for each property, $E(N_k)$, and the algorithm (for the case where the frequency model is Poisson) is similar to Algorithm 2, with only a couple of differences: (a) the loss amounts are generated using Algorithm 3 and not Algorithm 1, and (b) the loss amounts range from 0 to IV and so that no separate allowance for layers above MPL is required:

1. For each scenario, sample the number of losses from a Poisson distribution with rate $\lambda = E(N)$
2. Assign each of the losses to one of the K properties in proportion to $E(N_1), E(N_2)\ldots E(N_k)$, calculated for each property as per Equation (4.24)
3. For each loss: if property k is selected, sample a loss amount from 0 to IV_k using Algorithm 3.

Algorithm 4. *Generation of random loss variates for a contract in the case of the attritional/large model. Notice that losses between 0 and IV are simulated.*

Various tricks can be used to increase performance by reducing the number of losses generated, for example, by only generating losses above the LD. Also, the usual caveat is that this is only the basic form of the simulation engine and plenty of details need to be added in order to deal with all real-world complexities (see discussion at the end of Building the technical premium).

4.5.4. *Building the technical premium*

The rules for premium breakdown into expected losses, expenses, profit, etc. according to a "Cost+" framework (Parodi, 2014a, 2014b) will be broadly the same as for the traditional approach (Section 2.3.4). However, the fact that the exposure curve includes rules for dealing with MPL uncertainty means that it might not be necessary to include a separate allowance for the premium above MPL.

5. Limitations, Extensions and Further Research

This methodology allows to extend exposure rating beyond the case where all losses are scale dependent. This may help modelling more accurately the effect of changing the underlying deductible.

From a purely conceptual standpoint, this method appears better than the traditional approach of assuming that the exposure curve is valid only above the deductible and changes in the deductible only affect the overall ROV (Section 2.3). It can be considered broadly equivalent to the method of dealing with attritional losses and large losses separately, for example, by using a traditional frequency/severity model for attritional losses and another frequency/severity model for large losses (based on an exposure curve), and combining the results of the two models via simulation. However, our approach:

– simplifies the calculation of the deductible impact, because it does not require to assess the impact separately for the scale-independent and scale-dependent losses;
– is more compact as it uses a single exposure curve, which is calculated "in flight" based on the property's MPL and IV.

The most obvious limitation of the framework described in Section 3 is that, given a list of losses, it is not easy to separate the scale-independent and the scale-dependent losses. Section 3.3 shows how clustering techniques such as expectation-maximisation can be used to calibrate the losses under such circumstances, using a scheme that works when the MPL of every loss is known. Section 4 addresses this problem more radically in the case where information about the MPL is only partially available, by using a practical approximation (the attritional/large model) where scale-independent and the scale-dependent losses are assumed to be sharply separated by size. Of course, one limitation of the attritional/large model is that it is indeed only a rough approximation – the distinction between attritional and large losses is obviously more nuanced.

Another limitation is that the assumption that the proportion of attritional to large losses is constant across all values of MPL is only a rough approximation. We saw that adhering strictly to this assumption has unintended consequences when MPL approaches M_A (see Section 4.3.1). It is also not obvious that this assumption will hold as MPL increases beyond a certain amount.

5.1. BI losses

This framework applies mainly to property damage losses. Although in principle the methodology could be used for BI losses, in practice the concept of scale-independent losses is much less useful for BI, as BI losses will normally be proportional to the number of days of interruption and the MPL is assumed to be reached when that number of days equals the indemnity period.

5.2. Complexity of the methodology

At first sight, the methodologies described in Sections 3, 4.1 and 4.2 look more complex than the usual exposure rating method which uses a single exposure curve from the ground up. However, that is not a fair comparison since the simple method with a single exposure curve is not able to deal effectively with the impact of deductible changes or MPL uncertainty, and for this reason the traditional approach is to enhance that method in such a way that these issues can be addressed, as described in Section 2.3.

A fairer comparison is therefore between the traditional approach to D&F property pricing described in Section 2.3 and our generalised exposure rating methodology. The traditional approach uses an exposure curve for large losses, a transformation function $\mathbb{T}(SD, LD, MPL)$ to capture the deductible impact and rules describing the amount of premium allocated between MPL and IV. The generalised exposure rating framework uses a curve for scale-independent (attritional) losses, one for scale-dependent (large) losses and another for losses beyond the MPL. Although the details of the implementation may vary both for the traditional method and the generalised exposure rating method, the complexity of the two models (whether defined in terms of the number of parameters of the model or in terms of the length of the code necessary to describe it) is similar, and is ultimately a function of the number of things the model needs to take care of.

5.3. Using the insured's own experience

The methodology (in any of the incarnations discussed in this paper) can be easily extended to the case where the insured can share its own historical attritional loss experience. In this case the observed distribution can be plugged in to replace the portfolio/market attritional loss distribution. If an estimate for the portfolio/market large-loss frequency is also available, the ratio $\rho_\lambda = \lambda_L/\lambda_A$ (see Section 3.1) can be calculated from the data. The details of the implementation depend on which one of the methodologies presented in this paper (Sections 3, 4.1, 4.2) is adopted.

5.4. Further research

This work relies on the idea that losses can be divided into scale-dependent losses (the default assumption in exposure rating) and scale-independent losses. This reflects underwriters' intuition that attritional losses mostly do not depend on scale and are the ones that LDs are designed to get rid of. This paper develops this intuition and shows how the resulting framework can have the desired properties in terms of deductible impact, allowance for losses above MPL plus the standard workings of exposure curves.

More research is needed, however, to prove that this framework is fully supported by the data. Specifically, large-scale studies are needed to show the extent to which the distinction between scale-dependent and scale-independent losses is borne out in practice or is only a heuristic to obtain the desired behaviour, and whether the attritional/large model and the general model are a good fit for actual data.

The extent to which the attritional/large model is a good approximation of the general model should also be investigated. Its relationship with other possible ways of addressing the problem of the deductible impact, for example, the use of different curves for different levels of MPL, should be explored.

Another line of investigation was suggested in Parodi and Watson (2019), where the size of fire losses was modelled using weighted graphs (see Trudeau (2003) for a review of graph theory). Weighted graphs are networks of nodes representing property units (e.g., rooms) and arcs connecting units between which fire can propagate, with the weight of the arc representing the probability of fire propagation between the two units. If the propagation probability is below a certain critical threshold, fires remain confined to a small area of the graph, whose size does not depend on the size of the graph. If the probability is above that threshold, the size of the fire will be a significant fraction of the overall graph (and possibly the whole graph). If that graph-theoretic model captures something of the reality of fires, this gives a good theoretical grounding for the difference between scale-dependent and scale-independent losses and suggests that one can model both scale-independent and scale-dependent losses by using the same graph but with different weights. It would be an interesting research topic to assess what type of exposure curves result from using simulation based on these mixed graphs.

6. Implementation

An implementation of all the cases described in this paper, namely:

- the general case of Section 3 for a mixture of scale-dependent and scale-independent losses
- the attritional/large model of Section 4.1
- the alternative attritional/large model of Section 4.2

using MBBEFD models for both the scale-independent and the scale-dependent losses is available in spreadsheet format with code written in VBE in (Parodi, 2019). The spreadsheet only works under the assumptions stated in the paper, that is, $M_A < MPL \leq IV$. In a professionally designed tool, all the different sub-cases (see for example the discussion in Section 4.3) must of course be considered.

Note that all the figures provided in this paper were obtained with the spreadsheet tool mentioned here.

Acknowledgements

The methodology outlined here is currently at the basis of SCOR Business Solutions' property pricing model, which is part of SCOR's proprietary underwriting software, although obviously enhanced with all the necessary bells and whistles and consideration of all possible special cases.

Special thanks to Peter Watson who led the implementation of the property pricing model into SCOR's proprietary system, reviewed the methodology and proposed improvements where necessary. I would like to thank Prof. Mario Wütrich for his helpful suggestions relating to the calibration of mixture models. I would also like to thank Bertrand Le Gall, Francis Cheung, Homin Choi, Jean-Stephane Bodo, Joseph Lees, Kenneth Wells, Renaud Ambite, Renzo Cermenati, Rob Spaul, Robin Whiteley and Shane Lenney for many helpful discussions.

References

Bernegger, S. (1997) The Swiss Re exposure curves and the MBBEFD distribution class. *ASTIN Bulletin* **27**(1), 99–111.

Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*. New York: Springer.

Dempster, A., Laird, N. and Rubin, D. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, **39**(1), 1–38.

Devroye, L. (1986) *Non-Uniform Random Variate Generation*. New York: Springer-Verlag.

Dutang, C., Gesmann, M. and Spedicato, G. (2019) *Exposure rating, destruction rate models and the MBBEFD package*, https://cran.r-project.org/web/packages/mbbefd/vignettes/Introduction_to_mbbefd.pdf.

Guggisberg, D. (2004) *Exposure Rating*, Zurich: Swiss Re. http://www.swissre.com.

Lee, G.Y. and Frees, E.W. (2016) *General insurance deductible ratemaking with applications to the local government property insurance fund*, https://www.soa.org/resources/research-reports/2016/2016-deductible-ratemaking/.

Ludwig, S.J. (1991) An exposure rating approach to pricing property excess-of-loss reinsurance. *Proceedings of the Casualty American Society*, **LXXVIII**, 110–145.

Parodi, P. (2014a) *Pricing in General Insurance*, New York: CRC Press.

Parodi, P. (2014b) *Appendix 3 – Introduction to commercial property pricing*, http://www.pricingingeneralinsurance.com/home/additional-text.

Parodi, P. (2019) http://www.pricingingeneralinsurance.com/home/downloads/genexposurerating.

Parodi, P. and Watson, P. (2019) Property graphs – A statistical model for fire and explosion losses based on graph theory. *ASTIN Bulletin*, **49**(2), 263–297. doi: 10.1017/asb.2019.4.

Riegel, U. (2010) On fire exposure rating and the impact of the risk profile type. *ASTIN Bulletin* **40**(2), 727–777.

Salzmann, R.E. (1963) Rating by layer of insurance. *Proceedings of the Casualty American Society*, **LXXVIII**, 110–145.

Trudeau, R.J. (2003) *Introduction to Graph Theory*. New York: Dover Publications Inc.

PIETRO PARODI (Corresponding author)
SCOR Business Solutions, 10 Lime Street
London EC3M 7AA, UK
E-mail: pparodi@scor.com