Institute and Faculty of Actuaries

Validating Operational Risk Models

by the Operational Risk Working Party

Madhu Acharyya, Andrew John Couper, Patrick Oliver John Kelliher *,Edward Niall Vincent Maguire, Choong Ai Pang, Christopher Michael Smerald, Jennifer Karen Sullivan, Paul Matthew Teggin.

Presented to the Institute & Faculty of Actuaries

10 October 2023
Staple Inn Hall, London

Abstract

Operational Risk is one of the most difficult risks to model. It is a large and diverse category covering anything from cyber losses to mis selling fines; and from processing errors to HR issues. Data is usually lacking, particularly for low frequency, high impact losses, and consequently there can be a heavy reliance on expert judgement.

This paper seeks to help actuaries and other risk professionals tasked with the challenge of validating models of operational risks. It covers the loss distribution and scenario-based approaches most commonly used to model operational risks, as well as Bayesian Networks. It aims to give a comprehensive yet practical guide to how one may validate each of these and provide assurance that the model is appropriate for a firm's operational risk profile.

Keywords

Operational risk, Loss distributions approaches, Scenario based approaches and Bayesian Networks.

Correspondence details

*Correspondence to: Patrick Kelliher, IFoA, 1-3 Staple Inn Hall, High Holborn, London, WC1V 7QH

E-mail: professional.communities@actuaries.org.uk

Contents

Contents 4
1. Introduction 6
2. General 6
3. Loss Distribution Approach (LDA) models 8
3.1 Loss Data 8
3.1.1 Data Governance and Quality Assurance: 8
3.1.2 Losses captured: 9
3.2 Distribution Fitting and Results 11
3.2.1 Distribution Fitting Methodology: 11
3.2.2 Model Results 12
4. Scenario Based Approach (SBA) models 13
4.1 Scenario Analysis Results 13
4.1.1 Risk Coverage 13
4.1.2 Scenario analysis process 14
4.1.3 SMEs involved 14
4.1.4 Quality of background information provided to SMEs 14
4.1.5 Scenario outputs – General 15
4.1.6 Scenario outputs – Frequency 15
4.1.7 Scenario outputs – Loss quantification 16
4.1.8 Scenario outputs – Recoveries 16
4.1.9 Scenario Analysis Governance 16
4.2 Reviewing SBA model results 17
4.2.1 Recurring losses 17
4.3 Structured Scenarios 18
4.4 Scenario Analysis example 18
5. Aggregation and allocation 18
5.1 Correlation assumptions: 18
5.2 Aggregated Results 19
5.3 Aggregation across BUs and legal entities 19
5.4 Allocation 19
6. Causal Factor-based models 20
6.1 Bayesian Network Models 20
7. Conclusion 22
References 23

Appendix A – sample causal factors 24
Appendix B – Scenario validation example 25
Appendix B.2 Validation points I – scenario generation 26
Appendix B.3 Validation points II – modelling 26
Appendix C – Simple Bayesian Network model example 29
Figure C.1 – Causal Factor relationships 29
Appendix C.1 Bayesian Network Modelling 29
Figure C.2 – Sample Bayesian Model output 30

1. Introduction

The Operational Risk Working Party aims to assist actuaries and others in the modelling and management of operational risk. One challenge faced may be to validate models of operational risk as either internal or external reviewers. This paper sets out points which actuaries and other risk professional may wish to consider in validating such models, both in general and with respect to the following types of operational risk model:

- Loss distributions approaches (LDA);
- Scenario based approaches (SBA); and
- Causal factor-based approaches such as Bayesian Networks (BNs).

It also considers the validation of methods for aggregating different types of operational loss allowing for diversification, and the allocation of aggregate figures by risk type and legal entity.

The paper focuses on financial services companies and in particular insurers, banks and asset managers, but it is hoped this also has wider relevance to other firms.

2. General

In validating operational risk models, the validator will need to first understand the design and structure of the model; its operation in practice; and how it fits in with other risk models. The following points should be considered:

a) **Operational Risk taxonomy / definition:** there should be a clear articulation of what is classed as operational risk – and what is not classed as operational risk – and how the model addresses the former. Without a clear taxonomy, operational risks may not be captured by the model and/or the model may make allowance for non-operational risks already modelled.

b) **Operational Risk profile and changes to this:** validators should seek to gain a reasonable understanding of the organisations current operational risk profile from regular risk reporting including any changes to this. This will help understand what the key operational risk types are where the validation should focus particular attention on, as well as recent or upcoming changes to risk profile (e.g. new distribution channels) which may not be picked up by modelling.

c) **Model choice:** what is the rationale behind the operational risk model chosen (e.g. LDA)?

   i. This should include an articulation of what the purpose of the model is – for financial services firms, the model will typically be developed with a view to assessing capital requirements for operational risks, but models could also be developed for wider management purposes e.g. to model the potential impact of cyber-attacks.

   ii. The rationale should be clear as to why a particular approach was chosen over other approaches, having regard to model purpose.

   iii. Parsimony – validators should always consider whether a model is overly complex for the uses intended, but particularly for operational risks where issues with data often mean that model complexity is spurious.

d) **Governance** – validators should consider the governance around operational and other models. Amongst other things, this should include:

   i. Model risk policy and standards setting out how models should be developed, validated, approved and reviewed as well as the roles of different stakeholders;

   ii. Assessments of model limitations and risks, with logs kept of limitations, key expert judgements required, and planned model developments to address limitations; and

iii. A regular model oversight committee to ensure models are regularly reviewed and revised where appropriate, and that model risk policy and standards are being adhered to.

A proper model governance framework does not necessarily mean an operational risk model is fit-for-purpose, but the absence of such a framework makes it more likely the model will be flawed. The Institute and Faculty of Actuaries (IFoA) Model Risk Working Party has produced a number of papers on model risk and how this can be mitigated, including governance around model[1]. The IFoA has also produced a paper (Ashcroft et al, 2015) on expert judgement and the framework for using such judgements.

e) **Model use** and integration with wider risk and capital management:

i. For UK and EU insurers looking to use an operational model as part of a wider internal model, validators will need to consider how the operational model satisfies the use test requirement in Article 120 of the Solvency II Directive[2].

ii. More generally, validators should seek evidence for how the model is used in practice, and whether it is being used for its intended purposes. If a model is not being used for its intended purposes and/or is not embedded in the risk and capital management system of a firm, then a question arises as to how serious the modelling effort is – if the firm isn't prepared to use model results in practice, it calls into question whether model results are suitable for say determining regulatory capital requirements.

iii. Alternatively, if the model is being used for purposes beyond its original scope there is a risk is may be mis-used, raising questions as to model governance.

iv. Note that when it comes to the use of operational risk models, the journey may be as important as the end result – more may be gained from the use of loss data or scenario analysis in business as usual (BAU) risk management than the capital figures arising from these inputs to models.

v. Model use is not a one-way street – there should be evidence of feedback from BAU risk and capital management as to how well the model is performing for its intended purpose, and that this feedback should be fed into a program of continuous model improvement.

f) **Documentation** – validators should consider whether documentation of such a standard that they or another independent knowledgeable third party would be able to understand the design, structure and operation of the model and how it interacts with other models. To the extent a validator is not able to fully rely on documentation but has to query points further with developers, this would indicate that model documentation is un-satisfactory and as such does not comply with the requirements of Article 125 of the Solvency II Directive[3].

More generally poor documentation is often the sign of poor model development practices; increases key person risk to those with the knowledge missing in documentation; and makes it more likely the model will be mis-understood and mis-used.

g) **Culture** - validators should seek to understand the culture of the organisation and how this influences modelling, loss reporting and risk management. For instance, CFOs and CEOs may push for lower capital requirements so they can improve dividends, and this could introduce a downward bias to subjective assumptions and hence model results.

[1] See the Model Risk Working Party page at Model Risk | Institute and Faculty of Actuaries

[2] This requirement is captured in sub-section 10 of the Solvency Capital Requirement - Internal Models part of the PRA rulebook – see https://www.prarulebook.co.uk/rulebook/Content/Chapter/212834/18-02-2022

[3] See sub-section 15 of the Solvency Capital Requirement - Internal Models part of the PRA rulebook https://www.prarulebook.co.uk/rulebook/Content/Chapter/212827/18-02-2022

Another aspect of culture is the extent to which staff own up to mistakes and weaknesses. A fearful corporate culture may inhibit the reporting of operational losses or weaknesses in controls, resulting in model results which understate the true operational risk profile.

It is acknowledged that culture can be difficult to define, let alone measure, but validators should try to assess the firm's culture and the extent to which this might affect operational risk management and modelling.

h) **Benchmarking** – finally, ORIC International[4], ORX[5], KPMG[6], E&Y and others produce regular modelling surveys including operational risks which can be useful to benchmark methodology and aggregate results.

3. Loss Distribution Approach (LDA) models

LDA models are based on fitting statistical distributions to internal and external loss data. Validation needs to consider first the loss data used, and then the results from distributions fitted.

3.1 Loss Data

The Operational Risk Working Party has produced a paper (Kelliher et al, 2016) on inputs to operational risk models which addresses loss data and which should be considered in validating loss data, but in essence loss data should be assessed for:

- **Accuracy** – are loss figures correct? What controls are in place to ensure accuracy of loss figures?
- **Completeness** – how far back do they go? How significant are events not in data (ENID)? Do they miss notable periods of stress? Do they fail to capture key elements of loss?
    - Are there any shifts in loss experience which may call into question the relevance of prior losses?[7]
    - Are there any outliers in data which should be excluded? Or conversely any data points excluded as outliers which should be included in loss data?
- **Appropriateness** – are historic losses relevant to the firm's current risk profile (noting this includes legacy exposure)?

Key points which validators may wish to consider:

3.1.1 Data Governance and Quality Assurance:

[4] ORIC is a consortium of insurers and other financial institutions which share operational loss data for modelling and management purposes. ORIC produces research into operational risk including regular benchmarking studies - see for example, ORIC's 2020 Annual Capital Benchmarking Survey, with the Summary Report available at: 44340f_2f07eaf9a5f545d9ba0c1af08a8edd64.pdf (filesusr.com)

[5] ORX is a consortium of banks and other financial institutions which share operational loss data. Like ORIC, ORX also produces research into operational risk which can help with benchmarking – see Operational risk management in financial services | ORX

[6] As well as the operational risk section in their annual Technical Practices Survey, KPMG have also produced a detailed market survey of operational risk modelling which they presented to the 2018 Life Conference – see https://www.actuaries.org.uk/system/files/field/document/F2%20Life%20Conference%20Operational%20Risk%202018.pdf.

[7] It is noteworthy that the PRA considers that the information value of operational risk losses generally diminishes over time as business models change – see pgh.8.24 of "*CP16/22 – Implementation of the Basel 3.1 standards: Operational risk"* at Chapter 8 – Operational risk | Bank of England

a) Is there a documented process in place for collecting operational loss data? Lack of such a process makes it more likely that loss data will be inconsistent, incomplete and/or corrupted.

b) Do 2nd line risk management and/or internal audit review of loss data? This would give some comfort as to accuracy, completeness and appropriateness of data.

c) Do senior manager provided attestation to data on losses arising from their area? Again, this would provide some comfort on the accuracy, completeness and appropriateness of data, but if senior managers don't sign off, there is a risk their staff may not pay due attention to recording losses accurately.

3.1.2 Losses captured:

a) Does loss data include "near misses"? or events which give rise to gains (e.g. dealing errors where the market moves in a firm's favour)? Excluding these could ignore significant information relating to risk events.

b) Are these still relevant? For example, legacy mis-selling losses may no longer be relevant if a firm no longer involved in providing advice.

c) Even if relevant, there may be a case to adjust loss amounts for inflation. There may also be a case to scale losses with an exposure measure such as the number of customers.

d) Does loss data capture impacts on the present value of future profits ("PVFP") such as caps imposed on future charges? These would be relevant for life insurers as they would affect Solvency II Own Funds and so should be included in loss figures.

e) Assessments of economic capital typically exclude new business and if so, loss data should exclude new business impacts of operational risk events.

f) Similarly, some impacts such as lapses may be covered elsewhere in the economic assessment so should be similarly be excluded.

g) Loss data might include costs such as management time which may be part of BAU costs as opposed to the marginal costs arising from the operational loss event. Such BAU costs should be excluded for the purposes of economic capital assessments.

h) Validators should consider whether operational loss data includes boundary events which may be captured under other risk models e.g. where an operational risk event gives rise to a credit loss, as this may be captured under credit risk, or where it affects insurance claims, as it may be implicitly allowed for as part of insurance risk.

   That said, where operational losses are assumed to be covered under another category, validators should confirm this assumption.

i) For insurers, some operational risk losses might be captured as part of expense analysis and so might be implicitly allowed for as part of base maintenance expense assumption and expense risk capital. If this is the case, then these losses should be excluded from LDA data to prevent double-counting.

   However, the validator should verify this by considering the expense analysis as sometimes losses which are assumed to be implicit in expense assumptions are in fact excluded from the expense analysis and hence from maintenance expense assumptions and stresses to these.

j) Recurring losses – to the extent losses may be recurring, there may be a need to consider whether these should be included in base maintenance expense assumptions and budgets as opposed to being covered as part of operational risk capital, in which case such losses should be excluded from LDA data.

k) Recoveries – best practice would be to model losses on a gross basis in the first instance before insurance and other recoveries, as modelling net losses may implicitly extrapolate recoveries beyond sum insured limits; and insurance arrangements may change.

l) A variation of insurance recoveries would be where losses are recovered from third party suppliers or passed on to customers. For the former, there may be limits to the indemnity provided by the supplier, or the ability to charge back losses may be disputed, while for the latter, there would be regulatory rules around treating customers fairly and other constraints. For these reasons, loss data should ideally be modelled on a gross basis in the first instance to avoid recoveries being implicitly extrapolated beyond recovery limits in model.

3.1.2 External Loss Data

Where a firm uses external operational loss data, validators should consider the following:

a) How are losses scaled for the firm's own size? What might be the impact of alternative methods of scaling on LDA results?

b) Linked to this, are any adjustments made to external losses to reflect risk profile? For instance, a mortgage lender may not have the same exposure to dealing errors as say an investment bank.

c) Are losses actually relevant to a firm's operations? For instance, unit linked pricing errors would not be relevant to an annuity writer.

d) Could similar losses arise given the firms control environment? For instance, a cyber loss arising at a peer due to the use of unsupported software may not be relevant to a firm if it keeps software up-to-date[8].

Ideally there would be 2nd line risk management and/or internal audit review and challenge of external loss data points included

3.1.4 Risk coverage / ENID

There is a risk that historic data may not capture the tail of operational risk loss events for a category and/or material operational risks. Thus, the validator should consider:

a) What is the distribution of losses by risk type? are there any categories in the firm's operational risk taxonomy with no data? Ideally there would be sufficient data to calibrate models for each Basel Level 2 risk category or the equivalent under the firm's own risk taxonomy[9].

b) What is the length and size of dataset? How many data points does it contain? How far does it go back? Could it omit any notable losses before data collections started e.g. misselling losses? or are there too few points to properly understand and model the tail?

c) Are there any large losses deliberately omitted from data? If so, what is the rationale for this? If these could plausibly recur, then they should be included in scope.

d) If external data is not being included in the LDA, validators should consider external loss events and whether there are any types of loss types which are not captured by internal loss data.

e) To the extent that losses below a certain threshold may be excluded, there is a need to understand the scale of these, lest in aggregate they are significant.

   This could perhaps be assessed by comparing recorded operational loss totals with any accounting figures for total operational loss, or perhaps unexplained variances from budget.

[8] Though consideration should be given to failure of current controls before excluding external losses.

[9] Basel Level 2 categories can be found in Annex 9 of Annexes to "International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version", June 2006 (bis.org).

To the extent that there are significant differences, consideration should be given to an add-on to operational risk capital to allow for these. Note however that these low-level losses may be included in expense analyses and so may be implicitly allowed for already in expense assumptions.

3.2 Distribution Fitting and Results

Having validated data inputs, the next stage will be to validate the approach to fitting distributions and the results from distributions fitted.

3.2.1 Distribution Fitting Methodology:

a) Frequency / Severity – a key model choice is whether to fit distributions to operational loss frequency and severity separately or just fit a single distribution to combined loss data. Most banks and insurers adopt the former approach[10], so the choice of the latter is something that will need to be explained.

   For separate frequency / severity models, the same severity distribution is usually assumed for all loss events, but the validator should challenge if this is reasonable. It may be the second and subsequent events be more severe than the first e.g. if the firm was less resilient as a result of a loss event; or subsequent losses could be less severe e.g. where the first event reduced exposure to subsequent losses. Where there is a significant probability of 2+ loss events arising, the validator should question the assumption of the same severity distribution for all events.

b) Choice of distributions to be fitted – the validator should review the rationale for distributions chosen to be fitted to loss data and whether there are plausible alternative distributions[11] which could be also used but which were not considered. Ideally the sensitivity of results to these plausible alternatives should be quantified.

   Often a compound severity distribution may be fitted e.g. a Log-Normal distribution may be fitted to the body of losses, with a Pareto distribution fitted to the tail. The validator should review the loss threshold at which the tail distribution is fitted including the rationale for the loss threshold chosen and the sensitivity of results to alternative thresholds.

c) Approach to parameter estimation: method of maximum likelihood v method of moments – the latter may give rise to biased estimates but may be simpler to implement.

[10] From figure C.1.2 of ORIC International "Annual Capital Benchmarking Survey - Summary Report" (ORIC, January 2020), 75% of insurers and investment managers surveyed model frequency and severity separately; while for banks, from page 6 of "*Observed range of practice in key elements of Advanced Measurement Approaches (AMA)*" (Basel Committee on Banking Supervision, July 2009), nearly all banks adopting the AMA model frequency and severity separately (see https://www.bis.org/publ/bcbs160b.pdf).

[11] For example, a plausible alternative to the Poisson distribution for frequency may be the Negative Binomial distribution.

d) Granularity – there will be a trade-off between (a) number of risks modelled and the homogeneity of each risk category; and (b) volume of data in each cell. Too many cells may result in insufficient data to credibly fit distributions but validators should be aware that too few cells could mean a distribution is fitted for heterogenous risks, akin to fitting a single distribution to motor and property claims in general insurance. Another issue is how granular is data e.g. loss data may not be broken beyond the equivalent of the Basel Level 1 categories.

e) Treatment of extreme losses – are these excluded from fit? If so, there needs to be a valid rationale exclusion.

   Alternatively, if extreme loss events are included in loss data, assumptions may be made as to the percentiles these form of fitted distributions. The validator should identify any such assumptions; ensure there is a rationale for the choice of percentile; and assess the sensitivity of results to alternative assumptions.

f) Maximum loss caps – and caps on losses modelled need to be justified in terms of boundary constraints for losses such as portfolio size, other measures of total exposure (e.g. total number of staff per location) and legal boundaries such as time-bars.

3.2.2 Model Results

In reviewing results, the validator should consider:

a) Goodness of fit – a wide range of goodness of fit statistics should be produced as part of the fitting process including amongst other things:

   i. Kolmogorov–Smirnov (K-S) test scores;

   ii. Akaike information criterion (AIC) and/or Bayesian information criterion (BIC) scores; and

   iii. Q-Q plots.

   These should support the choice of distributions fitted, though validators should also be wary of over-fitting complex distributions to immature loss datasets.

   Testing goodness of fit might also include tests for unimodality which assess whether distributions with a single peak value – like most distributions typically fitted – are appropriate for the data, or whether a bi-modal distribution would be better. Tests for modality includes Hartigans' dip test (Hartigan and Hartigan 1985).

b) Sensitivity analysis – amongst other things this should consider the sensitivity of modelled results to:

   i. Different distributions;

   ii. Different assumptions e.g. percentile corresponding to historic extreme loss events; or threshold for fitting separate tail severity distributions; and

   iii. Adjusted data sets e.g. excluding extreme loss events (if included in data); the impact of excluding new data; or different scaling approaches to external data.

   Validators should challenge models which are very sensitive to small changes in data, while the sensitivity results to different distributions and assumptions is something that should be clearly sign-posted in model documentations and communications of results to model users.

c) Simulation error – where separate frequency and severity distributions are fitted, the combined distribution of losses will usually be derived through simulation, but validators should seek to establish the extent of simulation error around results. In particular, the number of simulations chosen should be supported by convergence testing.

d) Kurtosis – immature datasets with few extreme loss events may give rise to thin tailed severity distributions which may not be fit-for-purpose where there is the possibility of extreme losses under a category.

e) Modelled recoveries – the validator should ensure these are in line with maximum sums insured; indemnity limits and/or consistent with what is possible to charge back to customers under consumer protection legislation.

   i. Validation should assess the possibility that losses could arise as a result of uninsured perils, or that there could be a dispute over coverage. It should also consider the possibility of disputes over liability with third party suppliers which could limit recoveries under indemnities

   ii. For losses charged back to policyholders, there is a need to consider limits such as regulatory rules on treating customers fairly or, for a with profits or participatory fund, its Principles and Practices of Financial Management (PPFM) or equivalent. For with profits, any modelled chargeback should be put to the With Profit Committee for their views on whether this would be acceptable.

4. Scenario Based Approach (SBA) models

LDA models are common in banks in part because the Basel II Advanced Measurement Approach (AMA) was built on such an approach, but also because banks often have large amounts of operational loss data. However, insurers typically have less loss data to work with, so the predominant approach is to model operational risk using a scenario-based approach (SBA) where models are based on loss scenarios derived by expert judgement, or a hybrid approach combining elements of LDA and SBA[12]. From ORIC's 2020 Capital Benchmarking Survey of insurers and asset managers, 70% of respondents used SBA while the remained used a hybrid approach[13].

4.1 Scenario Analysis Results

Scenario-based approaches are driven by the judgements of subject matter experts (SMEs) and will be inherently subjective. However, validators should ensure that scenarios are assessed in a structured fashion, as free as possible from bias. The Operational Risk Working Party's paper on inputs to operational risk models (Kelliher et al 2016) sets out good practice for scenario analysis in this regard. Taking this into account, key points that should be considered in validating SBA approaches:

4.1.1 Risk Coverage

SBA will typically focus on a limited number of scenarios, and validators should consider if this set of scenarios adequately captures the range of operational risks to which a firm may be exposed to:

a) Granularity – often risks will be considered by the equivalent of Basel Level 2 risk type which should ensure at least every Level 2 category is covered, but if not, there is a need to map scenarios considered to Level 2 category to ensure no such risk type is excluded from scenario analysis.

b) While it would be impractical to consider all lower level (Level 3) sub-risks, there should be evidence that these have been considered as part of the process for identifying representative

[12] One example of such a hybrid approach may be to use LDA for risk categories where there is significant loss data (e.g. unit pricing errors) and use SBA for categories where there is little data; another may be to calibrate the frequency distribution based on LDA but use SBA for the severity distribution. In approaching hybrid models, validators should consider the points in this paper for LDA and SBA for the respective loss-data and scenario elements of the model.

[13] See figure C.1.1 of the ORIC 2020 Summary Report at: 44340f_2f07eaf9a5f545d9ba0c1af08a8edd64.pdf (filesusr.com)

scenarios for a particular risk category. This might for example include detailed breakdown of risks by sub-type as part of background material to be considered by those involved in scenario analysis, and minutes of scenario workshops considering these sub-types.

c) The Working Party would be of the view that scenario analysis carried out just for the seven Basel Level 1 high level categories (or their equivalent) is unlikely to give a suitably broad coverage of risks.

d) Scenario participants should also be asked to consider internal and external loss examples to ensure no notable operational risks are missed.

e) Scenarios considered but discarded before arriving at final scenarios should be captured as part of scenario documentation to provide evidence they have been considered. Validators should consider if there is bias in scenarios chosen compared to those discarded e.g. is there a focus on more common but lower impact scenarios? Or vice versa on less likely, but higher impact scenarios?

f) Often scenario analysis will be based on previous years scenarios but with scenarios refreshed – these is a risk this might miss new risks emerging so there is a need to consider how well the scenario process covered changes in risk profile and notable new losses and emerging risks (e.g. increase in home-working with Covid-19).

4.1.2 Scenario analysis process

Given the subjective nature of SBA, it is important that the process for arriving at scenarios is robust, so validators should consider the following:

a) How is scenario analysis performed? Is it through workshops or some other process for eliciting SME views. One such approach is the Delphi method[14] where SMEs may be asked first to provide their responses individually rather than in a workshop. This can reduce the impact of certain biases, such as group think.

b) How are workshops moderated? There is a risk that these can be denominated by an individual and their concerns, to the exclusion of other valid scenarios.

c) How well are workshops minuted? Can the validator get a sense for the points discussed? And for differences in opinion between SMEs?

d) Linked to this, is there any evidence of bias in discussions? e.g. focus on recent events, and/or a lack of consideration for events which have not occurred yet.

e) What follow-up is carried out? A common weakness is where loss estimates are arrived at during a workshop without follow-up investigation to firm up loss estimates.

4.1.3 SMEs involved

The quality of SBA depends on the quality of SMEs contributing to the development of scenarios, so validators should consider how were these chosen – do they represent a suitable breadth of expertise across the organisation? This should include consideration of their level of seniority, length of service and relevant qualifications.

4.1.4 Quality of background information provided to SMEs

The quality of scenario outputs will be linked to the quality of background information supplied to SMEs. Validators should review the material considered by SMEs before arriving at scenarios. Ideally this should include:

[14] For more details, see Delphi method - Wikipedia

a) Definition of operational risk to be considered – needed to ensure clarity in terms of the scope of the operational risk being assessed and ensure that all sub-risks are considered as above, but also that there is no duplication of assessments.

b) Relevant historic losses – both internal and external, and ideally including “near misses” and gains.

c) Details of risk control assessments and any current risk and control issues.

d) Relevant exposure details e.g. current sales by product and channel for mis-selling risk, or the number of employees for employee relations risk.

e) Details of strategy and plans which may affect exposure e.g. planned investment in controls and system upgrades; outsourcing initiatives; or planned new products (especially if these new products are complex or involved new distribution channels).

4.1.5 Scenario outputs – General

a) Scenarios chosen and their associated frequency and loss estimates may reflect control failures. Scenario analysis outputs should be clear on what controls if any are assumed to fail. The validator should consider whether sufficient allowance is made for control failures in scenarios or if too much faith is being placed in controls with little consideration for the impact of failure.

b) For a frequency / severity approach where conditional loss estimates are sought i.e. loss assuming a loss event has occurred, a common issue is that SMEs confuse this with the unconditional probabilities. By way of example, we may have a scenario with a 1-in-5 year frequency and seek say a 1-in-20 conditional loss estimate. The latter should correspond to a 1-in-100 loss, but too often SMEs seek to articulate the loss at a 1-in-20 level. Validators should be aware of the potential for confusion in this instance.

c) Where we seek a typical loss and a more extreme loss estimates, it is common for the latter figure to be based on the typical loss albeit of greater severity, but it may be better to consider a different scenario for the latter e.g. for a cyber risk scenario, the typical loss could be based on a data breach, but the extreme loss could encompass not just a data breach but also a ransomware attack.

4.1.6 Scenario outputs – Frequency

Typically, a frequency / severity approach will be adopted with scenario analysis seeking to arrive at parameters for the frequency of loss events and a number of estimates for the loss given an event has occurred. For the former, it may help to ask SMEs to choose from a limited range of frequencies (e.g. once in every 2/5/10/20/40 years) to help ensure consistency between assessments and avoid spurious estimates.

The frequency parameter could be validated against historic loss event frequencies but for high impact, low probability risks, a firm’s own experience is typically not long enough for such risks to crystallise. Even where an event has arisen, it can be difficult to judge frequency based on a single data point.

Another issue to be aware of is whether the frequency parameter relates to the likelihood of a material loss event arising in a particular risk category, reflecting all sub-risks, or whether it just relates to the sub-risks covered by loss scenarios. If the latter, consideration needs to be given to how well other sub-risks are covered by modelling – it may be that the scenario addresses low frequency, high impact loss events but that high frequency, low impacts may need to be separately addressed (see section 4.2.1 below).

4.1.7 Scenario outputs – Loss quantification

Typically, two or more loss estimates will be derived to inform the severity distribution. The quality of these loss estimates is often an issue, particularly where these are arrived at in a workshop and not as part of a follow up analysis.

The validator should assess:

a) Loss elements assessed – validators should check whether certain types of loss are missed e.g. the impact on PVFP if a scenario impacts on future charge or premium income, which should be included in loss estimates as this may affect Own Funds. Validators should also consider if there are loss elements such as new business impacts which may not be relevant to an economic capital assessment; or certain costs such as management time which may be part of BAU costs as opposed to the marginal costs due to the operational loss.

b) Boundary losses – as for LDA, there is a need for validators to consider whether scenario losses relating to credit losses or insurance claims could already be covered as part of credit and insurance risk capital already.

c) Quality of data inputs used in assessing loss estimates e.g. validators should check the source and veracity of portfolio details used to assess exposure to cyber-attacks.

d) Assumptions including research used to set these e.g. for data theft, there are a number of studies published[15] which could be used to set loss assumptions and/or validate these.

e) Tools – validator should consider whether there is proper quality assurance around spreadsheets used for loss quantification.

f) Maximum loss caps – as for LDA, any caps to loss assumed should be justified in terms of boundary constraints for loss such as portfolio size, other measures of total exposure (e.g. total number of staff per location) and legal boundaries such as time-bars.

4.1.8 Scenario outputs – Recoveries

As for LDA, best practice would be to model gross losses in the first place, but to the extent that recoveries are allowed these need to be consistent with sums insured and supplier limits of indemnity.

a) Validation should also consider the possibility that the loss may arise as a result of an uninsured peril, or that there could be a dispute over coverage / supplier liability.

b) For losses charged back to customers, there is a need to ensure this is consistent with legislation and ideally reviewed by any committee responsible for the fair treatment of customers, such as the With Profit Committee for UK with profit products.

4.1.9 Scenario Analysis Governance

Given the high level of subjectivity around scenario analysis, it is important that there is robust governance around the process. Validators should look for evidence of the following:

a) 2nd line / Internal Audit independent review and challenge of scenarios.

b) Senior management sign-off – ideally, individual managers would be responsible for the sign off of scenarios relating to their area (e.g. HR manager signs off on employee relations

[15] IBM / Ponemon Institute and Verizon currently publish an annual study of cyber losses - see https://www.ibm.com/security/digital-assets/cost-data-breach-report/#/ and https://enterprise.verizon.com/resources/reports/dbir/?cmp=paid_search:google:ves_us:gm:awareness&utm_medium=paid_search&utm_source=google&utm_campaign=ves_us&utm_content=gm&utm_term=awareness&gclid=EAIaIQobChMIpLT1kaPU6QIVTtbACh3MeQCyEAAYASAAEgK4XvD_BwE&gclsrc=aw.ds

scenario). This should ensure senior management engagement with the scenario analysis process, and help drive up standards.

If senior managers are not required to sign off results, validators should consider whether the scenario exercise has been given proper consideration or whether it has been delegated to junior staff with little “buy in” from SMEs.

If senior managers do sign off on results, validators should also be wary of challenges and changes made by senior managers. For instance, a senior manager may be concerned about their reputation if the potential for a large loss is identified in their area, and may seek to artificially reduce figures to preserve face.

4.2 Reviewing SBA model results

Key aspects to consider in validating the model and results:

a) Deriving distributions from loss estimates: typically, we might collect 2+ loss estimates encompassing an expected / typical loss and one or more severe / extreme loss estimates. The former might be assumed to be the mean / median of the loss distribution, with severe loss estimates assumed to relate to higher percentiles, but results can be very sensitive to what percentiles are assumed.

   Sensitivity analysis should be performed on the impact on model results if different percentiles were assumed.

b) Kurtosis – depending on the loss estimate values, how close these are to each other, and the percentiles of distribution these are assumed to represent, the resulting loss distribution could be thin tailed or extremely fat tailed, so there is a need to assess the kurtosis of calibrated distributions.

c) Distribution – sensitivity analysis should be performed to understand the impact of alternative distributions on results e.g. for frequency, Negative Binomial instead of Poisson (former allows greater variance in events); or for severity, Weibull instead of Lognormal. However, given the subjectivity of scenario analysis inputs, the use of complex distributions with 4+ parameters is not likely to be appropriate.

d) Stability – often, Monte Carlo simulation is used to model the combined frequency and severity distribution, but results can be very unstable at the tail due to simulation error, and it may require 1m+ simulations to achieve stability for a tail loss estimate. The number of simulations used should be supported by convergence testing.

   Validators should also look for sensitivity testing to small changes in scenario inputs, including any changes as a result of scenario refreshes, and challenge models where small variations in inputs lead to large changes in results (though from (b), this could arise due to the gap between loss estimates and the resulting kurtosis of the scenario distribution).

e) Back-testing – having calibrated the model, it would then be useful to compare this against historic losses. If the model ascribes a low probability to a recent large loss, then it may be the case that the model is weak.

f) Alternative scenarios – it would be useful to consider the impact of alternative scenarios discarded as part of the process e.g. a lower impact but higher frequency scenario, though it may be the case that loss estimates for these may be less robust.

4.2.1 Recurring losses

Scenario analysis may capture infrequent, high impact losses but not high frequency, low impact losses. The validator should consider whether an additional allowance may be required for these on

top of scenario analysis capital modelled[16]. This could be based on the historic average of such losses, perhaps capitalised to reflect their recurring nature. On the other hand, it may be that such losses are implicitly included in base expense assumptions and expense risk capital (see 3.1.2 (h) above).

4.3 Structured Scenarios

A development to be aware of in the field of SBA is the use of what may be termed structured scenarios which seeks to adopt a more objective approach to scenario quantification. It does this by constructing formulas to calculate losses based on how the scenario affects underlying drivers, for example the number of customers impacted by a cyber-attack, or the value of physical assets which could be affected by a natural disaster scenario. This approach could be used to strengthen expert judgement and avoid some of the pitfalls associated with scenario analysis results outlined in section 4.1 above. Further detail on structured scenarios can be found in Kramer and Ramakrishna (2016).

4.4 Scenario Analysis example

Validating scenarios and modelled results involves a good understanding of operational risks as well as knowledge of risk modelling. Appendix B gives an example of the issues the validator may need to consider just for a single risk category, in this case financial reporting.

5. Aggregation and allocation

Under LDA and SBA, operational losses are generally modelled by risk type, so there is a need to aggregate these allowing for diversification between operational risk types and possibly with non-operational risks. The IFoA Operational Risk Working Party has produced a paper on dependencies (Kelliher et al, 2020) which gives a good overview of the issues surrounding operational risk aggregation but validators should consider the following:

5.1 Correlation assumptions:

a) Empirical assumptions: need to consider whether these may be compromised by lack of data, while for low frequency risk types, correlations may be systematically under-estimated[17].

b) There is likely to be some reliance on expert judgement to set these, but there needs to be 2nd line / Internal Audit independent review and challenge of these.

c) Correlation assumptions should ideally be supported by causal analysis, with underlying drivers of each risk type identified and compared between risk types to identify common drivers.

d) Macro scenarios could be help identify common linkages. Examples of these could include (see Appendix A):

   i. Change of government leading to a different legal and regulatory environment.

   ii. Pandemics like Covid-19 – amongst other things, these could disrupt service, increase backlogs and potentially increase cyber risk exposure due to increased working from home.

[16] Under Basel II AMA, the focus is on unexpected losses, whereas expected losses arising from high frequency, low impact may be assumed to be covered by expected profits emerging. For insurers however, this distinction is moot as such expected profits will usually be allowed for in Own Funds and thus either Technical Provisions or Solvency Capital Requirements should make some allowance for these.

[17] See section 6.1 of Kelliher et al, 2020.

iii. Economic downturns could lead to an increase in fraud, or expose existing loan and other fraud.

iv. Market falls could lead to mis-selling and other conduct losses arising.

v. Change programmes could go awry leading to system outages, data breaches and reporting errors amongst other things.

vi. If these have not been considered already, validators could use these scenarios to assess the reasonableness of correlation assumptions, probing SMEs on scenario impacts to understand common exposures and dependencies.

e) Consideration should also be given the impact of Reverse Stress Test, ORSA, regulatory and other scenarios carried out as part of wider risk management on operational losses and what this implied for correlation assumptions.

5.2 Aggregated Results

a) Sensitivity analysis should show the impact on results of different correlation assumptions (e.g.+/-25%).

b) Ideally sensitivity testing would also involve testing different methodologies e.g. if using a Gaussian copula, the impact of using a T-copula; of if using a T-copula, the impact of different degrees of freedom parameters. That said, the choice of copula may be spurious given the subjectivity of operational risk correlation assumptions.

c) Given correlation assumptions and undiversified requirements, it should be easy to aggregate the latter using a variance-covariance approach as a broad check on copula results, though for technical reasons, this is likely to be higher than Gaussian copula requirements.

d) Benchmarking studies often give details of diversification benefits which can be used to assess the level of own diversification benefits, but it should be noted that these will vary with the number of risks modelled by each firm, with those modelling more operational risks typically seeing greater diversification benefits.

5.3 Aggregation across BUs and legal entities

often operational risk is modelled at BU and/or legal entity level, and there is then a need to aggregate the results across these to arrive at a Group figure. This could be a simple additive approach, or allowance could be made for diversification between BUs and/or legal entities. If diversification is allowed for, the validator should consider the reasonableness of correlation assumptions using causal analysis as above, noting amongst other things that:

a) Weaknesses in group-wide governance, risk management and compliance could lead to diverse operational losses arising across BUs and legal entities

b) There may also exposure in respect of group-wide systems e.g. a key group system failing could affect multiple BUs and legal entities, or group systems could be exploited as part of cyber-attacks resulting in transmission across IT networks.

c) Reputation damage suffered by one part of the group could trigger downgrades and mass lapses across the group. This in turn could result in operational losses across the group e.g. as a result of backlogs arising in lapse processing, or the need to invoke deferral clauses on unit-linked funds due to mass lapsing.

5.4 Allocation

Having aggregated operational risk across categories, BUs and legal entities, there may be a need to allocate diversified capital back to these levels. The validator should have regard to the following:

a) Allocations should have regard to service level agreements between companies which may preclude the charging back of certain operational losses e.g. employee relations losses.

b) Operational risk capital allocated back to with-profits funds needs to be consistent with PPFM or other governance arrangements for the funds and what these say about charging losses to the Fund e.g. it would be inappropriate to allocate conduct risk capital to a With Profit Fund if these cannot be allocated to the Fund in practice.

6. Causal Factor-based models

LDA and SBA operational risk models have been widely criticised for not being dynamically risk sensitive. This is driven by many reasons, but most notably that historical loss data makes no allowance for current exposures of the firm, scenario analysis is subjective, and it is difficult to incorporate a direct link between risk and controls in existing model designs. Furthermore, operational risk models are often only recalibrated on an annual basis. Thus, operational risk models may not have a direct link to the dynamic risk exposures of the firm.

As a result, an emerging field in operational risk modelling is the use of causal factor-based models such as Bayesian Networks. These approaches assume that operational risk exposure can be described by a function of a set of underlying causal factors or drivers. By identifying and quantifying these factors, a risk sensitive model can be developed, based on how these factors link together to drive the frequency and severity of loss events. To the extent that data feeds can be automated to calibrate data-driven exposures, this can lead the way to real-time assessment of operational risk.

6.1 Bayesian Network Models

Bayesian networks are built upon a framework of causal factors, pulling together both the known (factors) and unknown (probabilities) into a visual node map to describe each risk process. The factors will be a mixture of firm-specific and external risk drivers.

Causal factors are used to identify key exposure metrics and conditional probabilities are used between the factors to estimate the frequency and size of loss when a given exposure is impacted. Together these can estimate the value at risk for a given unit of exposure. Appendix C provides a simple example of how causal factor modelling is overlaid with conditional probabilities for loss events, control failure and impact to provide a holistic distribution of losses.

Bayesian Network models can aggregate several risk processes or scenarios, via the use of conditional probabilities and common risk drivers, meaning that the resulting loss outputs are sufficient to calculate capital requirements, negating the need for assessing subjective correlations between risk types.

One of the key advantages of Bayesian Network models is that they can easily be visualised with the use of the node or factor map. These visuals are intuitive and easy to understand for any stakeholder in the business, allowing a much wider breadth of stakeholder challenge. This in turn will increase understanding of each business process and associated risk and controls, providing a much closer link between risk measurement and risk management approaches.

Conditional probabilities can be quantified to take into account the impact of controls and mitigating actions, thus providing a platform for what-if analysis, helping to analyse the impact of specific controls and identifying which controls have the most impact on resultant loss, which can be used in future planning and investment choices.

Bayesian Networks can be complex and difficult to set up initially, requiring specialised knowledge of key business processes and risk drivers. External consultancies and software packages can however be utilised to support with the initial set up, and the effort involved can be repaid by the greater insight into operational risks and how thee interact.

Validation of model design will be easiest when starting with the simplest designs and introducing additional complexity in a step-by-step process. Care needs to be taken to ensure the model does not

become overly complex too quickly, resulting in too many subjective estimates and resulting in spurious accuracy of model results. Each additional variable added should be tested to ensure that the value-add in accuracy of model output outweighs the additional complexity.

The model could be utilised across several areas in addition to capital, for example, operational resilience or recovery and resolution testing, as well as emerging areas such as climate and cyber risk quantification.

Validation points to consider are:

- What causal factors are driving losses? Are there any notable omissions? Appendix A below provides a list of basic causal factors which could be considered.
- Risk coverage – while it would be unrealistic for a Bayesian Network model to cover every sub-risk (of which there could be 300+), the validator should assess its ability to model key risks types across each Basel Level 2 or equivalent high-level category.
- Conditional probability assessments are difficult and can be subjective. Where these are data driven, data choice and reliability should be regularly validated. Where these are judgement driven, a log of judgments, SMEs contributing to the judgement and materiality should be kept.
- Discrete probability nodes may utilise a 'high, med, low' assessment, which when combined with other continuous probability nodes may give the user a false sense of accuracy to the output. Sensitivity testing of node materiality is important in communicating this.
- The model will be highly sensitive to the inter-dependence of nodes, particularly across different risk processes (akin to the high sensitivity of traditional models to dependency assumptions). These variables will likely be the most difficult to quantify given the limited data available and hence communication of this subjectivity will be very important.
- Back-testing of the model against material historical losses, annual losses e.g. in stressed periods and to existing scenario assessments will help validate model performance and identify any gaps.
- Quantification of the impact from specific controls failing may be challenging. SME judgement could be sought alongside analysis of near-miss event data.
- Scenario analysis: ideally, firms would supplement Bayesian Networks with forward looking scenario analysis as a check on the former model's ability to cater for scenario losses, but if not, a validator may wish to construct his/her own scenarios to test this, perhaps based on notable historic loss events.
- IT systems – some IT packages may struggle with multiple drivers of risk (e.g. where there is more than 2 drivers of a particular loss) and the computation this gives rise to. Validators need to understand any IT limitations of the Bayesian Network model, noting Article 245 (f) of Solvency II Delegated Regulations requires limitation of IT used in internal models to be documents.

7. Conclusion

Validating operational risk models can be a complex undertaking but it is hoped this paper will provide a useful starting point for this. While considerations will vary by type of model, a number of common themes should be borne in mind:

- **Risk profile** – validators need to have a good understanding the operational risk profile being modelled and whether any changes in this are being picked up.
- **Risk coverage** – operational risk is a diverse category with potentially hundreds of sub-risk types. While it is probably not feasible to model every single sub-risk type, validators should ensure the resulting model is broadly representative of the risks which may arise under a particular category.
- **Recurring losses** – modelling may naturally focus on high impact, low frequency operational losses, but validators should consider how low impact, high frequency losses are allowed for.
- **Loss data** – this may not be enough to calibrate loss distributions and/or capture tail exposures. Validators should also be aware that loss data may:
    - not capture certain relevant impacts (e.g. impacts affecting PVFP);
    - include impacts not relevant to economic capital (e.g. lost new business);
    - include losses no longer relevant (e.g. legacy mis-selling); and/or
    - include boundary losses that may be covered other risk models.
- **External loss data** – validators need to consider how this is scaled to a firm's size and whether it is relevant to the firm's business, control environment and risk profile.
- **Subjectivity** – expert judgement is likely to be required for most operational risk model and validators should ensure this is made as part of a structured process with the appropriate level of expertise involved and robust review, challenge and sign-off of assumptions.
- **Recoveries** – exposures should first be modelled on a gross basis and validators should consider whether modelled recoveries are in line with insurance policy limits and coverage; outsourcing arrangements and/or regulatory and PPFM constraints for amounts charged back to policyholders.
- **Aggregation and allocation** – correlation assumptions are likely to be subjective and should be tested against underlying causal factors, while allocations to legal entities and With-Profit Funds should be consistent with service levels agreements and PPFMs respectively.
- **Documentation** – in general poor model documentation which may highlight a lack of professionalism around how this the model was developed, while it is critical that scenario documentation is sufficient to evidence the breadth of risks covered and the quality of discussions around the choice of scenario and loss quantification.

Last but not least, validators need to have regard to intangible factors such as the culture of the firm commissioning the operational risk model and whether this could lead to bias and/or the understatement of risks; and/or the quality of SMEs and extent of senior executive involvement which can point to a lack to commitment to proper model development.

References

Ashcroft, M. et al (2015) "*Expert Judgement*", produced by the IFoA's Solvency & Capital Management Working Party and available for download from the profession's website at expert-judgement-paperfinal-8-june-2015-sessional.pdf (actuaries.org.uk)

Basel Committee on Banking Supervision (2006) "*Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework – Comprehensive Version"*, June 2006, available online from the Bank of International Settlements (BIS) website: https://www.bis.org/publ/bcbs128.htm

Basel Committee on Banking Supervision (2009) "*Observed range of practice in key elements of Advanced Measurement Approaches (AMA)*" available for download at: https://www.bis.org/publ/bcbs160b.pdf

Hartigan, J. A. and Hartigan, P. M. (1985) "*The Dip Test of Unimodality*", The Annals of Statistics 13 (1) 70 - 84, March, 1985 available at https://doi.org/10.1214/aos/1176346577

Kelliher, P. et al. (2016) "*Good practice guide to setting inputs for operational risk models*", produced by the IFoA's Operational Risk Working Party and available from the profession's website at: https://www.actuaries.org.uk/system/files/field/document/FINAL%20PAPER%20FOR%20WEBSITE.pdf

Kelliher, P. et al. **(2020)** "*Operational risk dependencies*", produced by the IFoA's Operational Risk Working Party and available from the profession's website at: https://www.actuaries.org.uk/system/files/field/document/Operational%20Risk%20Dependency%20Paper_0.pdf

Kramer, A. and Ramakrishnan, K. (2016) "*Operational risk scenario analysis: a structured approach*", Risk Americas presentation by Ernst and Young, which is available at: 2.15-AndrewKramerKarthikRamakrishnan.pdf (cefpro.com)

McGinnity, E. and Pang, N. (2018) "*Operational Risk Modelling Market Survey*" – this was presented by Nicole Pang and Eamon McGinnity of KPMG at the IFoA's November 2018 Life Conference, and is available for download at: https://www.actuaries.org.uk/system/files/field/document/F2%20Life%20Conference%20Operational%20Risk%202018.pdf

ORIC International (2020) "*Annual Capital Benchmarking Survey Summary Report*" available for download from 44340f_2f07eaf9a5f545d9ba0c1af08a8edd64.pdf (filesusr.com)

Prudential Regulatory Authority (2022) "*CP16/22 – Implementation of the Basel 3.1 standards: Operational risk"* available at Chapter 8 – Operational risk | Bank of England

Appendix A – sample causal factors

The following is not an exhaustive list but it may be useful for a validator to consider any gaps and explore the reasons for non-inclusion.

External factors

- Change of government leading to a different legal and regulatory environment.
- Litigation results in an adverse ruling for the firm or industry (e.g. Law Lords ruling on Equitable Life Guaranteed Annuity Options).
- Natural disasters such as floods could damage offices and/or prevent staff getting to work, leading to backlogs.
- Pandemics like Covid-19 – amongst other things, these could disrupt service, increase backlogs and potentially increase cyber risk exposure due to increased working from home.
- Widespread cyber-attacks like NotPetya and WannaCry.
- Economic downturns could lead to an increase in fraud, or expose existing loan and other fraud.
- Market falls could lead to mis-selling and other conduct losses arising.
- Geo-political events such as the Russian invasion of Ukraine leads to market turmoil and wider economic impacts including higher inflation and recession.

Internal factors

- Weak compliance culture could lead to losses across all categories including conduct failings and failure to adhere to internal controls.
- Ambitious targets could lead staff to cut corners and fiddle figures.
- Poor recruitment processes lead to unsuitable staff being recruited – either because they are dishonest; or because they do not have the education levels necessary to perform roles correctly and avoid errors.
- High staff turnover could lead to a loss of experience, increasing processing errors and exacerbating weaknesses in recruitment (this in itself could be driven by a culture of bullying and harassment, or by stress caused by over ambitious targets)
- Change programmes could go awry leading to system outages, data breaches and reporting errors amongst other things.
- Poor model controls lead to errors in pricing and valuation models.

Appendix B – Scenario validation example

The following outlines a hypothetical example of scenario analysis output for Financial Reporting Risk for a life insurer, the modelling of the 99.5th percentile based on a scenario-based approach, and the points validators may consider in reviewing this.

B.1 Hypothetical scenario output and modelling

A workshop of SMEs from Actuarial and Finance came up with the following scenario parameters:

1. Probability of a material loss (defined as a £100k+) crystallising under the Financial Reporting Risk category over the coming year *– 1-in-5 = 20%*, chosen from prescribed options of 1-*in-2/5/10/20/40 years*.
2. Typical loss event (assumed to be the median loss assuming a material loss event occurs) – a minor error in Report and Account disclosures gives rise to an extra £0.25m in external audit and consultancy fees.
3. Severe loss event (assumed to relate to 90th percentile of losses i.e. 9 out of 10 material losses are less severe) – an error in the calculation of with-profit option and guarantee costs results in a £20m increase in Technical Provisions / reduction in Own Funds, which is the materiality level which would trigger a re-statement of accounts.

This was then fed into operational risk modelling as follows:

4. Financial Reporting Risk losses are modelled using a compound frequency / severity model based on Poisson and Lognormal distributions respectively:
    - Poisson parameter (λ) = 0.2
    - Lognormal – with median and 90th percentiles of £0.25m and £20m we get the following parameters:
        - μ = -1.386294
        - σ = 3.419314
5. 1m simulations are generated to model the compound loss distribution, giving rise to a 99.5th percentile Financial Reporting Risk loss of £200m. Other percentiles of the simulated compound loss distribution:
    - 90th percentile - £0.2m
    - 95th percentile - £2.4m
    - 99th percentile - £67.8m
    - 99.9th percentile - £1.7bn
6. For this risk category, no allowance is made for recoveries under E&O and other insurance policies; compensation from third party suppliers under indemnity clauses; nor for chargeback of losses to with-profit policyholders.
7. No cap is applied to losses simulated.

Appendix B.2 Validation points I – scenario generation

On the face of it, the scenarios reflect a good level of understanding of the different types of financial reporting losses, but validators may wish to explore the following points:

a) How well was the workshop documented? Ideally documentation would give detail not just on the final scenarios selected, but the rationale for the selection of these scenarios as well as any alternative scenarios considered but rejected.

b) Who was involved in the workshop? What were the qualifications and experience levels of participants? Was there any function involved in financial reporting not represented?

c) What information was fed into the scenario workshop? The quality of scenario outputs will depend in part on how well workshop participants are informed. Amongst other things information supplied to participants should include detail historic loss events – both internal (if any) and external – as well as the state of financial reporting controls including any audit finding and, looking forward, proposed changes in financial reporting e.g. new accounting rules like IFRS17.

d) What risks and scenario were considered as part of the scenario workshop? – and what risks weren't considered? Ideally, workshop participants would be supplied beforehand with a detailed taxonomy of Financial Reporting Risk sub-types to consider and discussions would include which of these are material and the potential losses under material sub-types.

e) What is the rationale for the choice of frequency parameter? Is it based on historic loss events? Does it reflect any weaknesses in control and/or new reporting requirements which may affect the likelihood of a material loss arising?

f) What is the rationale for the typical and severe loss scenarios selected? How are these more appropriate than other loss events considered? For instance, the severe case loss is based on an error in the calculation of with-profit liabilities but what about liabilities for other products such as annuities? Or errors in the valuation of illiquid assets?

g) How were loss estimates arrived at?

    - Losses should be broken down into components (e.g. fines, extra audit costs, Own Funds re-statements etc.), but what is the rationale behind the figures for each component? Are they crude "guesstimates" or are they grounded in facts e.g. internal or external historic loss events, benchmarking information of regulatory fines etc.?
    - A common failing of scenario workshops is that the loss estimate is arrived at the workshop with no further analysis conducted offline with the result is that the loss estimate is often of poor quality.
    - In this example, the severe case loss was based on the minimum error that would trigger a re-statement of accounts, but has the potential for higher losses than this being considered?
    - Have and boundary constraints / upper limits to loss been considered?

h) Who reviewed and challenged the scenarios to ensure these were of appropriate quality? Risk Management? Internal Audit? Is there any evidence of senior management ownership of the scenario results e.g. sign-off by a senior executive?

Appendix B.3 Validation points II – modelling

Overall validation should encompass the choice of modelling approach and distributions used, but specific points to consider in this example:

i) Choice of distribution – for Financial Reporting Risk, there may be a case for using a Normal distribution as errors may be symmetrical i.e. an increase in Own Funds from correction of an error may be as likely as likely as a reduction. Other possible distributions include:

- Negative Binomial for frequency – this can have a greater variance than the Poisson distribution; and
- Generalised Pareto Distribution for severity – this could be calibrated based on typical and severe losses but would also require an assumption for the threshold loss (possibly based on the definition of material loss).

j) Even if Lognormal distribution and the approach to its' calibration is judged appropriate in general, results can very sensitive to the difference between typical (/median) and severe (/90th percentile) loss estimates.

- In this instance the large difference between the two calibration points (severe = 40x typical) results in a very "fat tailed" distribution with an extreme kurtosis value, which explains why the 99.5th percentile loss is ca.10x the severe case loss, so it is important to consider the kurtosis of the fitted severity distribution.
- Note that increasing the typical loss to £2m would perversely reduce the 99.5th percentile loss of the combined loss distribution from c.£200m to c.£70m as it results in a thinner-tailed severity distribution.
- Validators should also be aware that when typical and severe losses are close, then this approach could give rise to a "thin tailed" severity distribution which may be inappropriate for modelling operational losses.

k) Assumptions – the extreme kurtosis and high 99.5th percentile estimate are also a function of the assumption that the severe case loss represents the 90th percentile – in this instance it may be more appropriate to assume the severe loss represents say the 95th percentile.

l) Sensitivities – varying the percentile assumption for the severe case loss is just one sensitivity that should be tested. Amongst others that should be considered is the impact of using a different frequency parameter (e.g. $\lambda = 0.5$ or 0.1 based on the prescribed choices around the 0.2 selected); and the impact of varying either or both loss estimates (e.g by 10%).

m) Simulation error – while generally 1m simulations will be enough to give a stable 99.5th percentiles loss, even this number may not be enough where frequency is low and/or the severity distribution is very skewed. Validators should consider how stable the results are to different sets of random numbers.

n) Sense check on results:

- Loss values should be considered against limits to loss. For instance, if a 99.5th loss greater than the total value of liabilities may be implausible while if were simulating potential over-statement of asset values, any loss in excess of aggregate asset value would not be plausible.
- Ideally, we will know the number of losses generated in each of the 1m simulations, which we can summarise and compare with the Poisson distribution.
- If we also know the loss in each simulation, we can sum this over all 1m simulations, divide by the total number of losses and compare this with the mean loss from the calibrated severity distribution (in this case £86.5m).
- The 99.5th percentile loss of the combined loss can be approximated by looking at the "equivalent percentile" of the severity distribution calculated as:

$$Equivalent\ Percentile = [\ 1 - \{\ 1 / (200 \times \lambda)\ \}\ ]$$

So, in this case, with $\lambda = 0.2$ i.e. a 1-in=5 event, the 1-in-200 combined loss by looking at the 97.5th percentile i.e. a 1-in-40 loss level. This gives a value of £203m which is of the same order as the 99.5th percentile simulated.
Note however, this approximation tends not to work for $\lambda > 0.5$.

o) Aggregation considerations – as noted the 99.5th percentile of the combined loss distribution is 10x the severe loss due to the large gap between typical and severe loss estimates and the very skewed severity distribution resulting. The impact of this on aggregate operational risk capital will depend on the aggregation method:

- For variance-covariance matrix aggregation, a high 99.5th Financial Reporting loss would feed through into a higher 99.5th aggregate operational loss.
- For copula aggregation, however, the situation is nuanced as the contribution of Financial Reporting and other risks to the overall 99.5th percentile loss figure will be a lower percentile of the individual loss distribution. A highly skewed distribution may give rise to lower values for lower percentiles than a less skewed distribution.

  In this example, the 95th percentile of the combined loss distribution for Financial Reporting Risk is £2.4m, but if the typical loss was changed to £2m, the 95th percentile of the combined loss distribution would increase to c.£6.7m even though the 99.5th percentile is much lower (c.£70m v £200m with typical loss of £0.25m), so the highly skewed distribution could give rise to a lower aggregate requirement than a less skewed distribution.

This list of points to consider is by no means exhaustive, and the example may be simplistic, it but highlights some of the key considerations in validating scenario-based approach results.

Appendix C – Simple Bayesian Network model example

The following is a simple example of a Bayesian Network (BN) model looking only at process and fraud losses.

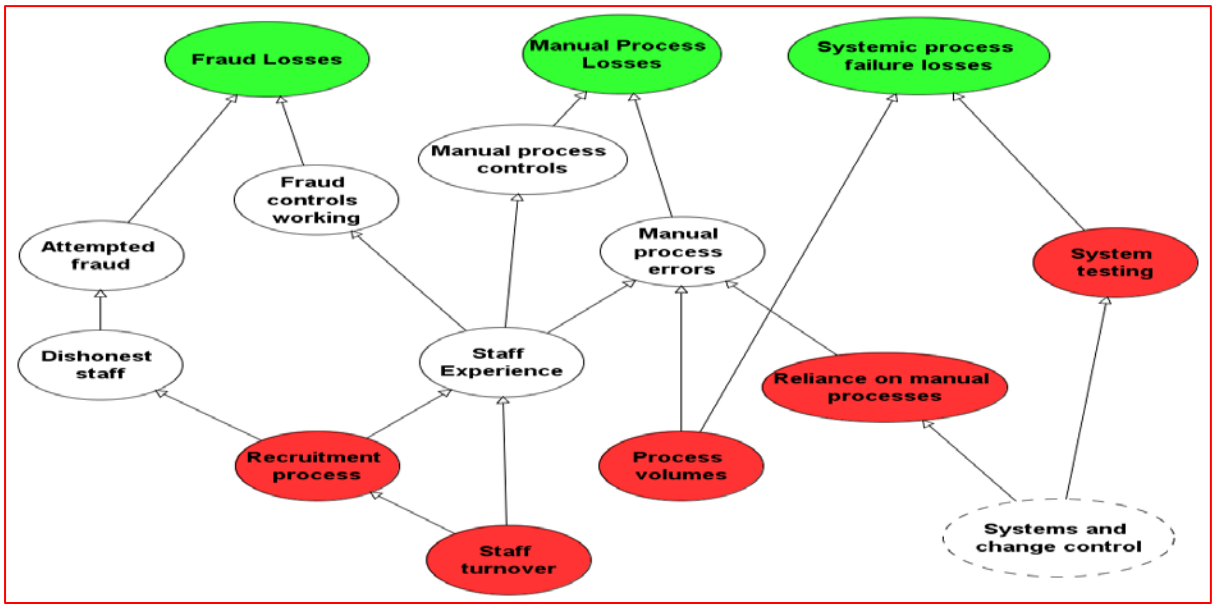
We may split process losses into system failures and manual errors. Both types of process loss will be affected by process volumes, but manual errors will also be affected by the:

- Level of reliance on manual processes which may vary by process…
- …and also the average experience levels of staff…
- … which will be a function of staff turnover and the quality of staff recruited.

Recruitment process failure may also lead to dishonest staff being hired and hence prevalence of fraud, with inexperienced staff less likely to pick this up.

These relationships are summarised in the following figure C.1 produced using AgenaRisk, a leading Bayesian Network modelling package[18]:

Figure C.1 – Causal Factor relationships



Appendix C.1 Bayesian Network Modelling

Bayesian Network modelling will first look to model the causal drivers above in red using a mixture of historic data, for instance on process volumes and staff turnover rates, as well as expert judgement. Note this could take the form of different states e.g. high / medium / low levels of turnover rather than precise values.

Modelling will then seek to derive conditional probabilities for fraud attempts and manual errors based on these variables, as well as the conditional probabilities that controls may fail to spot and rectify these.

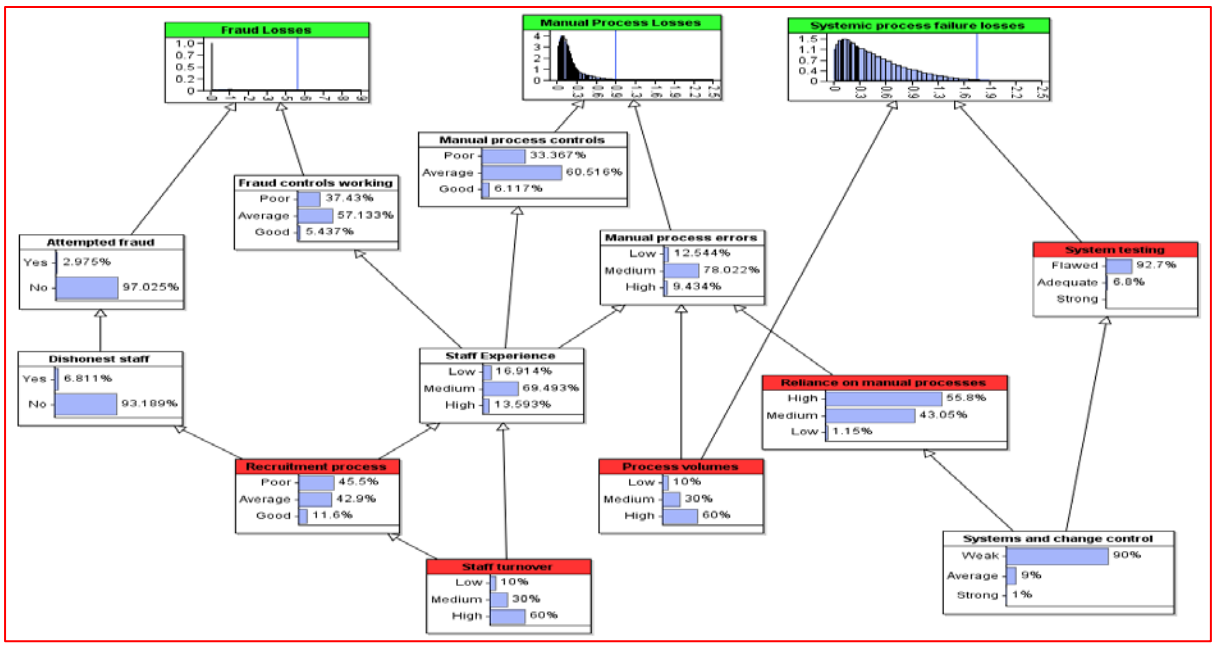
[18] With thanks to Neil Cantle, Bayesian Network SME at Milliman for this and figure C.2 below.

Like drivers, conditional probabilities may be expressed in terms of “states” e.g. error rates may be expressed as a function of high / medium / low levels of staff experience, and may be derived based on a mixture of MI and expert judgement. In this example, we may have to rely heavily on expert judgement to assess systemic process failures which tend to be low frequency, high impact events.

Based on modelling of causal drivers and conditional probabilities of loss events occurring based on these, we can model incidences of fraud and process losses, while loss severity may again be derived as a conditional function of causal drivers based on historic losses and expert judgement.

The following is an example of what the model output may look like using AgenaRisk.

Figure C.2 – Sample Bayesian Model output



Points to note:

- Firstly, we are modelling fraud and process risks together allowing for underlying drivers rather than modelling them separately through loss data or scenario analysis and then aggregating them in some way i.e. there is no need for a separate aggregation step.
- Secondly, modelling risks in this way requires lots of assumptions for conditional probabilities to translate modelling of causal drivers into controls failures and losses. For this reason, calibrating Bayesian Network is perceived to require more effort compared to more traditional modelling approaches which may be why so few firms use this approach at present.
- However, it may be possible to leverage BAU MI and data to help calibration, for instance 6-sigma scores of process failures could be used in parameterizing process error rates, so in practice the effort involved may not be that much more than traditional approaches.
- Finally, the process of identifying underlying causal drivers and seeking to link these to control failures and operational losses can yield some useful insights into operational losses and how these are inter-linked so the pain in terms of calibration may be outweighed by the gain in terms of understanding of operational risk exposure.

London

1-3 Staple Inn Hall · High Holborn · London · WC1V 7QH
Tel: +44 (0) 20 7632 2100 · Fax: +44 (0) 20 7632 2111

Edinburgh

Level 2 ·Exchange Crescent · 7 Conference Square · Edinburgh ·EH3 8RA
Tel: +44 (0) 131 240 1300 · Fax +44 (0) 131 240 1311

Oxford

1st Floor · Park Central · 40/41 Park End Street · Oxford · OX1 1JD
Tel: +44 (0) 1865 268 200 · Fax: +44 (0) 1865 268 211

Beijing

Level 14 · China World Office · No.1 Jianguomenwai Avenue · Chaoyang District · Beijing, China 100004
Tel: + +86 (10) 6535 0248

Hong Kong

1803 Tower One · Lippo Centre · 89 Queensway · Hong Kong
Tel: +11 (0) 852 2147 9418

Singapore

5 Shenton Way · UIC Building · #10-01 · Singapore · 068808
Tel: +65 8778 1784

www.actuaries.org.uk