# GIRO Working Party (2010 -2012)

# Practical Issues in the Solvency II Internal Model Approval Process (IMAP) for General Insurance Actuaries 08/2012

*By Jahan Anzsar, James Armstrong, Roger Austin, Adrian Bartliff, Chris Bird, Martin Cairns, Cherry Chan, Gabriela Chavez-Lopez, Andrew Dee, Gavin Dunkerley, Shane Latchman, David Menezes, Stephen Robertson-Dunn, Melinda Strudwick, Buu Truong*

WE ARE ACTIVELY SEEKING REVIEW BY INTERESTED PARTIES.  PLEASE SEND YOUR COMMENTS OR QUESTIONS TO:  giroimap@gmail.com

## Abstract

This paper sets out the results of the GIRO Internal Model Approval Process (IMAP) Working Party's surveys and interviews performed in 2010 and 2011, as well as views of Working Party members.

The focus of our work has been to identify practical issues facing the market under the broad categories of the "six tests" for internal model approval. The paper does not attempt to cover all aspects of Solvency II internal models, and is not intended as an introduction to Solvency II. We have however included a brief overview of the Solvency II internal model guidance as it stands at August 2012.

This paper is intended to be relevant both to practitioners currently involved in Internal Models, as well as actuaries who are less familiar with IMAP.

The issues covered in this paper are covered from the point of view of General Insurance actuaries in the UK, but may be of interest to actuaries in other fields.

### *Keywords*

### *Contact details*

Melinda Strudwick – melinda.strudwick@uk.pwc.com
James Armstrong - jamarmstrong@deloitte.co.uk
Stephen Robertson-Dunn – srdmail1@gmail.com

### *Reliance on this paper*

This paper is intended to be a springboard for discussion in the profession of some key areas of IMAP. It should not be relied on in discussions with the FSA or with Lloyd's.

The paper does not necessarily represent the views of the FSA, Lloyd's of London, GIRO, Institute of Actuaries or any other stakeholders in Solvency II. Where we have presented views of the authors on internal model issues, these represent the views of the individual authors rather than their employers, and are not necessarily the view of every Working Party member.

### *With thanks to*

# Table of Contents

# 1   Introduction

## 1.1   Approach

1.1.1   Solvency II is the new regulatory regime for the insurance industry likely to be implemented EEA wide on 1/1/2014, though further delays are possible.

1.1.2   The authors ('we', 'our') have undertaken two surveys and a number of interviews in 2010 and 2011 with both actuaries and non-actuaries working on preparation for Solvency II at insurance and reinsurance companies.

1.1.3   The questions that we included in our surveys and interviews were decided collaboratively by the working party members, and our emphasis was on practical issues that would be of use to those currently working on internal models.  The results of the 2011 survey are included in Appendix A.

1.1.4   The paper does not attempt to cover all aspects of Solvency II internal models, and is not intended as an introduction or comprehensive guide to Solvency II.  We have however included in Appendix B a brief overview of the Solvency II as it stands at August 2012.

1.1.5   The issues covered in this paper are covered from the point of view of General Insurance actuaries in the UK, but may be of interest to actuaries in Life or other fields.

1.1.6   Where we have presented views of the authors on internal model issues, these represent the views of the individual authors rather than their employers, and are not necessarily the view of every Working Party member.

1.1.7   Throughout this paper, we have referred to the Solvency II Level 1 and draft Level 2 texts. For the avoidance of doubt, the Level 1 text relates to the Solvency II Directive 2009/138/EC on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), and Level 2 refers to the draft implementing measures released by the European Commission for private consultation in October 2011.

1.1.8   Following on from the Solvency II Directive, in this paper, "insurer", "company" or "firm" is used to refer to "insurance and reinsurance undertakings".

## 1.2   Structure of the Paper

1.2.1   This paper is structured around the "six tests" for internal model approval; however the Statistical Quality Standards test is broken into two sections - one for Risk Ranking, the other for Expert Judgement.  There is also an additional section on External Models.

1.2.2   Further information on the six tests is included in Appendix B.

# 2 Executive Summary

## 2.1 Introduction

2.1.1    This paper sets out the views of the Working Party members, informed by the results of the GIRO Internal Model Approval Process (IMAP) Working Party's surveys and interviews performed in 2010 and 2011.

2.1.2    Under Solvency II, all firms are required to calculate a Solvency Capital Requirement (SCR), using either the 'standard formula' or their own internal model.  Approval of an internal model is the focus of this paper.

2.1.3    In this paper, we explore eight practical challenges for general insurance actuaries in meeting the tests and standards for approval of their internal model. We summarise industry practice in each area, where possible, and put forward our own views or examples of 'good practice' where industry practice is still evolving.

2.1.4    This paper is structured around the "six" tests for internal model approval i.e. Use, Statistical Quality, Calibration, Profit & Loss Attribution, Validation, Documentation (Articles 120 to 126) as they apply to a firm's Internal and External Models.

2.1.5    Overall, the paper explores a proportionate response (i.e. a practical response) to the legal requirements.  In assessing proportionality, we have considered what work will be useful to firms, both to modellers and users of the model.

2.1.6    A key thread running through the paper is the opportunity that exists for actuaries to meet the tests in a way that delivers value to the business, and to the decisions the model is intended to inform.

2.1.7    For example, the requirement for a model to 'rank risk' appropriately (i.e. 'quantify risk' appropriately) is fundamental to any good model.   In chapter 4, we provide examples of management information which enable the user and management to engage with and challenge the model's ability to rank risk.  Through this review and challenge process, the model evolves so that it is a more accurate reflection of the risks it is intended to quantify; and the underlying risks are better understood by management and users of the model.

2.1.8    It is important for the profession and regulators to engage in this debate so that a 'proportionate' bar for model approval emerges, and actuaries are able to continue to support the business in the way they have done for many years.

2.1.9    We would like to thank the many individuals who have contributed to this paper through what has been a very busy period for the profession, and particularly capital modelling actuaries.

## 2.2  Use Test

2.2.1    In its broader sense, it is reasonable to conclude that model use is the primary reason for internal model development.

2.2.2    Five years on from the publication of the Framework Directive, it is important in our view to be realistic about what models can and can't achieve.

2.2.3    Most certainly, models can't and shouldn't replace management's responsibility and judgement in decision making.  Models can, however, inform, challenge and provide insights into those decisions which would not necessarily have been there without it.

2.2.4    We summarise some strong examples of model use from the general insurance industry (e.g. SCOR).

2.2.5    We also highlight some lessons learned by actuaries working with more mature models, about how best to embed the model in decision making.  These include:

- identifying sponsors in the business,

- investing time in training and one-one coaching of model sponsors;

- transparency in the model methodology and output reports;

- sufficiently granular output, to answer key questions from the business;

- simplicity in the model, and speed of run-times;

- clear documentation of the model, to facilitate wider review & challenge; and

- identifying 'early wins' (e.g. reinsurance restructuring) where the model has generated profit for the business.

2.2.6    Our view is that there is typically a three year lead time from development of a model and initial discussion of outputs with model users, before the model carries its full weight in decision making.  Given the immaturity of many Solvency II models in the market, this raises an interesting challenge for firms in meeting the use test at the point of entry to IMAP.

2.2.7    We discuss a realistic target for use of a general insurance model early on in the model's development, and thus at the point of IMAP entry.   For a smaller or less complex firm, we suggest that it's realistic to target use of the model in capital assessment and in quantifying risk exposures.  For larger or more complex firms, we would also expect the model to be used in setting risk tolerances (e.g. market risk as a percentage of net asset value at a 1 in 200 year level), and monitoring risk exposures against those limits.  For all firms, we would expect the model to inform reinsurance purchase and other ad-hoc strategic decisions (e.g. M&A).

2.2.8    We also summarise possible approaches to demonstrating compliance with the use test, and present our survey results with regard to firms' planned uses of the model under Solvency II.

## 2.3    Statistical Quality Standards – Risk Ranking

2.3.1    Within Statistical Quality Standards, our focus was on Risk Ranking.  The reason for this is that the internal model's main purpose is to quantify risk, and a review of the model's risk ranking ability ensures that this is done appropriately.

2.3.2    The ability of the model to rank risk can be assessed using both quantitative and qualitative approaches; we look at each in turn using worked examples.

2.3.3    More generally, we conclude that risk ranking works best if a basket of metrics and approaches are tested (at different points in the distribution) as this reduces over-reliance on any one approach.  Different metrics may be better targeted to different uses or users of the model.

2.3.4    Communication of results is key to enabling management to understand the model and challenge the model's risk ranking ability; and to informing their view of risk.

## 2.4    Statistical Quality Standards – Expert Judgement

2.4.1    Use of expert judgement is a key part of the development and subsequent use of any internal model under Solvency II.  It is likely to be applied in the selection of methods, assumptions and analysis of model results, particularly where credible data does not exist to support the selection.

2.4.2    In the wake of the financial crisis, CEIOPS originally set a very high hurdle for the use and validation expert judgement.  Whilst the Level 2 requirements are more proportionate the key principles remain:

- Methods and assumptions should be based on the expertise of persons with the appropriate knowledge of insurance risks.

- Key assumptions (or judgements) should be justified, and their limitations should be transparent to management and users of the model.

- Management and model users should be made aware of the degree of reliance on expert judgements.

- There should be independent validation of material assumptions in the model.

2.4.3    The challenge for firms is therefore to improve the governance and reporting of expert judgements as they are applied within the internal model.

2.4.4    We discuss how to demonstrate the suitability of an expert, including:  relevant experience and qualifications; the relationship of the expert to the model and users of the model results; training/briefing received by the expert on the internal model and the governance process.

2.4.5    We also give some thoughts and practical examples of approaches to governance, validation and reporting of expert judgements as they are applied in the internal model.   We provide a worked example of the expert judgements applied in pricing credit default obligations prior to the financial crisis.

## 2.5   Calibration

2.5.1    This section of the paper focuses on the issue of moving from the typical "to ultimate" measurement basis of liabilities, adopted under the ICA regime, to the measurement of liabilities over a 1 year time horizon as required for the SCR.  Many undertakings we spoke to considered the new basis purely as a regulatory requirement, and not useful for other purposes.

2.5.2    We discuss each of the Calibration methods in turn and the respective advantages and disadvantages are mentioned. These include: proportionate emergence, actuary-in-the-box, perfect foresight, Merz- Wüthrich, and hindsight re-estimation.

2.5.3    Empirical testing conducted by the working party in 2010 (see Appendix G) against historic FSA returns, suggested that no one method was a better predictor of the one-year movement than another.  With this in mind, it may make sense for companies to use methods which are easiest for them to implement, or which facilitate management's review.

2.5.4     We note that all these methods are inherently limited in that they do not allow for drivers of reserving uncertainty outside the sample data set being used to parameterise the method, including any tail factor development.  We highlight these methods should be applied with caution, particularly in exercising judgement in reviewing the parameter selections and the result.

2.5.5    In our experience, it is useful for firms to perform a comparison between ICA and SCR, from old to new model.   It should be possible to explain the change in basis, at least at a high-level, to senior management; even where the model has changed.

## 2.6   Profit and Loss Attribution

2.6.1    We note that profit and loss attribution as defined by Article 123, and as interpreted by Lloyd's, primarily relates to the identification of sources and causes of historic profits and losses.  Losses are then compared to the model's structure, to confirm that all risk drivers are captured within the internal model.  For example, in hindsight, exposures to Thai Floods in 2010 were not captured by some insurers in their catastrophe modelling.

2.6.2    The use of the internal model results in validation of assumptions (back-testing) and in business planning are, strictly speaking, covered by the validation and use tests respectively.

2.6.3    Firms we spoke to who make use of their model within the business planning and capital allocation processes, felt this definition of profit and loss attribution was too limited.  The main purpose of the internal model was in their view to inform management's understanding of past and future variability in profits and losses, and the associated return on capital.  A key function of the model is therefore to assist in explaining these drivers of profit and loss, to internal and external stakeholders (e.g. rating agencies, shareholders and management).

2.6.4    It's our view, therefore, that the assessment of the model's ability to model the drivers of profits and losses should be both forward and backward-looking.  In particular, the use of the model in business planning and validation are in our view key to meeting this test.

2.6.5    We discuss practical challenges in performing profit and loss attribution, possible solutions, and look at a worked example of the distribution of profit and losses output from the internal model.

## 2.7  Validation

2.7.1    Solvency II requires all firms to have a regular cycle of validation in place.  The firm's approach to validation should be approved by the Board in the validation policy, which forms a core part of the IMAP submission.

2.7.2    Validation is a difficult area of the Directive for firms to respond to.  As well as being critical to approval of a firm's IMAP submission, and subject to Board and Regulator scrutiny, it demands expert input (e.g. actuarial, underwriting, treasury, etc.) and therefore is highly resource intensive. Firms responding to our survey said they expected validation to be 2 to 4 x as extensive under Solvency II as under ICA, in year 1; and 1.5 to 3 x as extensive on an ongoing basis.

2.7.3    In particular, we focus on the key differences between validation under ICA and under Solvency II:

- How to define the purpose and scope of the validation.

- How to apply the concept of materiality within the validation work, in order to deliver value to users of the model.

- Which validation tools are required.

- How to assign roles and responsibilities for validation i.e. governance framework, including an appropriate amount of independent review.

- How to formalise the technical analysis - for example, is it possible to define pass/fail criteria for validation tools in a practical way?

2.7.4    We discuss what it means for an internal model to be 'appropriate' (as required by Art. 124) in the context of sensitivities and uncertainty in the modelled result.  This definition sets the bar which the validation should clear.

2.7.5    Amongst other things, we mention: that the SCR should reflect all available information about the 1 in 200 year position, including historic data, market data, and information about the firm's risk; that methods, assumptions and results should be justified; that the SCR is estimated on a consistent basis year-on-year; that the model meets the tests & standards for model approval; and finally that in the view of an "independent knowledgeable third party", the SCR is estimated using methods and assumptions that are reasonable.

2.7.6    We summarise the tools required by the Level 1, 2 and Lloyd's guidance.  We also include worked examples of validation tools in action, including back-testing, expert review and scenario testing.

2.7.7    We note that by applying validation in a proportionate way, firms we spoke to saw value in the validation exercise, including increasing understanding of the model and the results amongst the wider management team, reducing model error, and confirming to the Board that the model was ready for IMAP submission.

## 2.8  Documentation Standards

2.8.1    Documentation is a key channel of communication about the internal model to the regulator, senior management, Board, risk management, and other users of the model.  It will form the basis of the supervisory authority's views, as well as the views of the independent validator, on the appropriateness and reliability of the internal model.

2.8.2    We note the internal model is not defined by Solvency II, but extends beyond the 'calculation kernel' (or the engine of the capital calculation).  We assume that documentation should cover all aspects of the internal model, including all inputs and outputs of the model, underlying models or analysis quantifying risk (e.g. quantification of operational risk within the risk register), as well as the policies and procedures governing its use.

2.8.3    The Level 2 text requires firms to produce sufficient documentation for "any independent knowledgeable third party ... to understand the design and operational details of the internal model" and "Outputs of the internal model shall in principle be reproducible using the internal model documentation and all of the inputs into the internal model".

2.8.4    It is particularly challenging for firms to document the extensive technical details of the internal model when the level of detail and granularity required by the supervisor is unclear.

2.8.5    For practical purposes, we have assumed that an "independent, knowledgeable third party" is likely to be a financial modeller with the appropriate experience (i.e. experience of modelling insurance liabilities) and qualifications (i.e. mathematical or statistical university degree, but not necessarily a qualified actuary) who is independent of the design and build of the internal model.

2.8.6    We identify some key attributes of 'good' documentation in the views of such "knowledgeable third parties".  These included a clear explanation of model purpose and use, a plain English description of methodology used, flowcharts of model structure and processes, justification of methods and assumptions used, summary of model results and summary of validation results (including sensitivity testing, stress & scenario testing).

2.8.7    It was felt that detailed mathematical proofs were not helpful to an independent reviewer.

2.8.8    We also note that documentation alone is not sufficient for a successful, independent review of the model.   Individuals we spoke to highlighted that access to the model (i.e. ability to inspect the model components and parameterisation spreadsheets), ability to run tests on the model, or validate assumptions or results, and an audit trail were all important to their review.

2.8.9    We believe that creating a good 'culture' of documentation, i.e. enable a third party to understand the model, within a firm is vital if it intends to seek internal model approval.  If documentation is largely seen as a compliance exercise, it is unlikely to work in a sustainable way for ongoing approval.

2.8.10   We also discuss potential software solutions and include a possible framework or checklist for Solvency II documentation.

## 2.9    External Models - Catastrophe Modelling

2.9.1    Companies often use external models, provided by third parties, within their Internal Models. In this section, we consider catastrophe models as an example of an external model.  The main catastrophe models used by firms are: RMS, AIR and EQECAT.  Their approach is different and each one has both strengths and weaknesses.

2.9.2    External models need to satisfy the model tests set out in articles 120 to 125; that is, the Use Test, Statistical Quality Standards, Calibration Standards, Profit and Loss Attribution, Validation Standards and Documentation Standards.

2.9.3    This is a particularly challenging area for firms, given that the catastrophe model vendor has control over the methods and assumptions used within the model, one that contains proprietary information.  Detailed knowledge of the external model, including key assumptions and limitations (such as exposure data quality, switches, modelled and non-modelled exposures) may be limited to the catastrophe modelling team.

2.9.4    Management's review of catastrophe models and their use has focussed historically on aggregation monitoring and pricing.  As a working party, we have seen more focus on the use of catastrophe models in capital modelling in the last few months.

2.9.5    Recent publications released by the ABI and Lloyd's around the use of catastrophe models have substantially aided firms in determining how they respond to the Solvency II requirements.

2.9.6    We do not seek to repeat the guidance issued here, but instead explore some of the practical challenges identified by our discussions with firms, and give some examples to help firms implement their recommendations. We focus on documentation and validation, and in particular the balance between documentation/validation evidence available to firms under their license agreement, and documentation/validation that is produced by firms.  We also discuss use and governance of catastrophe models under Solvency II.

## 2.10 Conclusions

2.10.1   The main objective of this paper is to provide practical examples on the IMAP process to enable other industry professionals to understand the Directive and requirements behind it.

2.10.2   With this objective in mind, we are actively seeking feedback on this paper by interested parties.  Please send your comments or questions to: [giroimap@gmail.com](giroimap@gmail.com)

2.10.3   As a working party, we have made an effort to discuss and consider specific challenges faced by Insurance Companies during the process of meeting the IMAP requirements in order to provide a clearer idea of the practical implementation challenges of the Directive requirements.

2.10.4   It is worth noting that each element of the IMAP process can be considered in terms of its usefulness to the overall modelling goals of the insurer, and materiality to the modelled result.   Throughout this paper, we have found it helpful to consider decisions on the amount of work needed to comply with the standards in this light.

2.10.5   The final interpretation of the requirements by the UK FSA and other European Regulators will of course only emerge over time, once the Directive is transposed into national law and internal models start to be formally approved.

2.10.6   We acknowledge that there are a wide range of areas within IMAP that would be worthy of further investigation!

# 3 Use Test

## 3.1 Introduction

3.1.1    In its broader sense, it is reasonable to conclude that model use is the primary reason for internal model development within an insurance or reinsurance undertaking.

3.1.2    Demonstrating this use is a formal regulatory requirement under the Solvency II Directive[1] for firms seeking internal model approval. In particular the 'Use Test' challenges firms to evidence that the model is widely used in the firm's governance, risk management and decision making.  Particular uses include:

- Business Planning[2], including profit & loss attribution[3]

- Risk Management[4], including capital setting, and risk mitigation/reinsurance purchase.

- Strategic decisions[5]

3.1.3    The principle of the Use Test is such that it encourages companies to find ways in which the model can add value to the business.  The model's use within the business then becomes a natural part of the business process with future iterations being facilitated by improvements to the internal model itself.

3.1.4    It is important to acknowledge that different firms and models are at very different points on a maturity curve and as a result the Use test could mean different things for different firms.  Firms should make a realistic assessment of the level of development of their model and tailor their use targets accordingly. To help internal stakeholders understand and use the model fully is a long process often referred to as embedding. Firms should expect for it to take time to reach their ultimate use goals.

3.1.5    It is also important in our view to be realistic about model limitations and circumstances where the model doesn't work. Every model has limitations and no model can be expected to cover all risks.  Hence, a good understanding of the design, scope and principals underlying the model will help stakeholders interpret key outputs from the model for decision making purposes.

---

[1] Article 120, Council Directive (EC) 2009/138/EC 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II)
[2] Article 45, Own Risk & Solvency Assessment
[3] Level 3
[4] Article 120, Use
[5] Article 45, Own Risk & Solvency Assessment

3.1.6    Models can, however, inform, challenge and provide insights into those decisions which would not necessarily have been there without it.  Capital figures should be presented to management in a way that aids them in decision making, but judgement is still paramount.

## 3.2   Summary of Framework Directive

3.2.1    The Use Test is described in Article 120 of the Framework Directive.  In particular, firms are required to evidence use of the model in:

- "their risk-management system ... and their decision-making processes;

- their economic and solvency capital assessment and allocation processes".

3.2.2    These concepts help to meet one of the over-arching principles of the Directive: that of policyholder protection.  The Directive also requires management to show that they use the model in their own economic decision making with possible examples including performance monitoring and business planning.

3.2.3    The key requirements of Article 120 are outlined further in the Level 2 text:

- to be able explain the uses of the internal model in decision making, including consistency of outputs for different uses

- for the internal model to reflect the business structure ('fit to business')

- to be able to demonstrate senior management understanding of the internal model.

- to be able to demonstrate discussion of the internal model and its results by senior management

- for the internal model to "support" decision-making.

- for the internal model to be integrated into the risk management framework

- that the model reflects changes to the risk management system

3.2.4    In addition, Level 2 requires insurers to "use the internal model in their risk-management system and in their decision-making processes in a way that creates incentives to improve the quality of the internal model itself".

## 3.3   Key Practical Challenges

3.3.1    Our group focussed on the challenges associated with meeting the Use Test from a practitioner's perspective.  In this regard, there were clearly three themes facing actuaries working towards successful model use:

- The first was to understand what the Use Test actually means for firms with models at different maturities.

- The second was centred on what firms should be doing to ensure that business users are sufficiently bought into the internal model and are motivated for it to succeed.

- The third was ensuring that firm was capable of demonstrating compliance with the Directive.

## 3.4 Practical Solutions

*The Maturity Curve and the Use Test*

3.4.1    The requirements of the Use Test are such that firms must be able to demonstrate that the internal model is widely used in their system of governance.

3.4.2    However it is reasonable to assume that the level of reliance a firm will place on its internal model is largely dependent on how mature that model is.

3.4.3    In 2007, the FSA's insurance sector briefing spoke about a five year development period for designing, building and embedding an ICAS model[6].  In reality, where firms have introduced new models for Solvency II, against this timescale these models are only now in year 1 or 2 of development:

| Year | Activity | Resources |
|------|----------|-----------|
| Year 1 | Build first take model | Capital Team |
| Year 2 | License a dedicated model platform | Support by finance, actuarial, reinsurance functions |
| Year 3 | Improve model calibration<br><br>Apply more widely in the business | Involvement of the CRO/Finance/underwriting/pricing claims/actuarial/reinsurance/treasury/ risks & compliance |
| Year 4 | Ensure wide usage in planning, pricing, reserving, capital allocation, internal and external reporting | Embedded in all key business functions under CRO control, with resource implications |
| Year 5 | Objective is for it to be embedded to an acceptable level<br><br>Further development and calibration will continue as required | Embedded in all key business functions under CRO control with resource implications throughout |

3.4.4    Thus a key challenge for firms facing the Use Test is therefore deciding on what is an appropriate level of use given the development of their model. There is a danger that firms either attempt an exaggerated level of use, which ultimately lacks credibility given the level of development of their model, or conversely the extent of use of the model is below what might be expected given the maturity of the model.

---

[6]  http://www.fsa.gov.uk/pubs/other/icas_isb.pdf

3.4.5    In this light, we present in the diagram below an example of what the maturity – use trade-off may look like. We have identified possible appropriate types of use for a given level of maturity. Please note that this is a generalisation and specific firms could well follow a different progression than that outlined here.

**1 -2 years: Challenge**

In the early stages of the model, focus would be directed to ensure appropriate calibration of the model. SCR figures could be compared against ICA figures and an analysis of change could be conducted to contrast the two capital figures. Results inform firm's view of risk, and risk management decisions (especially reinsurance purchase)

**3-5 years: Practical Decision Making**

As the model matures, the internal model would become more prominent in setting and monitoring risk appetite, setting performance targets and monitoring performance (incl P&L attribution), as well as business planning and risk management decisions.

**How the model influences decisions**

**Number of years of using the model**

3.4.6    The following sections highlight examples of likely model use at **lower and higher maturity levels based** on our personal views of the general insurance industry:

| Likely levels of use | Immature model i.e. 1-2 years after design | Mature model i.e. 3-5 years after design |
|---|---|---|
| **Smaller / less complex firm** | Assessment of capital requirements for business plan and risk profile* (i.e. SCR) | Business planning, monitoring level of profits/losses against expectations as modelled |
| | Basis for projection of capital requirements* over business planning time horizon (say 1- 3 years) | (i.e. use of model in profit & loss attribution) |
| | Purchase of reinsurance | |
| | Quantification of risk and prioritisation of risk exposures | |
| | Ad-hoc strategic decisions (e.g. M&A) | |
| **Larger / more complex firm** | As above, plus: | As above, plus: |
| | Use of model in setting and monitoring risk against target risk appetite | Setting performance targets for lines of business, capital allocation |
| | (e.g. market risk as % of net asset value | Identifying more efficient uses of capital, e.g. via |

| Likely levels of use | Immature model i.e. 1-2 years after design | Mature model i.e. 3-5 years after design |
|---|---|---|
| | at 1 in 200 year;  aggregate annual catastrophe as % of net asset value at 1 in 200 year) | portfolio 'optimisation' Setting and monitoring investment strategy |

\* included in the firm's own risk and solvency assessment (ORSA)

> 3.4.7 The following sections highlight three **specific examples of model use at higher maturity** levels based on publically available data[7].

**Business planning**

> 3.4.8 A good example of use of the internal model for business planning has been exhibited by SCOR[8].  SCOR have integrated capital and risk considerations into business planning through the creation of a Capital Deployment Tool which they have named CaDeT.  CaDeT provides real time information on the impact of exposure (or business plan) changes on capital requirements and diversification benefits.

> 3.4.9 Since its creation, SCOR have used CaDeT as a tool for strategic decision making, looking at portfolio composition and potential business acquisitions. It has also been used to cement understanding of risk and capital across the organisation due to its accessibility. The pragmatism of using a simplified and therefore more accessible view of the internal model is a great example of how organisations can get non-technical parts of their company involved and make capital relevant.

The image below shows an example of how CaDeT can be used to examine how changes in exposure will affect the standalone and diversified capital.

---

[7] We are seeking permission from the relevant firms for referring to this information : confirmed by Andrew Hitchcox and Andrew Dee; awaiting response from SCOR.

[8] This was taken from Scor, November 2010, Enterprise Risk Management (ERM): A risk-based approach to the management of a (re)insurance company, Focus series.

## Fig. 50: CaDeT example (Simplified example)

| Risk Factor | Exposure | | | Standalone Capital | | | Diversified Capital | | | Diversification | | | Risk Monitor |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Current | Change | New | Current | New | Change | Current | New | Change | 2009 | Plan | | |
| P&C New Business | 1,500 | 50% | 2,250 | 800 | 1,100 | 38% | 415 | 700 | 69% | 48% | 36% | ↘ | ● |
| P&C Reserves | 5,000 | | 5,000 | 800 | 800 | 0% | 625 | 600 | -4% | 22% | 25% | ↗ | ● |
| Life | 1,300 | | 1,300 | 850 | 850 | 0% | 450 | 415 | -8% | 47% | 51% | ↗ | ● |
| Investments | 5,000 | | 5,000 | 350 | 350 | 0% | 65 | 55 | -15% | 81% | 84% | ↗ | ● |
| Other | 2,000 | | 2,100 | 150 | 150 | 0% | 30 | 25 | -17% | Click to close the zoom | | ↗ | ● |
| Total | | | | | | | 1,585 | 1,795 | 13% | | | | |

Changes in exposure (different granularities available)

Are assessed real-time in terms of risk and capital

### Available & Required Capital - Internal Model

**Available Capital**
2,000 (Current)  2,100 (New)  100 ↗

**Required Capital**
Diversified (net) capital
1,615 (Current)  1,825 (New)  210 ↗

**Solvency Ratio**
= Available/Required
124% (2009)  115% (Plan)  -9% ↘

Solvency monitor: ●

**Reinsurance strategy**

3.4.10   A good example of embedding the internal model use into decision making can be seen in Legal and General Insurance (LGI) when considering their reinsurance purchase.  LGI have used the internal model to analyse the effects of various reinsurance strategies for the last two years.

3.4.11   The model is used as above by LGI to test the varying levels of losses that may be experienced under each treaty.

3.4.12   When the various reinsurance programmes are being considered by LGI the capital implications as well as the potential losses under each strategy are presented to the board alongside the expected cost. This allows the risk appetite and capital considerations of LGI to be embedded within the reinsurance purchase as well as a greater level of understanding of the benefits gained from each reinsurance structure.

**Ad hoc use**

3.4.13   As discussed at GIRO in 2010[9], Andrew Hitchcox demonstrated use of the model within the decision making process in a number of examples, these include:

3.4.14   A large scale acquisition which lead to a significant increase in wind and flood business. This required a decision to be made as to which vehicle should be written into and therefore a good understanding of the capital requirements and return on equity considerations. A multi-month exercise was undertaken to adjust models and analyse the outcomes before a decision was made.

3.4.15   A decision was needed on a significant syndicate merger so a project to understand the likely effect of merging the two syndicates was undertaken. It was vital that the capital implications could be understood properly as both of the capital providers needed to be persuaded.

3.4.16   A rearrangement of the risk profile of a syndicate.  An excess of loss reinsurance programme of significant size had to be modelled and its affect on the capital requirements and return on equity analysed and understood.

### *Getting buy-in to the model*

3.4.17   We spoke to firms with more mature models about their experiences of getting management in buy-in to and use the internal model in decision making over the past 5-10 years.

3.4.18   Lessons learned include:

- identifying sponsors in the business,

- investing time in training and one-one coaching of model sponsors;

- transparency in the model methodology and output reports;

- sufficiently granular output, to answer key questions from the business;

- simplicity in the model, and speed of run-times;

- clear documentation of the model, to facilitate wider review & challenge; and

- identifying 'early wins' (e.g. reinsurance restructuring) where the model has generated profit for the business.

3.4.19   It is important for the model to generate the basis and granularity of results required by management.  As the examples below show, it can be seen that even small differences can hamper understanding, and this can make embedding of the model more difficult:

---

[9] http://www.actuaries.org.uk/research-and-resources/documents/plenary-5-beyond-2012-embedding-internal-model

- From a business management perspective, the 1 in 200 SCR may not be the most important measure. Some management teams may have their own calibration such as a 1 in 100 TVaR. Others may be more focused on Earnings at Risk.

- The granularity of the model should align to underwriting teams for the figures to make the most sense to product line heads. Where they differ, it is difficult for management teams to understand and challenge figures.

- Consistency of definitions between the internal model and other processes such as business planning, pricing and reserving can undermine the value of modelled output. For example, it is possible for a large loss threshold to differ between a pricing model, a reserving definition and a capital model.

3.4.20    'Early wins' can be helpful to gain credibility and buy-in to the model. For example:

| Stakeholder | Benefit |
|---|---|
| Risk Management | *Capital management, ORSA, risk appetite setting / monitoring, reinsurance, and investment decisions* |
| Actuaries | *Reserve uncertainty, pricing* |
| Board members | *Business planning and strategy, development of risk strategies and understanding material risks* |
| Underwriters | *Setting and monitoring performance of business units and classes against targets*<br>*Identifying opportunities for more efficient use of capital (e.g. portfolio optimisation)*<br>*Understand underlying risk profile of the business* |
| Reinsurance | *Facilitate better and more efficient reinsurance arrangements* |
| CEO | *Use model to explain variability in key drivers such as profits and losses, and the associated return on capital, in current and future years* |

3.4.21    It was also noted that simpler models are easier to embed, easier to explain, and faster to run.

## Demonstrating compliance with the Use Test

**Forms of use**

3.4.22     Participants in our e-survey were asked "How do you intend to use the model?" The results below highlight some of the key expected uses including capital management, input into ORSA and risk management, reinsurance, and integrating capital modelling with business planning. It should be noted that the survey results are likely to be underpinned by significant differences in use between smaller and larger firms.

**Percentage of Survey Respondents Adopting Different Types of Use**

| Type of Use | Percentage |
| --- | --- |
| Capital Management | 97% |
| ORSA | 90% |
| Reinsurance | 90% |
| Business planning/strategy | 87% |
| Development and monitoring of risk appetitie | 81% |
| Measurement of material risks | 71% |
| Development of risk strategies | 52% |
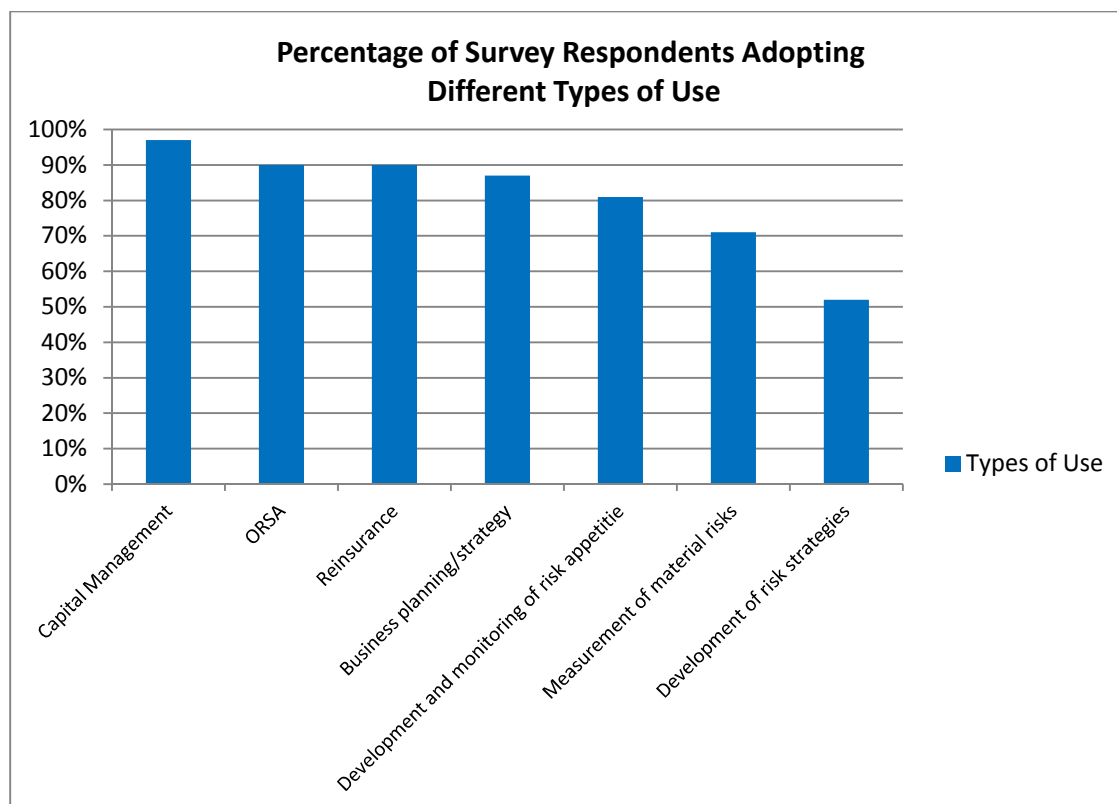
3.4.23     Given the ICA regime background of most UK companies applying for Internal Model approval, it is likely that they will be relatively advanced with regards to using the Internal Model to aid capital management.  This is the most obvious model use and was most popular in our survey.  Allocation of capital was less common.

3.4.24     ORSA is an important use and really picks up integration of the model with the risk management framework.  Measurement of material risks and development of risk strategies could be considered as parts of the ORSA.  There is certainly a regulatory need and management desire to focus on using the model in the risk space.  The challenge that most companies face is to start to use the internal model to set and monitor quantitative risk tolerances or limits, in what has historically been a relatively qualitative world of risk registers and risk profiles.

3.4.25   Reinsurance purchasing is often cited as a use.  Many firms have detailed reinsurance programmes in their internal models and are able to produce detailed granular outputs.  Use in this area will be developed by improving the calibration of large losses and creating a level of consistency between reinsurance broker models, pricing models and the internal model. Credibility will be needed prior to reinsurance buyers accepting the internal model as a reasonable alternative or addition to specific exercises already conducted.

3.4.26   Other uses which were not as wide-spread across e-survey participants are shown in the graph below.



**Percentage of Survey Respondents Adopting Different Types of Use**
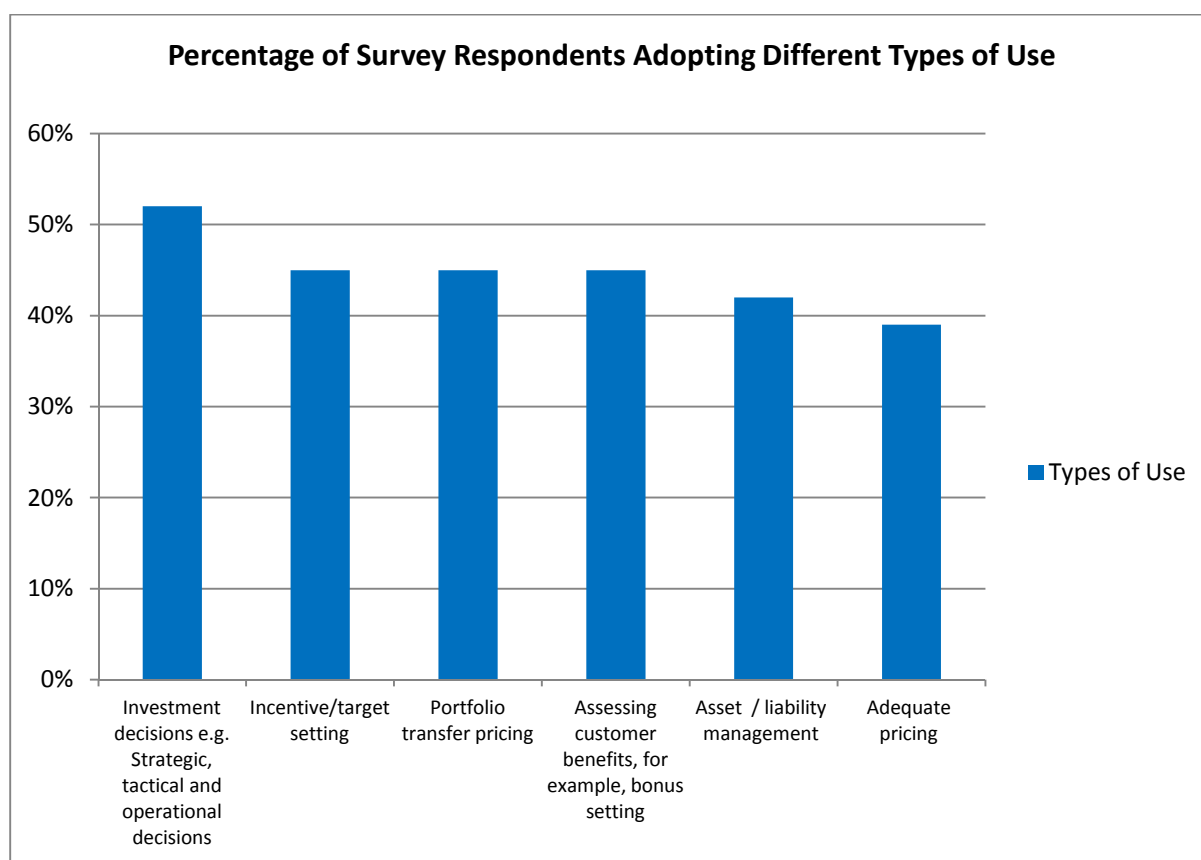
**Evidence and documentation**

3.4.27   Questions G.1 and G.2 of the FSA's Self-Assessment Template[10] cover the key aspects of how firms must demonstrate they are meeting the requirements of the Use Test, including the responsibility of the Board to ensure that the model is appropriate to the risk profile of the firm, and that it is used in the firm's decision making processes, solvency calculations and capital allocation.

3.4.28   Some examples of types of evidence are tabulated below.

---

[10] FSA, February 2012, Internal Model: Self Assessment Template

| Use Test Requirement | Evidence |
|---|---|
| Model embedded in organisation | Policy / process documentation (including flowcharts) |
| | Roles, responsibilities, terms of reference for committees |
| | Management Information |
| Users understand model | Training – particularly of Board and Senior Management |
| | Evidence of feedback from model users to the modelling team (e.g. notes/actions from meetings) |
| | Interviews with management |
| Model produces appropriate output | Process documentation |
| | Management Information |
| | Validation results |
| | Evidence of feedback from model users to the modelling team (e.g. notes/actions from meetings, model change requests) |
| Output influences key business decisions | Management Information |
| | Meeting minutes |
| | Interviews with management |

3.4.29   The amount of evidence expected to be required to pass the Use Test ranged among those we interviewed from 50 pages to 2,000 pages.   In our view, 2,000 pages are excessive!

3.4.30   Some interviewees indicated they have used the IMAP as an opportunity to revamp the management information that gets provided to the Board (e.g. graphical representations linking to the Probability Distribution Forecast (PDF) rather than just a single statistic).

3.4.31   A problem a number of companies may encounter is how to document examples of use and embedding, in a way that does not slow down decision making.

3.4.32   One approach we have seen is to use a tiered system, whereby key decisions are documented in great detail (usually with accompanying reports or papers) but smaller projects or ad hoc use may just have email-logs and calculation results as documentation.

3.4.33    All evidence is then held in a central document, indicating how the model was used, who used it and if the use indicated any potential model weaknesses. Examples of these central documents can be seen in the images below.

## LEVEL 1 EVIDENCE OF USE TEST

### REGULAR DECISION MAKING EXAMPLES

| No. | Use | Description | Frequency | Documentation | User / Committee / Board | Relevant part of risk distribution *key below | Was model successfully used? | Any identified limitation of the model | Used for: a. RM system b. Decision making process c. Capital assessment and allocation |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Setting regulatory Capital (ICA) | Annual capital assessment for Lloyd's | Annually | Board paper July (SBF submission) Board paper September (Updated numbers) | RJK Board, Audit and Risk Committee (ARC) | Tail | Y | N | C |
| 2 | Assessing economic capital | Basis of capital assessment for TM Head Office and Kiln Board | Annually | RJK Board paper (Sept/Dec) Head office standard form report | RJK Board and HO | Tail | Y | N | C |
| 3 | Capital buffer | Business planning tool for possible mid year plan change. Used for Kiln management purposes. | Annually | RJK Board paper (Sept) | RJK Board, Audit and Risk Committee | Tail | Y | N | C |
| 4 | Raising capital from capital providers | Ad-hoc activity on a needs basis | Ad-hoc | RJK Board paper | RJK Board, Audit and Risk Committee | Full dist | Y | N | B |
| 5 | Approval of business plans | A proposal on business planning requiring Board sign-off | Annually | RJK Board paper (Sept) | RJK Board, Underwriting Board | Full dist | Y | N | B |
| 6 | Capital management | Ongoing process of assessing business actually written vs carried capital | Quarterly | RJK Audit & Risk paper (March, May, Sept) | Audit & Risk Committee | Tail | Y | N | A |
| 7 | Three year mid-term business plans | Annual (rolling) activity for TM HO | Annually | Not yet completed. Due Oct 2011. | RJK Board, Audit and Risk Committee | Full dist | Y | N | A |
| 8 | ORSA | Annual ORSA report with quarterly updates | Quarterly | RJK Board (Dec) RJK ARC (March, May, Sept) | RJK Board, RJK Audit & Risk Committee | Full dist | Y | N | A |
| 9 | Evaluation of risk appetite | Annual proposal on risk appetite to be signed off by the ARC | Annually | RJK Audit & Risk paper (Dec) | Audit & Risk Committee | Full dist | Y | N | A |
| 10 | Capital allocation | Review of the methodology/ies applied and the implications of the results | Annually | UMC paper (Due Oct/Nov 2011) | RJK Audit and Risk Committee | Tail/Full dist | Y | N | C |
| 11 | Reinsurance decisions | Evaluation of RI purchases Review of Reinsurance programmes (design) | Annually | RSG paper (Sept) | - Reinsurance strategy committee - Underwriters | Full dist | Y | Y | B |
| 12 | Six-monthly reporting to TMHO | To help TMHO measure their overall Group exposure and capital responsibilities | Six-monthly | Standard format HO report (TG net) (March & Oct) | KGL Board | Full dist | Y | N | A |

## LEVEL 2 EVIDENCE OF USE TEST

### SPECIFIC USES (in date order)

| No. | Use | Description | Frequency | Documentation | User / Committee / Board | Relevant part of risk distribution *key below | Was model successfully used? | Any identified limitation of the model | Used for: a. RM system b. Decision making process c. Capital assessment and allocation |
|---|---|---|---|---|---|---|---|---|---|
| 1 | New reinsurance programme for S1880 (risk appetite of 1880 and balance with 510) | 1880 reinsurance purchase: The model was used to gauge the impact on capital of a brand new reinsurance programme for S1880. Various structures are being considered and modelled and are expected to verify the decision to buy an XL programme. | Jul/Aug '11 | KGL Board paper (Sept 2011) | RJK Board, Audit and Risk Committee | Full dist | Y | N | B |
| 2 | Testing efficacy of reinsurance programmes | Quantitative testing of the effectiveness of Kiln Syndicate/Division's reinsurance arrangements through: -Historical performance. -Realistic Disaster Scenarios -Performance against Capital Model | Jul/Aug '11 | RSG paper (11 Aug 2011) | Underwriting Board | Tail/Full dist | Y | Y | A |
| 3 | Model changes | RMS version 11 upgrade | Apr– Jul '11 | - RJK ARC paper (March and Sept 2011) - Group Stats project folder | RJK Board, Audit and Risk Committee | Tail/Full dist | Y | N | C |
| 4 | Writing extra business | A&H active underwriter deciding to write Fac business where he could no longer exclude the NBC componet. | Jun '11 | Group Stats use test email archive | A&H Active Underwriter | Qualitative only | N | The model was not used. The questions asked were too detailed for the model to be able to provide a sufficiently sensitive assessment. | B |
| 5 | WNC acquisition | The model was used to evaluate the impact of acquiring WNC and allowed an insight into the impact it would have on the existing business mix and reinsurance programme. | Mar–Apr '11 | KGL Board paper (April 2011) | RJK Board | Tail/Full dist | Y | N | B |
| 6 | Syndicate merger | Model was used to understand the effects of merging two syndicates (510 and 807) | Jan–Mar '11 | - RJK ARC paper (June 2011) - Underwriting Board paper (July 2011) | RJK Board | Full dist | Y | N | B |
| 7 | $100m xs $120m layer for Canada | Specific reinsurance efficiency study. Investigating the use of reinsurance to save capital. | Jan '11 | - Group email - Group Stats projects file | Property underwriter | Tail | Y | N | B |
| 8 | New classes of business | The model was used to investigate the impact of new classes of business. Following this three large categories of new business were established: - cyber (Enterprise risk team) - liability (Property) - warranty (Enterprise risk team) | Q4 2010 | - Group Stats Project file - Underwriting packs (Sharepoint) | UMC | Full dist | Y | N | B |
| 9 | New class of business: Space | A portfolio analysis tool was created using the model to determine a decision regarding Space business in Europe. This was also used in parametrising this new class of business in the model | 2010 | - Group Stats Project file - Underwriting packs (Sharepoint) | UMC | Full dist | Y | N | B |
| 10 | New business: Police scheme for S308 | The model was used to assess the implications of taking on a new large Binder for life insurance for the police authorities. | Aug-09 | Group Stats projects file | 308 Active underwriter/Division | Full dist | Y | N | B |
| 11 | Set up of syndicate 1880 | The set up of Syndicate 1880 including modelling of Ratatosk and TM Group and Kiln reinsurances | 2008 | JSC papers (Oct 2008) | JSC (committee since disbanded) | Full dist | Y | N | B |
| 12 | Setting up a transfer pricing structure/transfer of business to Bermuda | The model was used to investigate a proposed reinsurance agreement between Kiln Underwriting Limited (KUL) and Kiln Reinsurance Limited (Kiln Re). A substantial transfer pricing document was written by Kiln with input and review from PWC. The model was the primary tool for setting the pricing structure. | Q1 2007 | KUL Board paper (April 2007) | KUL Board | Full dist | Y | N | A |
| 13 | Underwriting pricing cycle | The model was used to investigate sensible capital management strategies to deal with the external pricing environment. | 2007 | RJK Board (2007) | RJK Board | Full dist | Y | N | A |

**Setting target levels of use**

3.4.34    As with a number of concepts within the IMAP process, it can be difficult to define how much use is sufficient to meet the use test requirements.  Model use is an area that the market and indeed regulator accepts will develop over time, so what is a target level of use acceptable for model approval?

3.4.35    One approach we have seen is to set a clear definition of this target level which can then be explained to the supervisor.  The definition will need to identify the core uses of the internal model and how much use is expected by the application date within each of these.  Specific examples and feedback on these uses is likely to be required to evidence progress. A clear model change process will aid the development of the model and ultimately help users to use the model more frequently.

3.4.36    By setting a target level of use, both the regulator and management team will have a clear set of goal posts. This also facilitates more detailed and specific project planning.  Though the FSA will be unlikely to agree these prior to formal submission, this discussion very much helps the debate and focus prior to application. Areas where the model has not been used to inform risk management decisions will need to be justified.

3.4.37    Use targets might include:

- List of target uses

- Identified users (individuals or committees)

- Training provided to users

- Documented objectives and process for each use

- Management information available at the required time and in the required format

- Standing agendas, internal reports, and minutes of forums where the model is used

- Feedback from model users, for example requests for model enhancements or different output

3.4.38    It is likely that the target level of use to achieve Internal Model approval expected by supervisors, will also increase as models and industry practice matures.

3.4.39    We note that a clear model change policy, responding to feedback from model users will support the development of the model over time; and ultimately help users to make use of the model more frequently.

## 3.5   Conclusion

3.5.1    In the chapter we discussed the practical challenges in meeting the Use Test.

3.5.2    Firstly we considered the issue of understanding what an appropriate level of use is for an insurance undertaking to aim for given the level of maturity of the internal model. This will be specific to individual firms, and we presented an example of the types of uses which firms could consider at various points on the maturity curve.

3.5.3    In our view, it is important for firms to make a realistic assessment of how developed their model is and adjust use accordingly.  Most certainly, any model, regardless of its level of development, can't and shouldn't replace management's responsibility and judgement in decision making.  Models can, however, inform, challenge and provide insights into those decisions which would not necessarily have been there without it.

3.5.4    Secondly, discussed the timescales and activity needed to gain buy-in to a model over time.  In particular, we covered: ensuring that the model produces appropriate output to meet the needs of stakeholders; gaining senior sponsorship for the model; and training / one-one coaching of senior management.  We also noted that simpler models and model output is easier to embed.

3.5.5    Finally, we discussed how firms would demonstrate compliance with the Use test. A key issue here is how much use is actually sufficient to pass the Solvency II requirements. We discussed one approach used by firms, where specific targets are set for use to provide clarity to the business and the supervisor on the extent to which the model informs decision making.

# 4    Statistical Quality Standards- Risk Ranking

## 4.1    Introduction

4.1.1    The Framework Directive [11] places a requirement on firms to ensure that the internal model's ability to rank risk is sufficiently developed such that the model is widely used and plays an important role in governance, the risk management system, decision making and capital allocation actions taken by the firm.

4.1.2    The Directive does not define "risk ranking" but CEIOPS and the Level 2 text set out a number of principles that a model must meet in order to 'rank risk' effectively.

4.1.3    The question remaining for companies interpreting the Directive is how to practically evidence risk-ranking within the internal model, for all material risks.

4.1.4    To answer this question, we have found it helpful to view the risk ranking requirement from the perspective of management, and their use of the model.

## 4.2   Summary of Framework Directive

4.2.1    Article 121 (paragraph 4) of the Framework Directive (Statistical Quality Standards) makes explicit reference to risk ranking.

4.2.2    Level 2 advice (Article 221 TSIM11 and Article 222 TSIM12) suggest four principles for the model's risk ranking ability:

- Coverage: all material risks should be covered by the model, and the risk ranking ability should exist for all material risks covered by the internal model;

- Resolution: the ability to rank risks shall provide a risk-ranking that is sufficiently precise for the purposes of risk-management , decision-making and capital allocation;

- Congruence: The ability to rank risk shall be consistent with the classification of risks used in the internal model and the classification of risks used in the risk management system;

- Consistency:  Risks of a similar nature are ranked consistently through the insurance undertaking and over time.  The risk ranking should reconcile with the capital allocation.

4.2.3    In addition, risk ranking forms part of the Use Test.  Firms must be able to show that the internal model plays an important role in:

- Their risk management system;

- Their economic and solvency capital assessment and allocation.

4.2.4    Finally, Article 223 TSIM13 notes requirements for the approach used to measure diversification benefits between risks included in the model.

---

[11] Article 121, Level 1 Directive

## 4.3   Practical challenges

4.3.1   This section describes a number of key challenges or considerations when demonstrating the internal model's ability to rank risk.  For each area, we have identified some solutions.  In addition, we have provided an explicit case study in risk ranking operational risks within Appendix E.

4.3.2   We have identified the challenges facing firms in this area:

- What does "risk ranking" mean i.e. what is the definition of risk ranking?
    - o   Qualitative
    - o   Quantitative

- How can one assess the model's risk ranking ability?

- Allowance for diversification

- Communicating the results to management

- Reliability of model results

- Use of the model in risk ranking

4.3.3   We refer to the 'Use Test' (above) for a discussion of how the firm might evidence use of the model in its risk management system.

## 4.4   Possible Solutions

*Definition of risk ranking*

4.4.1   We start from the assumption that the model's main purpose is to quantify risk in the firm's basket of exposures, by attributing losses to the downside risk that the firm faces.  To be used by management, the model must be able to quantify losses at the required level of granularity e.g. by risk, line of business and for the portfolio as a whole.

4.4.2   Our working definition of risk ranking is that "the model quantifies or measures risk effectively", i.e. the firm can use the model to measure risk and calculate return on capital in the *current* portfolio, and optimise risk-reward in *future* exposures e.g. by risk mitigation or business plan.

4.4.3   At a simplistic level, risk ranking enables model users to understand the significance (or materiality) of each risk type within the overall model results.

*Assessing risk ranking ability*

4.4.4   As with risk coverage, the ability of the model to rank risk can be assessed using both quantitative and qualitative approaches.  The assessment should focus on whether the model captures the relative importance of different risk drivers in relation to other risk drivers.

4.4.5    The approach taken to assessing risk ranking should be flexible, and can vary according to the requirements of users and modelling approach.

4.4.6    In our view, the objective should be to assess the model's ability to rank risk and that risk ranking is a test of the model rather than simply an output of the model.

## Qualitative Assessment

4.4.7    The table below sets out a worked example of the risk ranking assessment, applying the definition above:

| Level 2 Principle | Practical examples |
|---|---|
| **Coverage**: risk ranking ability should exist for all material risks covered by the internal model; | Results of (say) the reinsurance credit risk module and underwriting risk module should be useful to management and losses should be measured consistently at different percentiles<br><br>Any dependency effect, or diversification benefit, between reinsurance credit risk and the insurance results (e.g. increased defaults following underwriting losses) is also captured by the model. |
| **Resolution**:  the differentiation between the various risks and risk drivers has to be sufficiently precise to allow senior management to take appropriate decisions; | If reinsurance defaults on new business are modelled in one block, the model cannot be used to inform the firm's risk appetite for credit risk (e.g. no more than 20%of future exposure to counterparties rated BB or lower)<br>If a high-layer US property book is predominantly exposed to catastrophe losses, the appropriate drivers for loss experience would be US windstorm patterns as captured in a catastrophe model.   In this scenario, modelling loss experience using a simple frequency/severity model (parameterised using historic claims data) would not adequately inform pricing or outwards reinsurance purchasing decisions. |
| **Congruence**: The ability to rank risk shall be consistent with the classification of risks used in the internal model and the classification of risks used in the risk management system; | For example, a model structured <u>only</u> into attritional, large and catastrophe losses for the group as a whole, and not the firm's individual lines of business, would not be useful to management. |
| **Consistency**:  Risks of a similar nature are ranked consistently through the insurance undertaking and over time.  The risk ranking should reconcile with the capital allocation. | 1 in 200 year ultimate loss ratios on different lines of business are comparable, i.e. they are measured consistently between classes and year-on-year.<br><br>The output probability distribution forecast (ULR%) and capital allocated to Marine Treaty (say) reflects the riskiness and diversification of that line against the remainder of the portfolio.<br><br>The internal model is used to allocate capital to risk drivers, i.e. the approach to capital allocation is consistent with the quantification of pre- and post-diversified risk in the model. |

*Quantitative Assessment*

4.4.8    For a quantitative assessment of the model's risk ranking ability, it is necessary to define the metric used.  We note that, as with all aspects of Solvency II, judgement is necessary in interpreting the result of the analysis and the model's ability to capture the firm's risk profile using the chosen metrics.

4.4.9    Available approaches include:

4.4.10   **Stochastic measures** require metrics to be calculated from the sampled output results.  For stochastic methods, various ranking approaches can be deployed including ranking drivers based on the sample mean, variance, coefficient of variation, value at risk (say 99.5%) or based on the capital allocated under a particular method.

4.4.11   **Graphical representations** of model output (e.g. showing drivers of risk at different points in the output distribution) are also useful in helping management challenge the risk ranking ability of the model.  For example:
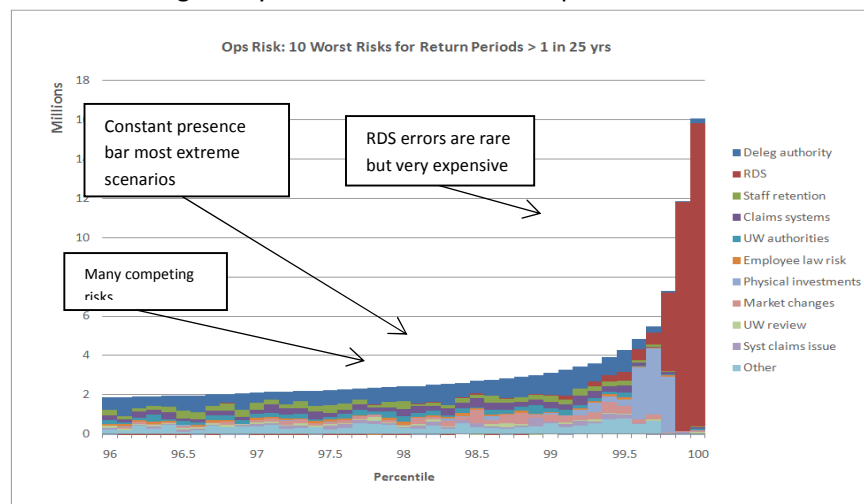


4.4.12   The objective of the analysis should be to order the model results, using the chosen metric.  Management or users of the model can then assess the firm's risk ranking ability, using judgement or stress and scenario tests.

4.4.13   **Stress and Scenario tests** are commonly used and these involve defining several scenarios and estimating the cost associated in each.  For stress tests, risks are ranked based on the estimated cost in each scenario at a given confidence level (e.g. 1 in 100).  The results can be compared to the model results at or around the same confidence level.  The key strength of stresses is that they are easily understood.  Over time, they often become a heuristic means of quickly grasping the relative importance of different drivers.

4.4.14   **Sensitivity tests** can be carried out to identify key risk drivers in the model.  A qualitative comparison can then be made to stress and scenario tests, including any inter-dependencies identified.

4.4.15   We note the results of a risk ranking assessment may be biased by the metric chosen or method used when constructing the ranking or drawing the graph.  For example, many capital allocation methods are influenced only by situations that occur infrequently and which are subject to limited risk mitigation.  In contrast, decision makers are likely to be focussed on everyday situations (e.g. 1 in 10 year Earnings at Risk).

4.4.16   For example, different metrics and confidence levels have been used to rank operational risks in the graph below.  Depending on the use, highest ranking risks at the mean (UW Authority) or in the tail (Delegated UW Authority) may be of more interest to the reader:

| Option | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Rank | Mean | SD | CV | Var 97.5 | Capital Alloc |
| 1 | UW Auth | RDS | RI | Del UW | Del UW |
| 2 | Staff Ret | Del UW | RDS | Claim systems | RDS |
| 3 | Emp Law Risk | Staff Ret | Phys Inv | Staff Ret | Staff Ret |
| 4 | Claim systems | UW Auth | Other | UW Auth | Claim systems |
| 5 | Del UW | Claim systems | Other | UW Review | UW Auth |
| 6 | IT | Phys Inv | Mkt Change | Emp Law Risk | Emp Law Risk |
| 7 | UW Review | Emp Law Risk | Other | Proj 1 | Phys Inv |
| 8 | Fraud Claim | UW Review | Other | Syst Exploit | Mkt Change |
| 9 | Syst Exploit | Mkt Change | Other | Fraud Claim | UW Review |
| 10 | Dis Rec Plan | Systemic claims issue | Other | Dis Rec Plan | Systemic claims issue |

4.4.17   More generally, we conclude that risk ranking works best if a basket of metrics and approaches are constructed (at different points in the distribution) as this reduces overreliance on any one method.

4.4.18   Different metrics may be better targeted to different uses or users of the model.

## *Allowance for diversification*

4.4.19   Risk ranking should take diversification benefits into account.  Accordingly, diversifiable and non-diversifiable risks should be dealt with separately.

4.4.20   We focus our attention on the diversifiable risks.  The challenge is how to allocate the diversification benefits fairly across all risks.

4.4.21   The factors affecting the amount of allocation depend on the method used, the tail end of the loss distributions of the risks and to certain extent the correlation structure.

4.4.22   We have listed 3 possible methods in dealing with this issue:

4.4.23   Option 1) use a pro-rata method – For example, if a model has an undiversified insurance risk charge of £25m and diversified amount is £20m.  Each of its component, say the underwriting and reserve risks can be prorated down by 80% (= £20m/£25m).

4.4.24   Option 2) using a TVaR allocation method.  This can be done by extracting the worst simulations of a particular risk e.g. insurance risk (using the example above) from the internal model and measure the mean sampled underwriting risk and mean sampled reserve risk for this tail dataset.  This would reflect the relative contributions of each aspect to the insurance risk.

4.4.25   Option 3) Using an averaging simulation method.  This involves extracting a reasonable number of simulations around the 1 in 200 years risk point.  For example, simulation number 150 is tracked down as the simulation corresponding to our sampled $20m insurance risk number.  Extracting, say the 40 simulations around this point would enable us to calculate the average diversified underwriting and reserve risk amounts.

4.4.26   There are still a few issues that we have not addressed, for example, the simulation results of the components may not be consistent with the overall risk capital.  For example, you have may have negative retained profit on the underwriting risk at the time of the positive retained profit on the reserve risks.  The averaging method should reduce this likelihood.

4.4.27   In addition, the methods do not allow for the period of exposure.  As an example, the exposure to the underwriting risk is shorter than the reserve risk.   In theory, this can allowed for by applying a "penalty" for the risks that have longer exposure but this may be difficult to implement in capital model in practice.

4.4.28   In any case, any significant influence of the allocation method used on the ranking of risk from the model, should be made transparent to management and users of the model.

## *Communicating the results*

4.4.29   It is important to communicate the results effectively to senior management.  If they are not communicated efficiently then it would be difficult for management to assess the ranking of risk in the model, and whether or not it reflects the firm's risk profile.  Model results may then not be used, because management does not understand them (or, model results may bias decision making inappropriately).

4.4.30   This presents an interesting additional challenge, on the one hand a basket of metrics/approaches helps to illustrate different facets of the risk being considered, but too many methods are likely to cloud matters in the decision maker's mind.

4.4.31   In practice, graphical tools can often help in this regard.  Several methods, including treemaps and stacked tail plots, are discussed in Appendix E.

4.4.32    One approach is to target output from the model at users.  For example:
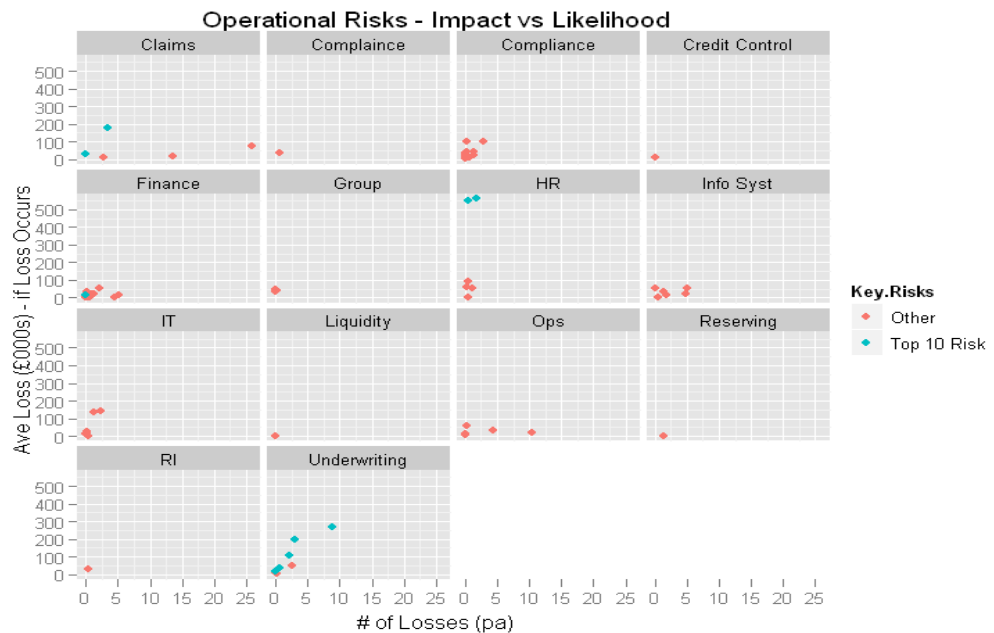


4.4.33    Recommendations with reasons are likely to be required when representing the results to senior management.

## Reliability of the model results

4.4.34    Management will only be able to use the model if it has sufficient confidence in the results.  The model should be reconcilable (to business plans and balance sheets) as well as detailed and stable.  In other words, it would be much easier to communicate the results of a model to senior management when you can evidence that it reconciles with other parts of the business.

### Achieving resolution

4.4.35    By "detailed" we mean granularity and risk ranking should be prepared at a variety of levels.  There is no prescribed level for monitoring.  As a starting point, ranking should be relevant and useful to the company.  Consequently, ranking at some/all of the following levels of granularity will help:

- *Within* risk categories – e.g. weighing up the significance of catastrophe vs. inflation exposures on the overall insurance risk charge.

- *Between* risk categories – e.g. assessing the relative importance of insurance risk vs. market risk when generating an overall capital figure.

- Between lines of business – i.e. understanding which classes drive the bottom line result

- Between reinsurers.

- Between operational entities – e.g. understanding the importance of UK vs. Bermudian subsidiaries.

### *Stable results*

4.4.36   "Stable" results may only be achieved if the internal model exhibits convergence in the tail.  Running more simulations helps to improve stability as more certainty will be provided about the results produced at 99.5 percentile. Tests should be run to check stability when different seeds are used by the underlying random number algorithm.  Where stability is not practicably achievable, e.g. 500k or 1m simulations plus needed, workarounds will need to be employed.  Options include, but are not limited to:

- Running the model several times with different seeds and averaging the results;

- As above, but selecting the highest result;

- Applying a fixed/variable add-on (for simulation error) to the model's output;

- Adjusting the model's input parameters in order to factor additional instability and therefore force a higher capital amount as a means of prudence.

4.4.37   Once results are stable, it should be possible to explain movements in the modelled result, year-on-year or under different input scenarios.

## *Use of the model in risk management*

4.4.38   Firms must be able to show that the internal model plays an important role in ranking risk and in their risk management system.

### *Ranking Risk*

4.4.39   The Level 2 advice on risk ranking highlights the use of the model in ranking risk.

4.4.40   In particular:

- Quantification of risk exposures – identifying key risks, so management can prioritise effort in managing and mitigating risk.

- Capital allocation and performance Measurement (e.g. performance against return on capital or equity targets, or other risk-adjusted measures[12]).

---

[12] The aim of accepting risk is to achieve returns that are higher than risk free rates.  The higher the risk is, the higher the expected return will be.  For forward looking risk areas, the analysis can be extended to allow for returns generated on specific activities.

This will help management understand if the company is being rewarded for those areas deemed to have the greatest risk.  Below are two ideas of how these can be measured:

- Risk adjusted return on capital (RaRoC): this is computed as the return (e.g. underwriting profit for line of business "X") divided by the allocated capital (e.g. underwriting capital for the same line of business).  This can be used for comparing different portfolios.
- "Sharpe" ratios: traditionally this is defined as the mean excess return divided by standard deviation of excess returns.  Although typically the preserve of investment professionals, a similar metric can be used in a general insurance context by dividing the mean underwriting profit for lines of business X by

*Other uses*

4.4.41    We have identified the following additional uses of the model in risk management:

- Risk mitigation decisions e.g. Reinsurance purchasing decision.

- Informing risk appetite setting i.e. quantifying current risk exposures.

- Monitoring risk profile i.e. risk exposures against risk appetite.

- Portfolio optimisation e.g. underwriting or investment strategy.

4.4.42    We note the usual challenges in use of the model in these ways:

- While risk measures can help inform decisions - e.g. the benefit of placing a certain reinsurance contract or deciding where to focus the entity's growth plans - it is often (if not always) inappropriate for the model results to dictate decision making.

- Effective use of the model results is always reliant on the judgements exercised by users of the information.  Decision makers may believe that the allocated capital amounts are "real" figures; scorers may try to "game" any model parameters particularly where they have a hand in shaping the inputs.

- Risk measures may also be inappropriate for judging individuals' performance – as many other factors come into play.

4.4.43    We refer to the Use Test section of our report for suggestions on how to evidence use of the model in these ways.

## 4.5  Summary

4.5.1    The guidance is principles based – this offers flexibility.

4.5.2    Companies will need to focus on materiality and consider both qualitative and quantitative approaches.

4.5.3    Providing a "basket" of measures will:

- Help achieve the requirements;

- enhance communication with Board members *et al*; and

- avoid overreliance on one method.

4.5.4    Diversification benefits can be allocated to the risks in various ways.   However, care needs to be taken to ensure consistency within risks and understanding the drawbacks of these methods.

---

the standard deviation of the lines' underwriting profit. This can be used for comparing individual risks within portfolios.
We note in practice it is often difficult to apply these measures.

4.5.5    Communication of results is key to enabling management to understand the model results, inform their view of risk exposures and challenge the model's risk ranking ability.

# 5  Statistical Quality Standards - Expert Judgement

## 5.1  Introduction

5.1.1  Use of Expert Judgement is a key part of the development and subsequent use of any internal model under Solvency II.  It is likely to be used in a number of areas, including instances of when credible data does not exist, as well as:

- when selecting appropriate methods to use,

- estimating parameters for those methods,

- assessing the reasonableness of outputs from a model validation exercise.

5.1.2  In the wake of the financial crisis, CEIOPS originally set a higher hurdle for the use and validation expert judgement than is currently explicit in the Level 2 draft text. In particular, reliance on expert judgement should be transparent; the judgement should be 'falsifiable' (i.e. circumstances should be identified, in which the judgement would fail, e.g. equity market shock); and the judgement should be validated.

5.1.3  The key principles remain in Level 2, however:

- Methods, assumptions should be based on the expertise of persons with the appropriate knowledge of insurance risks.

- Key assumptions (or judgements) should be justified, and their limitations should be transparent to management and users of the model.

- Management and model users should be made aware of model limitations, in particular the degree of reliance on expert judgements.

- There should be independent validation of material assumptions in the model.

5.1.4  The challenge for firms is to improve the governance and reporting of expert judgements as they are applied within the internal model.  In particular, loadings or adjustments to catastrophe models, quantification of reserve risk and dependency structure and assumptions are some of the most sensitive assumptions underlying the modelled result.

5.1.5  We expect the 'bar' for compliance in this area to emerge over time as firms progress through the IMAP process.

## 5.2  Summary of Framework Directive

5.2.1  The Level 1 text makes limited explicit reference to the use of expert judgement within the internal model.

5.2.2  CEIOPS advice to the Commission set a higher hurdle for use of expert judgement than is currently explicit in the Level 2 draft text.  In particular, their requirements included:

5.2.3     The expert judgement must be falsifiable, i.e. circumstances under which the expert judgement would be considered false can be clearly defined even though they may only be realised at a point in time far in the future.

- The expert must be able to make transparent the uncertainty surrounding the judgement, e.g. by providing the context of the judgement, its scope, basis and limitations.

- The expert judgement must be validated. Validation may include assessing the track record of expert judgements to assess reliability; challenging the expert judgement using scrutiny from other experts; comparing the expert judgement with existing and emerging data.

5.2.4     However in the Level 2 text, the requirements for expert judgement are more proportionate.  Article 4 stipulates,

- "Insurance and reinsurance undertakings shall choose assumptions...based on the expertise of persons with relevant knowledge, experience and understanding of the risks inherent in the insurance or reinsurance business thereof (expert judgment)."

- "Insurance and reinsurance undertakings shall, taking due account of the principle of proportionality, ensure that internal users of the relevant assumptions are informed about its relevant content, degree of reliance and its limitations."

5.2.5     In addition, Article 230 TSIM19 requires

- "The validation process shall be independent from the development and operation of the internal model. "

- "Insurance and reinsurance undertakings shall identify the most significant assumptions underlying the internal model and assess the effect on the results of the internal model to changes in the assumptions. Where the results of the internal model are highly sensitive to an underlying assumption, insurance and reinsurance undertakings shall be able to explain the underlying reasons and how this is taken into account in their decision-making process."

5.2.6     We also note that requirements around expert judgment are also included within the Board for Actuarial Standards' (BAS) TAS R regulation[13].  In particular there is a need to state all material assumptions used or recommended by an insurance or reinsurance undertaking, and to distinguish which of these assumptions are based on fact, judgements based purely on evidence and judgments influenced by person opinion.

## 5.3  Key Practical Challenges

5.3.1     The main challenges that firms are facing around the area of Expert Judgement that we identified during our interviews were as follows:

---

[13] Board for Actuarial Standards, November 2009, "Technical Actuarial Standard R: Reporting Actuarial Information"

- How do you **define/demonstrate** the suitability of an expert? (i.e. how do you demonstrate that someone *is* an expert),

- In which areas are expert judgements to be made (i.e. the **scope** of expert judgement), and

- How do you review/**validate** any expert judgements that have been made and what types of **governance** need to be placed around the expert judgement process.

5.3.2    A key concern for firms was how to justify the judgements made, where these by their nature are often made in the absence of internal or external data. It was important for many actuaries to find a way of meeting the requirements in a practical way so that they were free to investigate areas of uncertainty and apply judgments as they saw fit; responding to issues as they emerge.

## 5.4  Practical Solutions

### DEFINING/DEMONSTRATING THE SUITABILITY OF AN EXPERT

5.4.1    The Level 1 Solvency II Directive gives no requirements on how firms should go about defining or demonstrating the suitability of an expert.  As well as this, there are no specific requirements around justifying the suitability of an expert.

5.4.2    Being overly prescriptive around who would and would not constitute an 'expert' could potentially limit the use and application of expert judgement in developing and using the internal model.

5.4.3    An 'expert' need not be an individual and could in fact also be a panel/group of people that have formed a collective view based on their experience of the area on which an expert judgement is being made.

5.4.4    The current draft of the Level 2 text says that expert judgements should be based on the expertise of persons with:

- relevant knowledge,

- expertise, and

- understanding

of the risks inherent in the insurance/reinsurance business.

5.4.5    The above requirements form a useful set of guidelines to assess the suitability of an expert and could be used as set of criteria against which to measure credibility of an 'expert'.

5.4.6    In the diagram below, we provide an illustration of the types of professionals who could potentially be considered to be experts. It is our view that it is 'good practice' for all relevant experts to be consulted on key judgements in the model.

*EXAMPLES OF POSSIBLE EXPERTS*
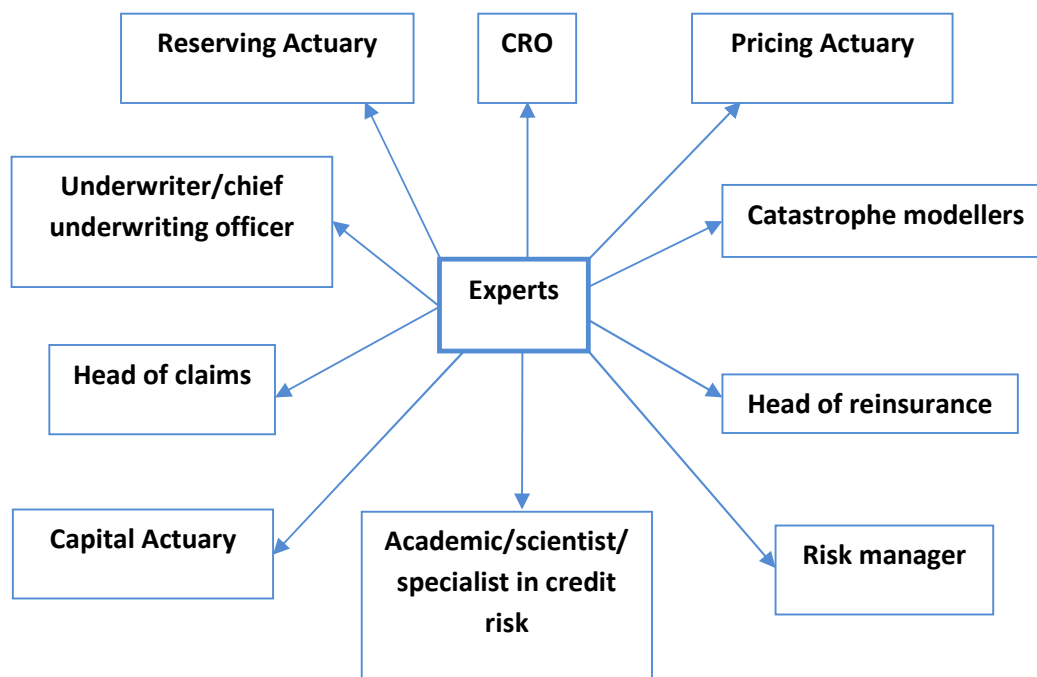


5.4.7    It is also important to document the suitability of an expert.  Taking the suggested requirements from the current draft of the Level 2 text would suggest that it would be useful (at a minimum) to document the following:

- The name of the expert(s);

- The relevant experience of the expert(s) (e.g. x years non-life experience/x years post-qualification experience);

- The relevant qualifications of the expert(s) that would help further justify their use as an expert (e.g. FIA/FFA or FCAS);

- The relationship of the expert to the internal model, and users of the results, including any conflicts of interest and relevant safeguards;

- The training/briefing received by the expert around the internal model and the way the expert judgement will be used (so as to demonstrate that the expert has a good understanding of the nature of the assumption and its intended use in the internal model, thereby reducing the risk of any mis-understanding); and

- The process/governance for making and reviewing judgments.

5.4.8    There may well be instances when experts disagree on the use of an expert judgement and such instances should also be documented together with details of the conflict resolution process and how a decision was reached on which view to adopt when selecting the final assumption to use in the internal model.

*SCOPE OF EXPERT JUDGEMENT*

5.4.9     One of the issues debated by this working party was that of the scope/areas to which expert judgement could be applied.

5.4.10    The conclusion reached by the working party was that expert judgements apply across **all** areas of the internal model, and not just around the data.   This includes methods, assumptions and review of model results.

5.4.11    However, it was also agreed that when classifying an 'expert judgement', *materiality & proportionality* are key and focus should be maintained on assessing the impact of **material** expert judgements only.

5.4.12    The Draft Level 2 Implementing measures[14] suggest that expert judgement is directly linked to the **assumptions** used in the internal model, as expert judgement is the way by which model assumptions are selected.

5.4.13    For example, expert judgement could be used in the estimation of a parameter used in the model, but also around the **choice of method** used in the model structure and the method to estimate those model parameters.

5.4.14    CEIOPS' advice to the Commission highlighted the following areas where expert judgement may be used to supplement data [15]:

- in selecting the data to use (i.e. data quality assessment),

- selecting the time period of the data,

- adjusting the data to reflect current and future conditions,

- adjusting for outliers and adjusting industry data to reflect the undertaking's circumstances,

- to act as a substitute for appropriate data and allow risk assessment which would otherwise have been impossible,

- to act as a complement to existing data,

5.4.15    Lloyd's guidance on **Solvency II Technical Provisions**[16] also recommends the use of expert judgment in:

- allocating whole account reinsurance contracts into segments,

- allocating expenses, and

- allowing for binary events.

5.4.16    In our interviews some of the **common areas in practice for applying expert judgement** included:

---

[14] Article 4 of European Commission, 31st October 2011, "Draft Implementing measures Solvency II".

[15] The Level 2 Advice issued by CEIOPS in response to Consultation Paper 56 (Tests & Standards for Internal Model Approval) suggests the following areas to which expert judgement may be applied.

[16] Lloyd's of London, March 2010, Detailed Guidance Notes Section 7 – Technical Provisions.

- Parameterising correlations,

- Predicting targets or frequency of terrorism events,

- Parameterising tail factors for reserves,

- Adjusting catastrophe model output, and

- Allowing for unmodelled catastrophes.

### EXAMPLE OF JUDGMENTS MADE, AND REPORTING

5.4.17    The Level 2 text requires key judgements and assumptions to be reported to management, including reliance on and limitations of those judgements.

5.4.18    This requirement links to CEIOPS' Advice earlier advice to the Commission,

"The expert judgement must be falsifiable, i.e. circumstances under which the expert judgement would be considered false can be clearly defined even though they may only be realised at a point in time far in the future." [17]

5.4.19    It's useful to refer to the recent financial crisis, and in particular the models used to price collateralised debt obligations (CDO's) to understand why the guidance focuses on reliance on expert judgements. CDO's were used to re-package blocks of US residential mortgage loans into tranches, where the highest-grade tranche often carried a triple-A credit rating.

5.4.20    Pricing models for the new issue CDO's were based on current and historic market prices of credit default swaps (CDS's), and an assumption that the correlation between default rates between loans could be represented using a Gaussian copula.  The basis of the model was flawed in that the market data used to parameterise the copula did not adequately reflect the strength of dependencies between defaults in extreme circumstances (i.e. systemic house price falls seen during the financial crisis of 2008-2009) and that under the Gaussian copula, the strength of the dependency does not vary across the distribution.  In addition, the sensitivity of the issue price to small changes in the assumed correlations factors was not understood.  Despite academic research at the time, criticising the model and the assumptions made, the judgements applied within this model were not understood by users of the models.[18]

5.4.21    In hindsight, it is easy to conclude that reliance on the judgements in these models were not clear to users of the model.   It is harder to foresee circumstances where judgements underlying capital models for general insurers might fail.

---

[17] CEIOPS, 2009, CEIOPs' Advice for Level 2 Implementing Measures on Solvency II: Articles 120 to 26 Test and Standards for Internal Model Approval
[18] http://www.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all

5.4.22    Examples of significant expert judgements applied in particular to general insurance models include correlations/dependencies between risk categories, use of external catastrophe models, quantification of operational risk, etc.

5.4.23    Wherever possible, the nature of these judgements and the uncertainty which attaches to them should be made transparent to the Board and users of the capital model.[19]

## REVIEW AND GOVERNANCE OF EXPERT JUDGEMENT

5.4.24    Having a robust governance process around the application of expert judgement together with sufficient validation/supporting evidence on expert judgements and documentation of the processes involved will be key to supporting any internal model application.

5.4.25    The risk the Directive is trying to mitigate here is that the judgments applied within the internal model are wrong, or the sensitivity of the modelled result (SCR) to those judgements is not communicated to management and users of the model.

5.4.26    This risk would be greater if:

i)    Judgement is set by one expert and is not challenged by another, with the appropriate skills and experience to ask the right questions about why the judgements made are appropriate e.g. hurricane models, dependency structures.

ii)    Judgement is made by an expert without sufficient knowledge of the underlying risk exposures (e.g. reserve CV or underwriting CV set without knowledge of the underlying business mix or trends in claims behaviour).

iii)    There is a conflict of interest between the expert applying the judgement and the model user (e.g. underwriter reviewing parameterisation of the distribution, with target RoC based on model output; capital actuary under pressure to reduce level of modelled capital)

iv)    Insufficient discussion of key assumptions by management – judgements not challenged or understood.

The risk is of course more significant also for more material model assumptions

5.4.27    To mitigate this risk, companies we spoke to were enhancing governance and documentation of the judgements applied in the model.

**Governance around Expert Judgement, including Reporting**

5.4.28    Some companies we interviewed already had an expert judgement governance policy in place and were following a robust procedure that included, for example:

- Who can make expert judgments and in which areas (i.e. assigning ownership of each parameter, data adjustment, methodology choice and guidance on whom to consult);

---

[19] This is a requirement under TAS R of the Boards for Actuarial Standards. For example see requirement C.4.

- Who should review expert judgments (which individuals / committees);

- Documentation requirements (including alternative judgments and the impact of the chosen judgment); and

- Requirements to communicate impact of expert judgments to senior management, including the uncertainty attaching to that judgement and the impact on the modelled capital result.

**Independence**

5.4.29     We have seen various approaches to ensuring there is sufficient independence in the review of expert judgements applied in the model:

- An Underwriter's input to a distribution reviewed by CUO or CRO, Capital Actuary (or all three).

- The Capital actuary's decisions are reviewed by a model governance committee, independent senior actuary (with relevant skills/experience) or external actuary.

- Material judgements are subject to independent validation as part of ongoing cycle of validation.

- A process of peer review of key assumptions is in place as part of an overall validation cycle.

- Clear responsibilities for sign-off of model assumptions and results.


**Documentation**

5.4.30     Documentation around specific expert judgements may include the following:

- The nature of the assumption, and its intended use in the internal model,

- The reasons why the assumption is based on expert judgement (e.g. lack of relevant data),

- Assessment of the materiality of the assumption, to include the quantitative and qualitative indicators on which this assessment is based on,

- Details of the expert involved (as set out in 6.3.8 above),

- Assessment of the uncertainty around the expert judgement,

- Date on which the expert judgement was made as well as when it was last reviewed, and

- The sign-off of the assumption used in the internal model (i.e. details of the organisation's functions involved in the sign-off of the assumptions, considering its materiality).

**Validation of expert judgements**

5.4.31    There are a number of approaches that could be used to validate expert judgements.  These include:

- Benchmarking (comparing to relevant external information),

- Challenge by an expert panel or by industry groups (though taking care to avoid systematic risks or "herd" behaviour),

- Back-testing (comparison to emerging experience), and

- Sensitivity analysis (i.e. varying the expert judgement and assess the resulting impact on the area for which the judgement is being made).

- Independent expert reports or peer reviews

# 6 Calibration Standards

## 6.1 Introduction

6.1.1 Calibration of the SCR is different from calibration of the ICA. Both are calibrated to a 99.5% confidence interval using a Value-at-Risk measure. However, the measurement basis of liabilities underlying the SCR is over a 12 month time horizon, rather than the "to ultimate" basis of the ICA.[20]

6.1.2 The SCR contains a provision for existing business (Reserve Risk) as well as new business written over the next 12 months (Underwriting Risk) - and reflects the expected volatility of losses over 12 months, including changes in the risk margin. Note that whilst the ICA also includes one year new business in its calibration, the SCR also includes all business bound at the closing balance sheet date (even if that business has not yet incepted).

6.1.3 One of the primary difficulties for actuaries is calculating a transfer value at the end of 12 months that is consistent with emerging experience over the first 12 months. After 12 months of very poor claims experience (the scenarios most relevant to the capital calculation), a human actuary would not blindly apply an unadjusted chain ladder model to experience in order to project a revised estimate of ultimate claims; their professional judgement would be applied. Determining how reserves would move given 12 months of new experience (in a way that can be automated for use in a capital model) is the focus of this section of the paper.

6.1.4 We discuss some common methods for calibrating the 12 month result, and some of the key issues in parameterisation of the methods.

6.1.5 At the time of our interviews and discussions with practitioners (2011) there was a wide variety in how progressed their companies were in relation to calibration. Some companies had not yet decided which method to use, while others were in the process of refining their assumptions. For the proportionate emergence methodology, in particular, best practice was still emerging.

6.1.6 It is worth noting that for the ORSA a number of companies are intending to use a "to ultimate" basis. Two reasons cited for this are:

- They feel it is a more appropriate for business steering (the 1 year time horizon may be "flattering" to long tail classes compared with short tail);

- As the 1 year methods are relatively new, they are not as trusted as the "to ultimate" basis.

---

[20] Art 122, Level 1 Directive. Note, however, that the insurer can use a different time period or risk measure for the SCR as long it can demonstrate that the results provide an equivalent level of policyholder protection.

## 6.2   Summary of Framework Directive

6.2.1    The Calibration standards are set out in Article 122 of the Level 1 text, but Article 101 is also directly relevant.

### *Article 122 of the Directive - Calibration Standards*

6.2.2    "Insurance and Reinsurance undertakings may use a different time period or risk measure than that set out in Article 101(3) for internal modelling purposes as long as the outputs of the internal model can be used by those undertakings to calculate the Solvency Capital Requirement in a manner that provides policy holders and beneficiaries with a level of protection equivalent to that set out in Article 101"

6.2.3    "Where practical, insurers should calculate the Solvency Capital Requirement directly from the probability distribution function generated by the internal model of those undertakings, using the Value-at-Risk measure set out in Article 101(3)."

6.2.4    "The supervisory authorities may allow approximations to be used in the process to calculate the Solvency Capital Requirement, as long as those undertakings can demonstrate to the supervisory authorities that policy holders are provided with a level of protection equivalent to that provided for in Article"

6.2.5    "Supervisory authorities may require insurance and reinsurance undertakings to run their internal model on relevant benchmark portfolios and using assumptions based on external rather than internal data in order to verify the calibration of the internal model and to check that its specification is in line with generally accepted market practice. ."

### *Article 101 of Directive - Calculation of the Solvency Capital Requirement*

6.2.6    "The Solvency Capital Requirement shall be calculated on the presumption that the undertaking will pursue its business as a going concern."

6.2.7    "The Solvency Capital Requirement shall be calibrated so as to ensure that all quantifiable risks to which an insurance or reinsurance undertaking is exposed are taken into account."

6.2.8    "It shall cover existing business, as well as the new business expected to be written over the following 12 months. With respect to existing business, it shall cover only unexpected losses. SCR should be calibrated to ensure that all quantifiable risks are taken into account."

6.2.9    "It shall correspond to the Value-at-Risk of the basic own funds of an insurance or reinsurance undertaking subject to a confidence level of 99,5 % over a one-year period. "

6.2.10   The SCR should cover at least:

- Insurance risk – including non-life, life and health; underwriting and reserve risk;

- Market risk;

- Credit risk;

- Operational risk.

*Summary of requirements*

6.2.11    The calibration standards aim to assess whether the SCR as derived from the internal model has the appropriate level of prudence.

6.2.12    The SCR is targeted to a 99.5% confidence interval, using a Value at Risk Measure, measured over a 1 year time horizon.

6.2.13    We note that Lloyd's syndicates will be required to calculate ultimate results as well as 1 year results, as it is likely that the ultimate results will be used to set capital requirements for Lloyd's names.

## 6.3   Overview of calibration methods

6.3.1    While a complete description of the methods is beyond the scope of this paper, below is an overview of the most common families of methods:

- Proportionate emergence;

- 'Actuary in the Box' / Simulated Re-Reserving;

- Perfect Foresight;

- Merz Wüthrich;

- Hindsight Re-Estimation;

- QIS 5 Undertaking Specific Parameters.

6.3.2    Within a stochastic capital model, deteriorations in claims experience (or savings against the opening reserve balance) are typically modelled simulation by simulation for each individual line of business, and origin years.  Most of the above methods (except for the Merz Wüthrich and QIS5) apply in this way.

6.3.3    Correlations or co-dependencies are then applied between years and lines of business, to generate an aggregate loss distribution over 1 year.   Alternatively, the 1-year method may be applied directly at the whole account level.

6.3.4    We note that these methods are inherently limited in that they do not allow for drivers of reserving uncertainty outside the sample data set being used to parameterise the method, including any tail factor development.  We highlight these methods should be applied with caution, particularly in exercising judgement in reviewing the parameter selections and the result.

6.3.5    We note that different methods or parameters may be applied to non-life reserve, underwriting and catastrophe risk, depending on the risk profile of the underlying exposures.

*Proportionate Emergence*

6.3.6　In this method the actuary creates a table of "emergence proportions" that represent the proportion of the ultimate deterioration or improvement for a simulation that would be recognised in the first year for any given origin period or class.

6.3.7　For each ultimate claim amount simulated, the emergence proportion is applied to the difference between the mean ultimate claims and the simulated ultimate claims, to give the 1-year results.

6.3.8　One formulation for the emergence method is:

- Mean ultimate claim + (Simulated ultimate claim – Mean ultimate claim) * Emergence proportion.

6.3.9　This would generate a booked ultimate, and a booked reserve can be calculated by deducting the simulated paid to that point in time (with 100% recognition)

6.3.10　Some actuaries have been considering using this method for setting the reserves only:

- Booked reserve = mean reserve + (simulated perfect foresight – mean reserve) *  emergence proportion.

6.3.11　The booked ultimate can then be calculated by adding on the paid claims.

6.3.12　There are advantages and disadvantages of each approach – the former probably requires parameters which are more immediately meaningful, and easier to peer review. But it could result in a negative booked reserve (if the paid claims which emerge in the year is higher than the recognised emergence – this would probably only happen if you had very volatile payment patterns).

6.3.13　The emergence proportion may vary depending on:

- The class of business - for example, the emergence proportion may be higher for property classes than for liability classes.  This is because, for a property class, it is often clear after one year whether losses are likely to be higher (or lower) than initially expected, whereas losses for liability classes typically take longer to emerge;

- The size of the difference between the simulated ultimate claims and the mean ultimate claims. It could be argued that, depending on the underlying liabilities, it would take a longer or shorter period of time to fully recognise movements away from the expected ultimate depending on the magnitude of the deterioration in 'to ultimate' result.  Actuaries have argued that larger movements would be recognised more quickly, but others have argued that they would be recognised more slowly.  The reader should therefore approach this issue with caution.

6.3.14　Emergence factors are typically applied deterministically, but they could also be stochastic.  This can introduce further variability to the results, but can also allow for the emergence factors to be correlated with the ultimate results.

6.3.15    For the proportionate emergence methodology best practice is still emerging.

## 'Actuary in the box'/simulated re-reserving

6.3.16    In this method, a triangle of actual paid or incurred claims forms the starting point for each simulation.  The next period's emergence of the paid or incurred data is generated to obtain a new diagonal.

6.3.17    The internal model then applies a reserving method to this triangle.  This should be consistent with the method used to project the next period's emergence (e.g. if bootstrapping was used the same underlying Chain Ladder assumptions should be used in the reserving method).

6.3.18    This generates an ultimate claim amount for this simulation.  Repeating for each simulation this gives a full distribution of the projected ultimates consistent with the 1 year results.

## Perfect foresight

6.3.19    This method assumes that the volatility in claims over 1-year is the same as the volatility of ultimate claims; for example, 100% of any reserve deterioration is recognised over the first year.

6.3.20    This can be seen as a special case of the proportionate emergence method, with 100% of deterioration (or release) being recognised in the first year.

## Merz-Wüthrich

6.3.21    This method estimates the volatility associated with the development factors underlying the chain ladder method, including parameter error (i.e. the parameter error associated with the selected development factor) and process error (i.e. the variance in the outcome expected given the development factor selected).

6.3.22    The Merz Wüthrich method operates on the same assumptions as the Mack method, which estimates volatility in the chain-ladder method to ultimate, but operates over a shorter time-horizon (1 year, if the claims data is annual).

6.3.23    The method can therefore be used to estimate the uncertainty in ultimate claims over a one year time horizon for Solvency II purposes.

6.3.24    The method is applied to the sample data by line of business, by year (to give a volatility estimate by accident year or year of account) and then aggregated (to give an estimate of the volatility for that line of business).

6.3.25    A one year distribution (e.g. lognormal) is parameterised for each distribution, based on the estimated volatility factor, and this is modelled stochastically with other lines of business to generate an aggregate distribution.

## Hindsight re-estimation

6.3.26  This method involves obtaining a triangle of historical estimated ultimate claim amounts, and using the historical movements in ultimates to obtain a distribution for future movements in ultimate claims.

6.3.27  It is necessary to select a distribution (e.g. lognormal) and then simulate the deteriorations over 1 year.

### STANDARD FORMULA / QIS 5 Undertaking Specific Parameters

6.3.28  The same methods used to obtain specific parameters for QIS 5 can be used to estimate standard deviations for the movement of ultimate claims over a 1-year time horizon.

6.3.29  In the Level 3 guidance[21], these include fitting a lognormal distribution to historic data (i.e. hindsight re-estimation) and estimating the mean-squared error of a paid claims development triangle (i.e. Merz-Wuetrich).

6.3.30  Under QIS5, two other methods were considered, including estimating the uncertainty in the one year reserve from historic outstanding claims, and using this estimate to fit a LogNormal distribution.

### Advantages and Disadvantages of Methods

6.3.31  Some advantages and disadvantages of the different methods are set out below:

| Method | Advantages | Disadvantages |
|---|---|---|
| Proportionate emergence | Parsimony | Crude method |
| | Ease of explanation to management | Parameterisation typically ignores likelihood of year 1 ultimate projection > actual ultimate. |
| | Consistent with ultimate volatility measures | |
| Actuary in the Box/simulated re-reserving | True one year method | Complex calculation |
| | Generally consistent with ultimate volatility measures | Reliance on standardised reserving approach- unrealistic for bad scenarios |
| | Can be extended to allow for BF as well as CL reserving | Difficult to control – can give unintuitive results |
| | No new assumptions/parameters are required | |
| Perfect Foresight | Easy to calculate | Not directly in compliance with |

---

[21] Draft proposal for Implementing Technical Standard on Undertaking Specific Parameters: Methods, December 2011

| Method | Advantages | Disadvantages |
|---|---|---|
| | | Solvency II |
| | Consistent with ultimate volatility measures | Biased – will almost certainly overstate risk emergence |
| Merz-Wütrich | Analytical | Complex formula |
| | Consistent with Mack Method | Assumes reserving is always by pure simple chain-ladder method, e.g. tail factors or BF methods not allowed for |
| | | Assumptions may not always be met in practice, e.g. no tail factor, independence / randomness in the underlying development factors |
| Hindsight Re-Estimation | Simple | Requires triangulation of best estimate ultimates-may not be available |
| | Grounded in experience | |
| | Easy to explain | Historic ultimates influenced by T&Cs and insurance cycle prevalent at that time – can't really "on level" old results |
| | | Similarly, will only consider results as extreme as history |
| QIS 5 Undertaking Specific Parameters | May be considered to be an industry benchmark | Method has not been used in the latest level 3 guidance |

## Empirical Testing of the Methods

6.3.32   The working party has conducted empirical testing of these methods; this testing was presented at the 2010 GIRO convention.  This testing did not indicate that any of the methods were significantly more accurate than the others.  With this in mind, it may make sense for companies to use the methods which are the easiest for them to implement, or for which they have the most data available, or which facilitate management's review.

## 6.4  Key Practical Challenges

6.4.1    As noted above, the key practical challenge facing non-life insurers is the measurement of liabilities over 12 months.

6.4.2    Further challenges include:

- The quantification of the risk margin.  From our survey results, it appears that firms are using a mix of stochastic and deterministic methods to estimate the risk margin (e.g. percentage of best estimate);

- The adjustment to best-estimate liability cash flows (as modelled) to allow for the impact of the risk margin.  In particular, the opening technical provisions on the Solvency II balance sheet include the risk margin; as do the closing technical provisions, as modelled at the 99.5th percentile.  The basis for the estimate of the SCR is not only the deviance between the mean and 99.5th percentile over 1 year (using best-estimate cash flows) but the net change in the risk margin between the opening and closing balance sheet.  This adjustment can be positive or negative, depending on the volume and nature of business and how this is changing over time.

6.4.3    A detailed discussion of these challenges is beyond the scope of our paper.

## 6.5  Possible Solutions

6.5.1    In this section, we summarise our findings on possible approaches to the 12 month measurement basis.

6.5.2    We note that, as many firms are applying these methods for the first time, it is still too early for our results to be conclusive.

### *Selection of Calibration Method*

6.5.3    In our survey, we asked which method people were using.  We note that based on our discussions with the market we believe many are using multiple methods to help parameterise the 1 year volatility factor, or have not yet decided which method to apply.

6.5.4    It has been suggested that companies tend to use the method that is easiest to implement in their software.  Anecdotally at least, Remetrica users are veering to proportionate emergence, while Igloo users may be more likely to use Actuary-in-a-Box.

6.5.5    We also note that although approximately 44% of respondents said that they were planning to use the Actuary in a Box method, the experience of the Working Party is that we have seen far more companies using the proportionate emergence method.  As discussed below, many companies appear to have tested the Merz Wüthrich and Actuary in a Box method, and often these are used to help calibrate the table of proportions required for the proportionate emergence method.

6.5.6    We note that many firms are applying different methods and comparing the results, as part of their validation of the 1-year time horizon.

6.5.7    Some applied bespoke methods.

## *Parameterisation of Calibration Method*

6.5.8    It appears from our interviews and discussions that the larger issue in calibration is not necessarily selecting the method to use, but how to parameterise it; particularly for the proportionate emergence method.

6.5.9    Many respondents who were planning to use the proportionate emergence method said that they would be using the bootstrapping volatilities (1 year: to ultimate), Actuary in the Box, QIS5 Undertaking Specific Parameters or Merz Wüthrich as an aid to coming up with their proportions. Judgement will be applied to select the resulting proportions based on the results of the various methods.

6.5.10   Many respondents said that they have looked at several approaches.

6.5.11   Some of the less common approaches we saw included:

- A stochastic parameterisation of the 'proportionate emergence' method – although it was found to be difficult to justify the assumptions;

- A bespoke model to estimate reserve deteriorations over 1 year;

- GLMs to examine the relationship between the 1 year and ultimate results.

## *Using different methods for different classes*

6.5.12   Some companies have said that they will be using different methods for different classes of business.

6.5.13   Those using proportionate emergence may be varying their table of proportions by class of business, particularly for instance between property and liability classes.

## *Validation of Calibration Method*

6.5.14   We note that this is a new area for general insurance actuaries, and it is a key assumption used in calculating the 1-year SCR. This means it is also likely to require a more in-depth validation.

6.5.15   Scenario testing is commonly used to test the reasonableness of the results coming out of the model.

6.5.16   We also found that pricing actuaries, or underwriters, were being consulted regarding the underlying risk profile of the business; whilst reserving actuaries were being consulted in order to understand the likely emergence of any deteriorations (or savings) in the reserves.

## 6.6   Other Issues

*Economic Capital*

6.6.1    The CEA's Solvency II glossary [22]defines economic value as the 'market consistent' valuation of assets and liabilities.

6.6.2    The CEA defines 'economic capital' as "the total of assets measured at market-consistent value, internally required by an insurer above the market-consistent value of obligations, in order to reduce the risk of not meeting the obligations to a defined risk measure (e.g. VaR, TVaR, EPD), and within a defined time period (e.g. one year)."[23]

6.6.3    There were a wide variety of responses in our survey to our question "how do you plan to define Economic Capital?"  Some of the responses were:

- Multiple of SCR;

- Other VaR measure;

- Against rating agency requirements;

- Economic assets minus economic liabilities;

- To be determined using ORSA;

- Lloyd's uplift;

- Under consideration.

*Use of the results*

6.6.4    It is important for all companies to ensure that their 1 year results are consistent with their ultimate figures, but different companies may have a different operational focus.

6.6.5    It appears to be a common approach that management information will contain capital figures on more than one basis – e.g. 1 year SCR, ultimate, economic capital.

6.6.6    Which measure of capital companies are using for management purposes may affect how the model is built or parameterised, even if this figure is not used for regulatory purposes.

---

[22]CEA Insurers of Europe, 2007, Solvency II Glossary accessible at
http://ec.europa.eu/internal_market/insurance/docs/solvency/impactassess/annex-c08d_en.pdf
[23] ibid

# 7 Profit and Loss Attribution

## 7.1 Introduction

7.1.1    The profit and loss attribution required for Solvency II extends to:

* Identification of the sources and causes of historic profits and losses

* An assessment of the model's ability to explain or project this variability

* The use of this analysis in

  o   decision making, i.e. helping management understand the variability in profits and losses
  o   validation, i.e. back-testing model assumptions against experience

7.1.2    We note that profit and loss attribution, as defined by the Directive, relates purely to the identification of sources and causes of historic profits and losses.  The *use* of this analysis in validation (both of the model's structure and assumptions) and in business planning are, strictly speaking, covered by the validation and use tests respectively.  [24]

7.1.3    Firms we spoke to who make use of their model within the business planning and capital allocation processes, felt a forward looking test is more useful.  The main purpose of the internal model was in their view to inform management's understanding of the variability in profits and losses, and the associated return on capital, in current and future years.  And to assist in explaining these drivers, to internal and external stakeholders (e.g. rating agencies, shareholders and management).

7.1.4    It's our view, therefore, that the assessment of the model's ability to model the drivers of profits and losses should be both forward and backward-looking.  In particular, the use of the model in business planning and validation are in our view key to meeting this test.

7.1.5    Profit and loss attribution is more prevalent in the Life Insurance sector, where emerging profits and losses are attributed to underlying drivers (e.g. investment, mortality, expense, lapse rates etc) in order to inform valuation or pricing decisions.

---

[24] We note Lloyd's have issued some helpful guidance in this area, which we encourage the reader to refer to (http://www.lloyds.com/The-Market/Operating-at-Lloyds/Solvency-II/Information-for-managing-agents/Guidance-and-workshops/~/media/Files/The%20Market/Operating%20at%20Lloyds/Solvency%20II/Workshops/Core%20Validation%20workshops%20slides%20final%20circ.pdf).

7.1.6    Within the general insurance sector, profit and loss attribution has not typically been performed for ICA models.  We comment on some potential approaches below for Solvency II.

## 7.2  Summary of Framework Directive

7.2.1    Article 123 requires the following -

Insurance and reinsurance undertakings shall review, at least annually, the causes and sources of profits and losses for each major business unit.

They shall demonstrate how the categorisation of risk chosen in the internal model explains the causes and sources of profits and losses. The categorisation of risk and attribution of profits and losses shall reflect the risk profile of the insurance and reinsurance undertakings.

7.2.2    The draft Level  3 proposed guidelines suggest the following –

### *Guideline 1– Definition of profit and loss*

7.2.3    Profit/Loss, for the profit and loss attribution, is the actual change in economic capital resources over the relevant period not attributable to capital movements (like dividend payments or public offerings). Without such movements the profit can be calculated as the difference between the economic capital resources at the beginning and at the end of the period. Economic capital resources in this context correspond to the surplus of assets over the technical provisions and other liabilities not treated as capital on an economic basis.

7.2.4    Note the Level 3 guidance does not define 'economic', but notes

- 'economic' profit/loss may well be different from the profit/loss shown in the financial statements

- that the probability distribution forecast (PDF) of profits/losses projected by the internal model should be consistent with the definition of profits/losses used the attribution.

### *Guideline 2 – Risk categorisation*

7.2.5    The risk categorization used by the undertaking for the profit and loss attribution should depend on the structure of the internal model as well as the intended applications of the profit and loss attribution in the use test and in the validation process. Typically, categorisation might be linked to risk types – e.g. cat risk, market risk, and/or lines of business.

### *Guideline 3 – Application of P&L in the use test*

7.2.6    The results of the Profit and Loss attribution provide valuable information for risk management and decision making. The output of the Profit and Loss attribution is thus a very important input to the Use Test. Therefore the undertaking has to evaluate and document on an ongoing basis where these results of the Profit and Loss attribution might be appropriately used within risk management and decision making. For this purpose, the undertaking should implement an appropriate process, on an ongoing basis, which includes a review of possible uses for profit and loss attributions that are already performed for validation purposes.

### *Guideline 4 – Application of P&L in validation*

7.2.7    The profit and loss attribution exercise provides information relating to how the model has performed in the past. Thus the results of the profit and loss attribution provide relevant information to feed into the regular validation cycle.

### *Summary of key requirements*

7.2.8    In summary, the current text requires;

- Analysis of profits and losses on an economic basis.

- Analysis of drivers of profit and loss by driver, based on the firm's:

    o   management structure (e.g. by business unit/line of business, as used in planning or performance monitoring; or asset class/category as used by management in setting the strategic asset allocation)
    o   risk management system and model structure
    o   Evidence that the results of the profit and loss attribution are being used to assess the model's ability to quantify risk (validation).
    o   Evidence that the results of profit and loss attribution are being used to inform decision making (the Use Test).

## 7.3   Practical Challenges

7.3.1    At a high-level, we found that firms are approaching the P&L attribution requirement with three main objectives:

1.  To analyse the drivers of historic profits and losses, and confirm that these are captured within the structure of the internal model (as described above).
2.  To assess the volatility in historic profit and losses, and use this analysis to validate the parameterisation of volatility in the internal model (as described above).  It was seen as a useful business planning tool to monitor the actual vs. expected experience, both to assess the viability of the business plan (at the mean) and risk in the business plan (i.e. body and tail of the distribution).
3.  To support management in assessing and managing the drivers of profits/losses in their business (as described above) as well as ensuring that they understand the capital model.

7.3.2    The key challenges we found firms faced in interpreting the requirements were:

- **Definition of Profit** – what basis of profits/losses to use for the analysis, for example Solvency II or Accounting bases e.g. IFRS, US or UK GAAP, Economic or Management's definition of profit.

- **Granularity** – should the analysis be conducted by Entity, Division or LOB, Insurance, investment class or by operational risk.

- **Historical Data** – how many years' historic data should be analysed. And, should historic results be "as if" adjusted to reflect subsequent market cycle movements when making such comparisons?

- **Trigger Levels** – how should trigger levels be defined?

7.3.3    Many were concerned by the extent to which the P&L test was a valuable use of expert modelling resources.  For example, for a Syndicate writing pure natural catastrophe exposure, with a conservative investment policy, it is more beneficial to devote time to validating the use of RMS or AIR results in the capital model than the historic model results.  Another example is where the current exposure bears little resemblance to previous years' business plan.

7.3.4    Yet again, it is important therefore to interpret the P&L test in a proportionate and pragmatic manner.

7.3.5    It was also noted that in many cases there is still a gap between the actuarial assumptions underlying the business plan (e.g. finance assumptions regarding expense ratios or allocation of reinsurance spend, or levels of prudence/optimism in the business plan) and assumptions underlying the model (e.g. on a best estimate basis).  Convergence of the internal model assumptions and the business planning process is required for P&L to inform management; this will take time to embed. it's the responsibility of both actuaries and capital modellers to work closer together.

## 7.4  Possible Solutions

7.4.1    We have spoken to a number of firms on their approaches to profit and loss attribution, and have formed our own views on the likely interpretation of the requirements.

7.4.2    We note that this is an area where firms are, only now, starting to invest resources. Therefore the comments below are not conclusive, as market practice in this area may well develop.

### *Definition of profit*

7.4.3    The Level 3 guidance does not define the required basis for profits and losses, other than referring to an 'economic' basis which may differ from the financial statements.

7.4.4    The Level 3 guidance requires that the probability distribution forecast (PDF) of profits/losses projected by the internal model must be consistent with the definition of profits/losses used the attribution.

7.4.5    In our discussion with actuaries, the following measurement bases for profit were being considered:

- Solvency II

- Accounting bases e.g. IFRS, US GAAP, UK GAAP

- "Economic" or "Management" basis e.g. Year of Account for Lloyd's Syndicates

7.4.6    In all our conversations with firms, there was strong support for:

- an analysis of Solvency II and accounting profits and losses at a total/aggregate level only.

- an analysis of economic profits and losses at a total level, and on a more granular basis as required by management to understand the drivers of profit and loss.

7.4.7    A detailed analysis of the underlying drivers by risk category and line of business was only felt to be useful to management, and the risk team, if conducted on the same basis as the management's view of profits.

7.4.8    There was concern amongst many of the firms we spoke to regarding the amount of additional resource a more detailed analysis on a Solvency II or Accounting basis would require.

## *Granularity*

7.4.9    There was strong support for targeting any more detailed analysis at the material drivers of risks:

- Insurance Risk - by Entity, Division/Business Unit, or Line of Business.

- Market Risk – investment returns should be analysed at an overall level rather than at an asset class level, unless the portfolio is non-homogeneous or risky.

- Currency Risk – this should be analysed at an overall level as the impact of currency movements is likely to be a single line item in the accounts.

7.4.10    Operational and Credit Risks were not felt to be material, and as such should be analysed in aggregate only, with more detailed inspection e.g. if individual operational risk events or near misses occur.

7.4.11    Many noted the rapidly diminishing returns offered by a more granular analysis, given the inherent volatility in smaller sample sizes.

7.4.12    It was generally accepted, however, that the model should be capable of producing results at a more granular level (if required) to support this detailed analysis.

## *Trigger levels*

7.4.13    Whilst the Directive requires 'triggers' to be defined for all validation tools (including profit and loss attribution and back-testing) it was felt these were in practice difficult to define and to apply.

7.4.14    It was felt by most that quantitative materiality thresholds (e.g. £5m deviation from plan) would only be useful as a rule-of-thumb to focus effort in the analysis.  Many observed that an element of judgement would always be required in determining whether or not to investigate further.

7.4.15    It was noted that *trends* in the emergence of profits and losses were more important to management's decision making than one-off movements in profits/losses.

7.4.16    Finally, it was noted that historic profits/losses should be used to sense-check the model results both in the body of the distribution and in the tail.

### *Historical data*

7.4.17    It was noted also that given that the business strategy or risk profile will evolve over time, the analysis of historical data, whilst helpful in identifying trends in experience, may have limited benefits..

## 7.5  Worked Example

7.5.1    The following is an example of the distribution of profit and losses output from the internal model.

### *Profit and Loss Attribution*

7.5.2    In this example, the key driver of the actual loss (-20) arises from insurance losses (-30) on the underwriting account with offsetting profits/losses in other areas.  These may be due to individual large losses, or trends in loss experience causing strengthening of the model reserves.

## Profit & Loss Attribution – a waterfall chart helps to track the key movements in sources of profit from plan
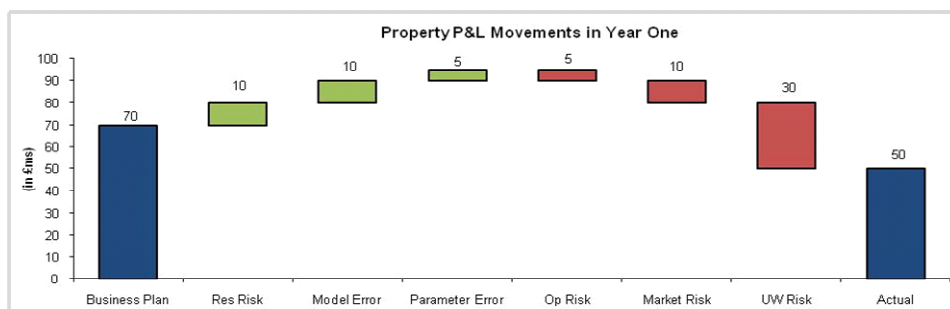


This enables
– the business to understand the areas of the internal model where differences have arisen from what was expected
– the actuaries to backtest volatility assumptions in the model, by looking at year-on year deviations, or more importantly trends

7.5.3    In this example, changes to volumes of business (model error) represent an adjustment for unplanned additional business volumes between 1/1/11 and 1/1/12, when the analysis was conducted. Changes to business plan assumptions (parameter error)' represent adjustments for revised business plan assumptions during the year e.g. rate increases or other expense assumptions (compared to the assumptions underlying the plan at 1/1/11).

7.5.4    At a more granular level, the source of the underwriting profit/losses can be attributed to individual business units or lines of business:

### Model Output

7.5.5    In this example, a comparison is made between the model output and the business plan. There is a difference between the business plan profit (70) and the mean of the modelled distribution of profits and losses (approx 63). In this case, it indicates a level of optimism in the plan.

7.5.6    The total business plan profit (70) is indicated by the red square, in the diagram below. The total, actual profit (50) is indicated by the green triangle, in the diagram below.

## Graphical Display of Profitability
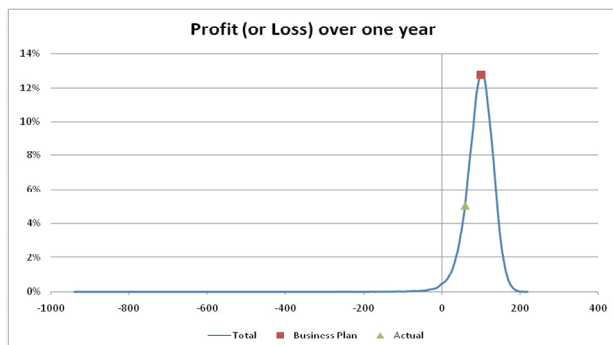


The variability in profit comes from a variety of **sources**:

- Lines of business (ie. property, motor etc)
- Risks (ie. non-life, market, operational etc.)
- Terms of trade & commission arrangements
- Business Strategy

And can be controlled by levers that **cause** profit variability:

- Investment portfolio
- Reinsurance protection
- Pricing & underwriting
- Terms of trade & commission arrangements
- Business Strategy

7.5.7     This information can be used as part of the business planning or performance monitoring cycle, to support management in their assessment of profits and losses. The model can be used to inform management's assessment of the return period of the actual profits/losses experienced. For example, on the profit charts below, property/channel 1 exhibits approximately a 1 in 8 year profit, whereas property/channel 3 exhibits approximately a 1 in 50 year loss.

## Graphical Display of Profitability
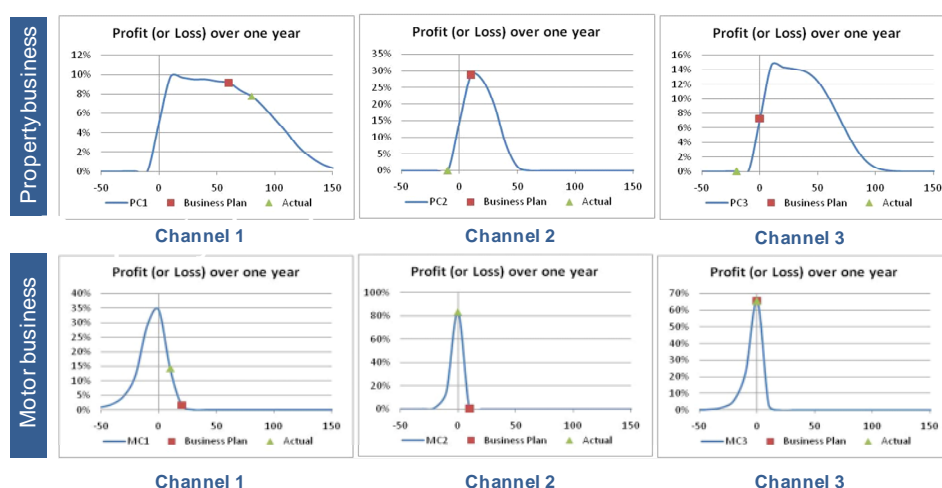


7.5.8     We note the business plan mean contains management adjustments and therefore may be different from the modelled mean for example given above.

7.5.9    The test of profit and loss attribution then consists of the following four components:

## Assessment of categorisation of risk in the internal model

7.5.10   In this example, it's likely to be reasonably straightforward to conclude that the model captures the drivers of loss experience excluding volumes and rates.   In this model, underwriting risk is modelled stochastically (including underwriting and reserve risk) and the model produces results by channel and line of business.

7.5.11   Had the losses arisen from e.g. US$ claims experience, in a year when the Dollar moved materially against Sterling, the conclusion might have been different.  Many models model insurance risk in consolidated GBP.  If the US$ liabilities were a material exposure and a material driver of profit/loss on the underwriting account, the structure of the model above would be inadequate if the calculation kernel did not consider currency risk.

## Application of profit & loss in the use test

7.5.12   The analysis of profit and losses above is only useful to the firm if the results are used to inform decision making.  For example, if the graphs above are reviewed by Finance or the Underwriting Committee as part of the business planning / performance monitoring cycle.

7.5.13   Given time, the firm would be able to cite examples where the results had informed decision making (e.g. growth in a profitable channel).

7.5.14   It is necessary to ensure that profit and loss distributions are of sufficient granularity for use within the business (e.g. they are at a line of business level).

## Application of profit & loss in validation

7.5.15   Again, the analysis of profit and losses above is only useful to the firm if the results are used to enhance the modelling approach, where necessary.

7.5.16   For example:

- Management may consider that a confidence level of 1 in 25 years for the loss observed on property/channel 3 (-25) is too extreme for the nature of the losses experienced.  In this case, the firm may revise the volatility assumptions for channel 3 upwards in order to increase the likelihood of this level of loss to 1 in 10 say.

- Management may seek to enhance the model structure to model the link between currency and insurance losses explicitly, using their economic scenario generator to model volatility in exchange rates.

- Alternatively, management may decide not to react to either of the scenarios above until there is a sustained trend in loss experience.

# 8  Validation Standards

## 8.1  Introduction

8.1.1    Solvency II requires all firms to have a regular cycle of validation in place.  The firm's approach to validation should be approved by the Board in the validation policy, which forms a core part of the IMAP submission.

8.1.2    Validation is a set of tools and processes, which are used by the Board and external parties to gain confidence over the information produced by the internal model.

8.1.3    When compared to the ICA regime, Solvency II validation is considered to be much more extensive in scope and formality.  For example, Solvency II requires:

- Validation of all aspects of the internal model – not just the parameters, methods and output results, but also the IT systems, data (internal and external), documentation and governance of the model, underlying external models (e.g. catastrophe and ESG's), as well as the use of the model.

- Formalisation of the firm's validation approach, in the firm's validation policy.  This includes the firm's ability to demonstrate appropriate levels of independence in the validation, particularly where expert judgements have been applied.

- There is an expectation that the validation will provide challenge to the model – i.e. where the validation identifies problems or sensitivities in the modelling approach, follow up actions or remediation activities are identified.  For example, CEIOPS recommended definition of pre-set criteria for each validation tool; and the FSA and Lloyd's all require firms to explain 'what an unreasonable output would be'. [25]

- An ongoing 'business as usual' (BAU) approach to validation to support ongoing use of the model in the business, in the regular calculation of the Solvency Capital Requirement.

8.1.4    Validation is a difficult area of the Directive for firms to respond to.  As well as being critical to approval of a firm's IMAP submission, and subject to Board and Regulator scrutiny, it demands expert input (e.g. actuarial, underwriting, treasury, etc.) and therefore highly resource intensive. (Firms responding to our survey said they expected validation to be 2 to 4x as extensive under Solvency II as under ICA, in year 1; and 1.5 to 3x as extensive on an ongoing basis.)

8.1.5    Firms have said that the benefits of strong validation included[26]:

- Good Practice

- Appropriateness of modelling methods

---

[25]CEIOPS (now EIOPA) Level 2 advice (former CP56 8.142); Lloyd's Validation Report Guidance (June 2012); FSA 2011 Solvency II: Internal Model Approval Process.
[26] http://www.iiag.org.uk/events/documents/IIAGMeeting25thMay2012-Seminarslides.pdf

- Issues or errors in data and calculations

- Confidence in Outputs

- Confidence in using model to support strategic decision making & risk management

- Board responsibilities

- Assists the Board in meeting its responsibilities

- Transparency

- Assumptions and limitations of the internal model are transparent to key stakeholders

- Benchmark against peers

- Understand where model sits relative to peers and areas for development & improvement

## 8.2   Summary of the Framework Directive

8.2.1    Article 124 - Validation standards require the firm to have:

- "... a regular cycle of model validation which includes monitoring the performance of the internal model, reviewing the ongoing appropriateness of its specification, and testing its results against experience."

8.2.2    The validation should cover:

- "monitoring of the performance of the internal model"

- a review of "the ongoing appropriateness of its specification"

- "testing the results against experience" (i.e. "loss experience but also to all material new data and information relating thereto", by which we understand, exposure data)

- "an effective statistical process for validating the internal model which enables the insurance and reinsurance undertakings to demonstrate to their supervisory authorities that the resulting capital requirements are appropriate."

- "stability of the model"

- "sensitivity of the model to key assumptions"

- "validation of data quality"

8.2.3    Whilst the overriding objective of validation is not explicit in the Directive, we note:

- Level 1 focuses on confirming the quantum of the SCR, viz "the validation process shall ... demonstrate ... that the resulting capital requirements are appropriate"

- Level 2 expands the scope of the validation to cover "a. Data, b. Methods, c. Assumptions, d. Expert judgement, e. Documentation, f. Systems/IT, g. Model governance, h. Use test"

- In response, many – but not all - firms have designed a validation policy which covers both quantum and compliance with all tests and standards (Art 120-126)

8.2.4    There is significant further guidance in the Level 2 draft text and in CEIOPS' final advice on Level 2 implementing measures (former CP56). This requires firms to put in place a validation policy; for the validation to cover all aspects of the internal model's scope and all tests and standards for compliance (Art. 229 TSIM18).

8.2.5    The validation policy should include:

- "the procedure to be followed in the event that the validation process identifies problems with the reliability of the internal model and the decision-making process to address those concerns;"

- "an assessment of the independence of the validation process" including both internal and external validators.

8.2.6    The contents of the validation report are not specified by the Directive.

## 8.3   Practical Challenges

8.3.1    This paper discusses some key issues and considerations seen by the working party through the surveys, interviews and other discussions with actuaries involved in the Solvency II internal model approval process.

8.3.2    In particular, we focus on the key differences between validation under ICAS and under Solvency II:

- How to define the **purpose and scope** of the validation.

- How to apply the **concept of materiality** within the validation work, in order to deliver value to users of the model.

- Which **validation tools** are required.

- How to assign **roles and responsibilities** for validation i.e. governance framework, including an appropriate amount of independent review.

- How to **formalise** the technical analysis - for example, possible to define pass/fail criteria for validation tools in a practical way?

8.3.3    We comment in more detail on **validation of expert judgement** in chapter 6, and include a detailed description of validation tools in the appendices of this report.

8.3.4    We also provide worked examples of validation tools.

## 8.4   Practical Solutions

8.4.1    We have spoken to a number of firms on their approaches to validation, and considered our own experiences in forming the views set out below in this paper.

8.4.2    We note that validation is an area where market practice is still developing, and good practice will continue to evolve as firms progress through IMAP.

*Materiality*

8.4.3    The FSA's guidance on proportionality is that all areas of the Solvency II requirements should be covered, but the depth or extent of work in each area might vary depending on its importance to the firm.  This is particularly relevant for validation, where the scope of the model is wide, the purpose of the work may be unclear, and the depth of validation required varies depending on materiality of model components or assumptions to the SCR.

8.4.4    A reasonable starting point to deciding how much time to invest on specific validation tools, would be to identify those methods or assumptions that are the most material to the model results; or those aspects of governance and controls that prevent material errors in the SCR.   Firms we spoke to identified those areas through judgement, often backed up by sensitivity testing.

8.4.5    Some also ranked validation tools in order of importance – e.g. critical, high, or medium-low level validation tools – see 'validation tools' below.

8.4.6    Another area where materiality is important is in assessing the impact of validation results on the appropriateness of the model or the results.   One possible approach to the definition of materiality is set out below:

| Materiality | Impact on SCR |
|---|---|
| High | >10% impact on SCR |
| Medium | 5-10% |
| Low | <5% |

*Purpose of the validation*

8.4.7    As noted above, the overriding objective of validation is not explicit in the Directive:

- "The administrative, management or supervisory body shall be responsible for ensuring the ongoing appropriateness of the design and operations of the internal model, and that the internal model continues to appropriately reflect the risk profile of the insurance and reinsurance undertakings concerned." (Art120)

- "The model validation process shall include an effective statistical process for validating the internal model which enables the insurance and reinsurance undertakings to demonstrate to their supervisory authorities that the resulting capital requirements are appropriate." (Art 124)

8.4.8    A key challenge for firms is to understand what is meant by "appropriate" in the context of capital modelling.  The uncertainty surrounding the estimate of the SCR at the 99.5th percentile is significant.  This is because models are most subject to simulation error at this point; there is a paucity of historical data at higher confidence levels that can be used to benchmark the model's results; and given the necessary reliance on expert judgment.  Whilst reserving actuaries might cite a range of +/- 10% in best estimate claims reserves for some lines of business (say); anecdotally, capital modellers cite much wider confidence levels of perhaps +/- 20% in the SCR.

8.4.9    We note that Lloyd's proposed the following confirmation statements in Syndicates' validation reports in 2011.  The reference to 'materially mis-stated' refers to the CEIOPS Level 2 advice to the Commission (former CP56).  At the time of writing, we understand the form of the statements is likely to be revised from 'materially mis-stated' to 'appropriate' in line with the Level 1 and Level 2 draft requirements:

- That the SCR is calculated in line with applicable regulations and is not materially mis-stated

- That the one-year to ultimate capital calculation is not materially mis-stated

- That the internal model materially meets all relevant regulatory standards

- That key output information is appropriate for the business decisions it is used to inform

- That validation has been conducted in line with the Validation Policy

- That an appropriate level of independence has been maintained during the validation.

8.4.10    We discussed, what we meant by 'appropriate' in the context of this uncertainty.  The table below lists some key attributes:

| What is 'appropriate'? | Reference |
| --- | --- |
| SCR reflects all available information about the 1 in 200 year position, including historic data, market data, and information about the firm's risk profile at the date of the calculation[27] | L1, Art 121(2) |
| Methods, assumptions and model results in the internal model are justified/ rationalised in view of this information | L1, Art 121(2) |
| In the view of an "independent knowledgeable third party", the SCR is estimated using methods and assumptions that are reasonable | L2, Art 230 TSIM19 (3b) |
| The SCR is estimated on a consistent basis year-on-year, so that changes in the SCR result and underlying assumptions are explained/ rationalised | L2, Art 221 TSIM11 |
| The model is consistent with the business plan, technical provisions and model results are used in risk management, and decision making (e.g. aggregate monitoring) | L2, Art 212 TSIM2, Art 215 TSIM5, |

[27] In practice, this is probably implemented by comparing lower level historic events with comparable return periods from the model; and by management's view of risk (e.g. stress & scenario testing, reinsurance purchase assumptions) to higher return periods from the model.

| What is 'appropriate'? | Reference |
|---|---|
|  | Art 218 (4) |
| The model meets tests and standards for model approval, including documentation | L2, Art 229 TSIM18 |
| Material assumptions, expert judgements and sensitivities in the model results are transparent (eg dependencies, business plan loss ratios) and are justified | L2, Art 4 and Art 223 TSIM22 |
| Methods are implemented accurately in the model | L2, Art 218 TSIM8 (3e, f) |
| The purpose and limitations of the model are clear to management / users of the model | Art 213 TSIM3 |

8.4.11   No model is perfect, and all models can be improved.   It may be the case that improvements can be made in the areas of weakness identified by the validation. But even if the modelling cannot be enhanced in these areas, it is still valuable to understand those areas of weakness in the model.  For example, management can be mindful of model key assumptions, when using the model information in business decision making.

8.4.12   Whilst it will never be possible for the actuary to estimate the SCR to any degree of certainty, in our view challenge through the validation work should lead to increased confidence in the results.

## *Scope of the validation*

8.4.13   We note that Level 2 requires the validation to cover

- "all parts of the internal model"

- "all requirements set out in Articles 101, 112, 120 to 123, 125 and 126 of Directive"

8.4.14   The directive does not set out how a firm's internal model should be defined. Market practice is likely to develop as supervisors review and approve models. [28]

8.4.15   The following table shows the potential scope of a firm's validation.  In our view, the validation should cover all items in the table, subject to our earlier comments on materiality:

| Internal Model | | |
|---|---|---|
| **Risks** | **Model 'parts'** | **Compliance** |
| NL Insurance | Uses | Use |
| -   Reserve Risk | Results (i.e. SCR and other metrics) | Governance |
| -   Underwriting | Calculation Kernel (e.g. Igloo) | Documentation |

---

[28] Note Article 203 IM1 requires firms to explain how the internal model covers all material and quantifiable risks to which it is exposed.  An application for partial internal model approval should also include an explanation of the risks, business units and entities included in the application.

| | | |
|---|---|---|
| -　Catastrophe | Parameterisation (i.e. spreadsheets) | Statistical Quality Standards |
| Market Risk | Data and Inputs | Calibration |
| Credit Risk | Methods and Assumptions | Validation |
| Operational Risk | Governance and control environment | Profit & Loss Attribution |
| Aggregation & Dependencies | Catastrophe Models (e.g. RMS) | External Models |
| Other (e.g. material uses) | Economic Scenario Generators (e.g. ESG) | |
| | Pricing and/or Technical Provisions Models (only if used to inform material assumptions in the internal model) | |
| | Documentation | |
| | IT systems | |

8.4.16　We note validation of technical provisions models is required under Level 2, Article 255 x3 TP24.

## *Validation Tools*

8.4.17　The following table summarises the tools that are mentioned in the Solvency II requirements, and Lloyd's of London guidance issued during 2012.　We consider a 'tool' to be a piece of analysis or control which gives the company additional comfort over the appropriateness of its calculated SCR result.

8.4.18　The list of tools mentioned in the guidance is not comprehensive.　There are several tools (e.g. analysis of change) that are commonly used by capital actuaries and are indeed essential to an understanding of the SCR results, but which are not mentioned in the requirements.

8.4.19　Table of validation tools:

| Validation Tools | Lloyd's June 2012 | Level 1 | Level 2 Oct 2011 | Good Practice |
|---|---|---|---|---|
| Stress and scenario tests, "As if" losses | v | | v | v |
| Reverse stress test | v | | v | v |
| Sensitivity tests (including correlation / dependency) | v | v | v | v |
| Profit and loss attribution | v | v | v | v |
| Back testing | v | v | v | v |
| Testing results against experience | v | | v | v |
| Stability/ Simulation / convergence test | v | v | v | v |
| Model robustness | v | v | v | v |
| Analysis of Change | v | | v | v |
| Hypothetical portfolio | | v | | v |
| Standard SCR formula comparison | | | V | v |
| Statistical tests | v | v | | v |
| Inspection of outputs / reasonableness checks (sense checks) | | | v | v |
| Reconciliation of inputs and outputs | | | | v |
| Re-performance testing of model calculation | | | v | v |
| Checks on data accuracy, completeness, appropriateness | v | v | v | v |

| Validation Tools | Lloyd's June 2012 | Level 1 | Level 2 Oct 2011 | Good Practice |
|---|---|---|---|---|
| Qualitative reviews | v | | | v |
| Interviews with modellers, management, users | | | | v |
| Expert Judgement review | v | | v | v |
| Assumptions review and benchmarking by experts | | | | v |
| Assumptions review and benchmarking by business | | | | v |
| Assumptions review against market practice/benchmarking | v | v | v | v |
| Methodology review against market practice | | v | v | v |
| Review of model documentation | | | v | v |
| Assessment of model governance, and control | | | | v |
| Audit trail | | | v | v |
| External models | | v | v | v |

8.4.20    There is little guidance on the frequency of application of the different tools in the Directive.  The firm is therefore responsible for defining:

- which tools to apply to different components of the model, for example which assumptions to stress.

- how frequently to apply the tools, for example whether stress tests should be run quarterly or annually, etc.

- Reasons/explanations for stressing specific assumptions and the chosen approach.

8.4.21    The responsibility is therefore with firms to define an ongoing cycle of validation that is appropriate for their internal model.  For example, identifying:

- "Critical" validation tools, which are run quarterly or at least in line with regular updates of the SCR(e.g. analysis of change, sense checks of SCR results, review and sign-off of results by the Risk Committee or Executive).

- "Key" validation tools, which are run at least annually or at least in line with re-parameterisation of the model (e.g. review of underwriting risk parameterisation by underwriters, back-testing).

- "Low level" validation tools, which are run less frequently but which provide additional insight on key areas of the model or the firm's risk exposures (e.g. expert review of dependency structure and methodology).

8.4.22     Where we have seen good validation, this includes a mixture of validation tools providing 'top down' (e.g. analysis of change) and 'bottom up' (e.g. QQ plot) used to justify  the model results.

8.4.23    There is a natural overlap between the underlying modelling and parameterisation process and what is considered under Solvency II as validation.   For example, tools applied within the parameterisation process (e.g. goodness of fit testing) or applied by the capital actuary in sense-checking the model results (e.g. analysis of change) are examples of validation tools.

*Governance, Roles and Responsibilities*

8.4.24    A key aspect of the formalisation of validation is defining the roles and responsibilities of key individuals who are part of this process. In particular, there is a requirement for the validation to be independent of the design & operation of the internal model. (Article 229 TSIM18)

8.4.25    Firms we spoke to stressed that the requirement for independence in validation does not necessarily mean that they will employ third parties (such as consultancies), if sufficient independent expertise exists in-house to provide objective challenge over the model.

8.4.26    From the surveys and interviews being carried out, we saw a number of different models for the validation roles and in particular the independent validation being considered. Examples of these are as follows:

- A common model seen utilises the traditional three lines of defence. This typically involves model builders and operators as the first line; a distinct internal role (for example the CRO, assuming they were not directly involved in the model operation) as the second line; and validation performed by an independent party or internal audit as the third line.

- A high proportion of respondents said that they would use external consultants to some extent in the validation process. In the firms we spoke to, this varied from a full scope independent validation performed by external consultants (to provide the Board with a full, independent view of the model's readiness for IMAP) to targeted validation focussed on material methods, assumptions and model results.

- Where independent validation was primarily kept in house, examples of how this might be achieved included the use of internal audit, or by via reviews conducted by the central group function in the case of subsidiaries within a group. Internal audit may not have the required skills for some of the more technical reviews, but might commission external expertise for these areas.

- The way that roles are set out typically seemed to reflect the size of the undertaking. In some larger undertakings for example, there are separate internal model and validation teams, whereas for smaller undertakings there may be only one or two model operators with the CRO performing the second line validation.

- Inclusion of business experts formally in the validation process has also been considered. For example, the second line validation of underwriting risk might include underwriter sign-off. (Note: it would be important to demonstrate that there were safeguards against any conflicts of interest in the underwriter's review, e.g. review of all classes by the CUO or by an independent actuary.)

8.4.27    Closely related to the roles are the types of review carried out by each party. In our survey, examples of the styles of review included:

- Detailed audit

For example walkthrough of the working papers or parameterisation spreadsheets

- Peer review

Higher level than the detailed audit, this generally involves reviewing outputs and documentation before questioning any unusual features. This may be extended to involve a more detailed look at key areas.

- Benchmarking

This may typically form part of an independent third party review where they can compare model results to that seen by peers. This may also involve running an independent model of the company and comparing results.

## 8.5   Worked examples

8.5.1   In this section, we provide some simple examples of validation tools and their application.

8.5.2   We refer to Lloyd's 2012 Validation Guidance and workshop slides for further examples, including sensitivity, stability testing.

### *Back-testing against historic data*

8.5.3   Back-testing is often referred to in EIOPA texts as testing results against experience. The principle is to test assumptions and results against historic experience, possibly with adjustments made to bring historic data onto a basis that is representative of the period being modelled (for example, adjustments for claims inflation or the premium cycle).

8.5.4   Backtesting can be used at various levels of aggregation within the model.   In our experience, testing is particularly useful at higher levels.  For example, backtesting a directly fitted large loss distribution may not add much value as it may be just repeating the fitting process (i.e. fitting the distribution to historic adjusted data). But comparing the resulting aggregate loss ratio distribution might be more informative. The modelled loss ratio distribution will be a consolidated model of the combined underlying models for large losses, attritional losses, catastrophes and premiums. Hence, demonstrating the accuracy of the resulting loss ratio distribution will in turn give confidence in the underlying models and their interactions (dependencies). This is analogous in a way to a simple check-total spreadsheet test.

## Worked Example:  Back-testing distribution against historic data

| Objectives /Limitations | What is being validated | Description of validation | Expectations | | Results (see graph) | |
|---|---|---|---|---|---|---|
| Objectives<br>Test key assumptions against historic and emerging loss experience.<br><br>Limitations<br>Assumes historic experience reflects current exposures<br><br>Observed data likely to under-estimate true distribution (survivor bias)<br><br>No explicit allowance for underwriting cycle<br><br>ULR's on more recent years are estimates | <u>Assumption</u> regarding the shape (including mean and volatility/ skewness) of the selected, aggregate loss distribution for this class (including attritional and large losses) | Description<br>Confirm fit of historic data points to selected distribution in the body and tail of the distribution. Focus is on tail of the distribution (adverse scenarios impacting SCR)<br>Data required:<br>Historic loss ULR's between 2001 – 2010, unadjusted for rate changes / claims inflation<br>Information from the underwriter and claims team regarding changes in the mix of business over the same period, including rate changes and claims inflation | Expect: | | Results: | |

| # data points | %-ile | # data points | %-ile |
|---|---|---|---|
| <=1 | >90th %ile | 2 | >90th %ile |
| <=2 | >80th %ile | 3 | >80th %ile |
| <=8 | Between 0 and 80th %ile | 7 | Between 0 and 80th %ile |

Comments:

2005 is treated as an outlier, this appears to be inappropriate given book is stable.

Status:

Validation results outside of expectations; estimated impact on SCR < 5%

**Recommended Actions**

- Given likely survivor bias, re-parameterisation of this class required
- Re-review underlying analysis and justification of selected distribution or parameterisation
- Re-perform validation by Q2 2013
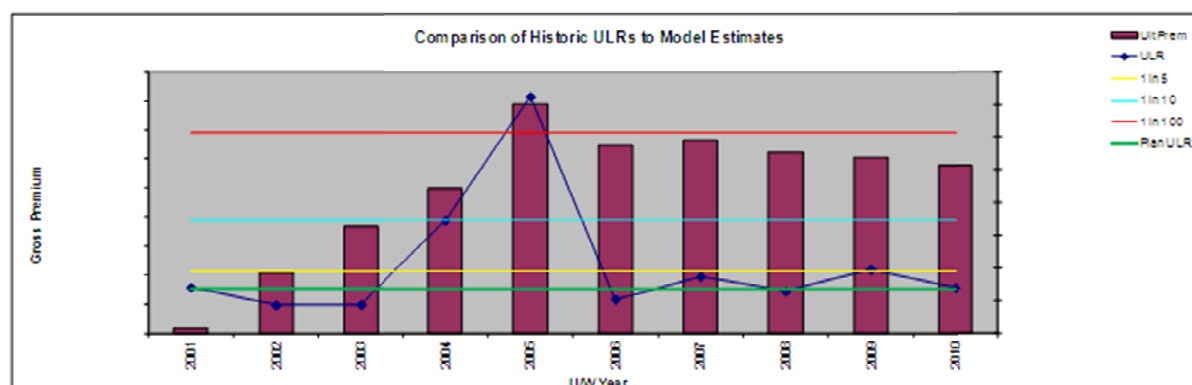
*Independent Expert Review*

**Example: Independent Expert Review**

| Objectives /Limitations | What is being validated | Description of validation | Expectations | Results |
|---|---|---|---|---|
| **Objective** Confirm that methods reflect good modelling practice, with any simplifications or approximations explained and justified.<br><br>**Limitations** Model methods vary widely between insurers, and good practice is not defined<br><br>The findings of the independent expert are subjective and reflect their own experience | Methods used to quantify diversification within insurance risk (between lines of business and accident years) | Review of the modelling approach by an appropriately qualified expert (i.e. qualified actuary with 5+ years' industry experience and 2+ years' relevant modelling experience)<br><br>Confirm through review of model documentation, walkthrough of the model components and Q&A with the modelling team that the methods applied reflect good practice<br><br>Data;<br>The model applies correlations using a Gaussian copula between lines of business and origin years as follows. An algorithm is applied in the model to sample from the marginal distributions; the sampling is based on the selected copula, but scaling is applied to increase the correlations between classes / origin periods with significant volumes of claims. | Expect:<br>In view of the independent expert,<br>- the methods and assumptions applied are appropriate vs current modelling practice, and are conceptually sound<br>- model results are reasonable given the risk profile and current modelling practice | Methods used to model diversification do not appear to be robust. In particular, the algorithm is based on volume of claims (rather than dependency) and does not generate sufficiently strong outputs correlations between origin years and lines of business. The methodology does not show sufficient dependency in the tail of the distribution. |

**Recommended Actions**

- Review of methodology and functional testing of its implementation in the software
- Justification of selected method against others in the market (e.g. tiered vs flat correlation structure)
- Re-validation of the methodology by Q2 2013

*Scenario Testing*

| Objectives /Limitations | What is being validated | Description of validation | Expectations | Results |
|---|---|---|---|---|
| Independent testing of the model design and result, by comparing the modelled SCR to a scenario test determined and quantified independently of the model. | Design of the internal model, and the results | Comparison of the modelled SCR to an independent scenario | Expect: <br>- Qualitative: expect drivers of risk in the scenario to be modelled <br>- Quantitative: expect scenario <£5m different from modelled result at the given return period | Qualitative <br>All risks identified by the scenario tests are also captured stochastically in the internal model |
| Limitations <br>May be difficult to reconcile a scenario test to output from the model; requires an assumption regarding the likely return period of the scenario being assessed. | | | Data: <br>*Model Results* <br>1in 200 i.e. SCR £110m <br>1 in 100 £85m <br>*Scenario Tests:* <br>(A) 2x Realistic Disaster Scenarios e.g. Florida Windstorm and Aircraft Collision Gross £150m <br>Net £75m | Quantitative <br>Difference between scenario (B) and modelled SCR is £20m, however this event is considered to be a 1 in 500 year loss i.e. significantly modelled <br>Modelled SCR adequately covers scenario (A). Given we assess scenario (A) to be a 1 in 100 event, there is a £10m difference between scenario (A) and the modelled 1 in 100 year loss. |
| Quantification of the loss to the insurer under any given scenario is often very judgemental. | | | (B) As above, with default of major reinsurer (50% loss given default) <br>Net £130m | |

| Objectives /Limitations | What is being validated | Description of validation | Expectations | Results |
|---|---|---|---|---|

**Recommended Actions**

- Scenarios A and B are very broadly in line with expectations for the internal model, though we do exceed our target criteria for model results (<£5m difference at the given return period)
- All risks included within the scenario are modelled
- Results should be reported to Risk Committee/ Board for their review & sign-off

## 8.6   Summary

8.6.1   When compared to the ICA regime, solvency II validation is considered to be much more extensive in scope and much more formalised. The requirements for aspects such as documentation and expert judgement to be validated demonstrate the increased scope. The requirements for independent validation and the expectation of the specification of triggers and thresholds with associated procedures further demonstrate the increased scope and the formalisation of this. Formalisation is most clear through the formal requirement for a validation policy and validation report.

8.6.2   A number of tools are set out in the EIOPA texts and in general these are already commonly used, though again may not be as formalised as required under Solvency II. This is demonstrated by the worked examples above, where we apply defined expectations and follow up actions in our analysis.

8.6.3   There are limitations attached to validation especially due the fact that the underlying assumptions and inputs with respect to the internal model are based on estimates.

8.6.4   Validation can add value to the business in various ways.

- Validation ensures that the methodology, assumptions and model results are subject to independent, objective challenge (i.e. dependency on one capital actuary is reduced)

- Model error is reduced or identified

- The validation results enable management and the Board to fulfil their responsibilities under Solvency II, and to better understand the model, key sensitivities and drivers of the SCR result

- Key judgements or sensitivities in the SCR are justified through further review and analysis

- In a business as usual environment, trends in the data or back-testing may indicate that certain methods or parameters are no longer appropriate, and need to be adjusted

- Model limitations and areas for future improvement of the internal model are identified

8.6.5    Whilst we have focussed in our worked examples on validation to support the appropriateness of the internal model and SCR result, validation should also encompass an assessment of compliance with the tests & standards for model approval.  The validation report can be used in this way to give the Board comfort that the model is ready for IMAP.

# 9  Documentation Standards

## 9.1  Introduction

9.1.1    Documentation is a key channel of communication about the internal model to the regulator, the Board and senior management team, the risk management function and other users of the model.  It will form the foundations of the regulator's views, as well as those of any independent reviewers, on the appropriateness and reliability of the internal model.

9.1.2    It is particularly challenging for firms to document the technical details of the internal model when the level of detail and granularity required by Solvency II is still not completely clear.   We discuss the level of documentation required for an 'independent, knowledgeable third-party' to understand the internal model and, in principle, reproduce the SCR result from given a set of inputs.

9.1.3    We believe that creating a good 'culture' of documentation within a firm is vital if it intends to seek internal model approval.  However, if documentation is largely seen as a compliance exercise, it is unlikely to work in a sustainable way for ongoing approval.  It is therefore important to invest the time and effort into getting this aspect right.

*Notes on our scope:*

9.1.4    We note the internal model as defined by Solvency II extends beyond the 'calculation kernel' (or the engine of the capital calculation).  We have assumed the internal model encompasses all inputs into and outputs from the calculation engine, including underlying models or other analysis used to quantify risk (e.g. not only including the calculation of insurance risks, but also the quantification of operational risk within the risk register, etc.) as well as the policies and procedures governing its use.

9.1.5    We note that documentation is also required for other areas of Solvency II (e.g. Pillar 2 and 3, as well as the IMAP submission itself).  Here we focus on documentation of the internal model as defined above.

## 9.2  Summary of Framework Directive

9.2.1    Level 1 Text, Article 125 - Documentation standards:

9.2.2    *"Insurance and reinsurance undertakings shall document the design and operational details of their internal model.*

*The documentation shall demonstrate compliance with Articles 120 to 124.*

*The documentation shall provide a detailed outline of the theory, assumptions, and mathematical and empirical basis underlying the internal model.*

*The documentation shall indicate any circumstances under which the internal model does not work effectively.*

*Insurance and reinsurance undertakings shall document all major changes to their internal model, as set out in Article 115".*

### *Key Requirements*

9.2.3    Paragraphs 9.40 and 9.53 of the Former CP56 set out CEIOPS's high-level view of the documentation requirements, and these have largely been pulled through into the draft Level 2 Text (Article 231).  This states that the technical documentation regarding the firm's internal model must be sufficient such that an "independent knowledgeable third party" can:

- Understand the design and operational details of the internal model; and

- Form a sound judgement as to its compliance with articles 101 (calculation of the SCR), 112 (general approval provisions), 120 to 124 (the five approval standards other than the documentation standard itself) and 126 (external models and data); and

- In principle, be able to reproduce the outputs of the internal model based on the documentation and all of the inputs into the internal model.

9.2.4    We note that the Level 2 text does not define an 'independent knowledgeable third-party'.  For practical purposes, we believe it is reasonable to assume that such an individual is likely to be a financial modeller with the appropriate experience (i.e. experience of modelling insurance liabilities) and qualifications (i.e. mathematical or statistical university degree, but not necessarily a qualified actuary).  Under independent, we assume an individual who is independent of the design and operation of the internal model (as for validation); this could be an internal or external person.

## 9.3   Practical challenges

9.3.1    As a working party we have spoken to a number of insurers, industry experts and have conducted a survey to identify the key challenges facing firms in meeting the IMAP requirements for documentation.  The following were the main areas identified:

- Level of granularity;

- Content  (i.e. what needs to be documented?);

- Key areas of focus;

- Resourcing;

- Documentation change control procedure;

- Document management solution.

9.3.2    We also include a possible framework or checklist of documentation required by former CP56 for the internal model.

## 9.4   Possible solutions

9.4.1    This section describes some examples of issues the Working Party has seen, along with some suggested solutions.

### *Defining Terms and Materiality*

9.4.2    One of the observations we have made is that some terms are used by different companies (or even different individuals within the one company) to mean different things.  Terms such as "parameters" and "assumptions" are an example of this.  Some terms are also specific to a particular company.  We believe it would therefore be useful for a glossary of terminology to form part of the company's documentation, to reduce the risk of ambiguity and confusion for both internal and external users, including within the modelling team(!)

9.4.3    Materiality is a concept embedded within the regulations, but it is not well-defined.  This is related to the concept of proportionality, which again is not well-defined.  During our discussions with industry experts, some have suggested that companies could benefit from having a defined "Materiality Policy".  This would then help to guide the drafting of documentation (including the level of detail), and would facilitate a consistent approach across the company to this difficult issue.

### *Level of Granularity*

9.4.4    Most companies will be producing different levels of documentation, depending on the intended audience.  For example, senior management will be provided with summary/high-level documentation around the internal model, whereas more granular documentation will need to be produced for those involved with the build of the internal model.

9.4.5    Producing documentation at a granular level and then summarising/tailoring as required depending on the audience is the approach that most companies are taking.  This is probably the most practical approach, although it presents the additional challenge of ensuring that the different levels of documentation are kept consistent when changes are made, which implies having good controls around documentation and capturing the interactions between documents.

9.4.6    We believe that it is important to ensure the company has a document which clearly sets out at a high level the methodology and framework of the internal model, as this will be useful to the Board and senior management team so that they understand the model at the highest level, without the distraction of detail.  Such a document would also be a key initial communication tool with the supervisor.  Once the high level methodology and framework is understood, the user will then be able to better understand the more detailed documents which would expand on the various aspects of the model.

9.4.7    Although companies are using a variety of approaches to granularity, we have observed that a number are planning to use three broad levels of granularity in their documentation:

   a) high level "executive" documentation primarily for senior management providing an overview of the model;

   b) a mid level providing more detail (e.g. on a particular area of the model); and

   c) a low level providing full detail on the model.

9.4.8    To illustrate how such an approach could work, let's consider the documentation of the SCR methodology.  The first level of granularity could provide a high level overview of the SCR methodology (i.e. how the SCR methodology works together as a whole).  The second level of granularity could cover details of the SCR methodology for each risk category.  The third level of granularity could then cover the specification of the calculations of the model for each risk category (which would in itself be consistent with the SCR methodology for that risk category set out in the second level of granularity of the documentation).

## Content - what needs to be documented

9.4.9    We have spoken to a number of modellers about what they would need to "form a sound judgment as to the internal model's compliance [with the approval standards]" and  "understand the design and operational details of the internal model".  We believe their views are important, as they potentially represent the expectations of a "knowledgeable third-party" responsible for independent review of the model, either as a regulator or validator.

9.4.10   Based on our research, we concluded the following are key aspects to an independent, third-party review of the model:

- Clear explanation of the model purpose and use;

- Plain English description of the methodology used (including model structure, methodology for setting assumptions, and calculating key output metrics).

- Justification of the methodology and assumptions used in the model and triggers for non-scheduled review, as well as an assessment of alternative methods/assumptions (e.g. sensitivity testing) and the model's strengths and weaknesses;

- Flowcharts and screenshots of the internal model;

- Description of key operational processes including who is responsible for each;

- Summary of the model results (e.g. SCR, broken down by risk category, pre- and post-diversification);

- Summary of the validation techniques used and the validation results (e.g. results of sensitivity, stress and scenario testing);

- Details of processes used to ensure data quality standards.

9.4.11   The following areas of evidence were also thought to be critical to a successful, independent review of the model:

- Access to the model (i.e. ability to inspect the model components and parameterisation spreadsheets).

- Ability to run tests of the model build, or validate the assumptions or results.

- A complete audit trail of the validation carried out on the model.

9.4.12   Our research highlighted that detailed mathematical proofs were not thought to be particularly useful for independent reviewers from a practical perspective where the methodology used was in line with industry practice.  Where relevant professional or research papers were available in the public domain to support the use of those methods, such aspects should be referenced within the documentation for audit trail purposes, but were unlikely to form a key part of the documentation from the reviewers' perspective.  On the other hand, non-standard methods were thought to require further documentation.

9.4.13   There is tension between the practical implementation of the documentation standards and some of the requirements.  One of the key aspects of this tension surrounds the "reproducibility" criterion set out in the Former CP56 9.40: "The documentation … shall be … sufficiently detailed and sufficiently complete … that an independent knowledgeable third party could … in principle reproduce the model outputs if all the parameters and exposure data were available".  This concept has been carried over into the draft Level 2 Implementing Measures produced by the Commission (Article 231 (2) TSIM20): "Outputs of the internal model shall in principle be reproducible using the internal model documentation and all of the inputs into the internal model."

9.4.14   The level of detail implied by a strict interpretation of this is significant, and we believe that some pragmatism will be required (drawing on the "in principle" wording) to achieve a level of documentation which is proportionate to the overall significance of the aspects of the model being documented. One industry practitioner suggested defining "reproducible" as "producing outputs which are not materially different to the model results".  This is another example of an area in which we feel that a company having a defined Materiality Policy would be helpful.

## *Key areas of focus for firms*

9.4.15   Close to 80% of the companies we surveyed said that a substantial re-write of existing documentation will be needed.

9.4.16   Areas where a substantial re-write may be required were mainly around:

- Validation documentation;

- Methodology documentation;

- Process mapping/flowcharts.

9.4.17   Interestingly, recent feedback from the FSA has indicated that validation documentation is indeed a key area of interest to the regulator, and an area in which they have identified significant deficiencies in the documentation many companies have provided so far.

9.4.18   Somewhat surprisingly, the areas where some respondents said that there would only be minor changes to documentation related to Use Test aspects such as:

- Executive summaries for senior management and users;

- Training material for senior management and users;

- Outputs/management information;

- Evidence of Use.

9.4.19   Most companies that responded to our survey were also planning to map their documentation to the Solvency II Directive, though a significant minority (approximately 25%) were not planning to do this. However, with the introduction of the new Self-Assessment Template from the FSA in February 2012, it is clear that the FSA expects firms to have a strong link between their documentation and the specific regulations the documentation satisfies.

9.4.20   Our research also highlighted that firms were facing particular challenges in the following areas:

- Expert Judgement;

- External Models;

- Validation.

9.4.21   We discuss these in more detail in the relevant sections of this report.

*Resourcing*

9.4.22   There is clearly a significant amount of documentation to be produced and a limited amount of time left within which to produce it. This was further underlined by communication from the FSA's Insurance Director in May 2012 about areas in which existing documentation was insufficient.

9.4.23   Many have underestimated the effort required to produce good quality documentation. , and this could cause some companies significant issues as their internal model approval submission dates approach.

9.4.24   It is likely that actuarial teams (e.g. the capital modelling team) may 'own' some of the detailed documentation around the internal model, but also utilise other resources to produce it.

9.4.25   Some of the larger insurers have been using a combination of actuarial/technical writers/risk management staff to produce documentation.

9.4.26    Some firms are using technical writers (who are not Solvency II or modelling experts, but are documentation specialists) to provide a framework on the standard of the documentation.  The subject matter experts (e.g. risk management, or the capital modelling team) then provide the detailed content required in the documentation based on this framework.

9.4.27    Whilst some firms are planning to utilise external consultants to produce/validate their Solvency II documentation, it was generally felt there are insufficient resources (internally and externally) with the appropriate skill-set to write good quality documentation.  For example, many actuaries are not necessarily the best at writing in a style that is easily understood by others), and this may compromise the overall quality of documentation.

### *Documentation change control procedure*

9.4.28    It is important that documentation is kept fully up to date as this reduces key-person risk (as well as being an ongoing requirement to meet the documentation standards).  We believe that this will be challenging.

9.4.29    Companies will need to put in place a well-defined process to ensure that documentation changes/updates are put through properly and that there is a strict version control process around this.

9.4.30    As part of any model change control procedure, we believe that the relevant committee within the insurer that has oversight of the change to the model (including input data) should also review/sign-off that all relevant documentation has been updated for the change in the model.  This should form part of a well-defined documentation ownership structure that is linked to any document management solution that is used by the company.

### *Document management solution*

9.4.31    A variety of software solutions are available to firms to organise and store their Solvency II documentation.

9.4.32    SharePoint appears to be the preferred document management solution at many insurers (around 2/3rds of the companies we surveyed are planning to use it).

9.4.33    SharePoint has many features that will be useful for Solvency II purposes.  These include version control as well as the ability to add categorisation fields to the documentation in order to map it to the Solvency II Directive.

9.4.34    Other options being considered include:

- Conducter (which includes a mapping to the Solvency II Directives);

- Quality Workstream – this is an HTML based documentation controller, which shows where documents are saved (via links) and whether changes have been made to a document and where it currently is in the approval/sign-off process;

- Nimbus;

- K Center;

- Documentum;

- ARIS;

- BPM One.

9.4.35    As well as the above, some insurers are developing bespoke in-house platforms.

## *Possible Documentation Framework*

9.4.36    The following is a checklist of key items referenced in CEIOPS' Level 2 advice, which may need to be documented.  Although it may not be exhaustive, we hope this checklist is useful to firms in scoping the documentation of their internal model.

**Data**
- Data Management approach
- Clear data dictionaries
- Description and construction of the databases
- Data flow chart covering internal model
- Assumptions for Data
- Process and controls for data, data flows and data quality
- External and Internal data
- Etc.

**Assumptions and Parameters**
- Summary of methodologies and formulae to estimate parameters
- Sources of data backing assumptions
- Expert Judgement
- Etc

**Technological Specifications / Systems**
- Description of the Information Technology platform used in the internal model
- Contingency plan, security policies and business recovery plans for the technological elements of the internal model
- User guide
- Source code
- Etc.

**Calibration Standards**
- Risk measures & time periods for different business units and justification of these
- Demonstrate consistency between SCR calculation and internal model
- If SCR cannot be directly derived from probability distribution then
  - How risk is rescaled and justify how bias is immaterial
  - Shortcuts used to reconcile outputs of internal model with distribution of Basic Own Funds
  - Etc.

**Calculation Kernel – (Methodology)**
- Detailed outline of the theory, assumptions and mathematical and empirical basis underlying the internal model
- Technical Provisions – best estimate and risk margins
- Capital/Solvency Requirements
  - Risk in scope/Out of scope
  - Business Units In scope/Out of scope
  - MCR
  - SCR
  - Recognition of risk mitigation instruments
- Aggregation policy and methodology
- Overview of the historical development of the internal model
- Simplifications/Approximations
- Etc.

**Use Test**
- Evidence of Use Test i.e. Integration of model within the business
- Senior management understanding of model
- Etc.

**Internal Model Governance**
- Policies & Standards
  - Validation Policy
  - Model Change Policy
  - Documentation Policy
  - Calibration Standards, Etc.
- Controls and Procedures
- Responsibilities and accountabilities
- Drawbacks and weaknesses
- Etc.

**Profit and Loss attribution**
- Profit and Loss Attribution Policy
- Results of Profit and Loss Attribution
- Material risks in the risk profile not represented by the internal model
- Etc.

**External Models and Data**
- Role and extent of use
- Decision/Rationale for choice of particular external model or data
- Demonstration of detailed understanding and knowledge of external models' and data's:
  - Methodological underpinnings
  - Basic construction
  - Capabilities
  - Limitations
- Demonstration of appropriateness in relation to:
  - Nature, scale and complexity of risks
  - Business objectives
  - Modelling methodologies
  - Availability of internal data
- Validation of External models and data
- Risks arising from use of external data and models e.g. strategic risk, contractual risk, etc.
- Etc.

**Expert Judgement**
- Description of where Expert Judgement is applied in the model
- Justification of use of Expert Judgement where used in the model
- Validation of Expert Judgement
- Etc.

**Validation Policy**
- Purpose and scope of validation
- Validation tools used
- Frequency of validation process
- Where, if anywhere, external review and systems are used
- Testing results against experience
- Analysis of Change
- Actual versus Expected
- Etc.

**Documentation Policy**
- List of all relevant documents and how these can be accessed
- Identify people responsible for maintaining documents
- Overview of historical development of the internal model including Methodologies, Assumptions and Data
- Version control process of internal model
- How requirements governed in Articles 120 to 124 have been taken into account and fulfilled
- Limitations in risk modelling
- Nature, degree and sources of uncertainty
- Deficiencies in input data
- Documentation Index
- Model Scope
- Etc.

**Statistical Quality Standards**
- Detailed description of Internal Model Methodologies and parameterisations
- Description of underlying assumptions
- Risk ranking and drivers of risk
- Etc.

**Internal Model Output and Reporting**
- Supervisory and external reporting
- Report to Supervisor (RTS)
- Solvency and Financial Condition Report (SFCR)
- ORSA – Economic Capital
- Internal reporting
- Etc.

**Model Change Policy**
- Definition of a major and minor model change
- Etc.

# 10 External Models – Catastrophe Modelling

## 10.1 Introduction

10.1.1    Companies often use external models, provided by third parties, within their Internal Models. In this section, we will consider catastrophe (CAT) models as an example of an external model.  External models need to satisfy the model tests set out in articles 120 to 125; that is, the Use Test, Statistical Quality Standards, Calibration Standards, Profit and Loss Attribution, Validation Standards and Documentation Standards.  Excerpts from the Level 1 and 2 texts, which are applicable to external models, can be found below in section 10.2.

10.1.2    From the survey[29], 93% of the companies who responded said that they were using CAT models.  In total, 44% of surveyed companies use one CAT model, 19% two models, 26% three models and 4% four models.  Responsibility for the use of CAT models in Solvency II varies across the companies surveyed; with 50% of companies assigning responsibility to their Capital Modelling team, 25% to their CAT modelling team,  4% to their underwriting team and 14% to other teams.

10.1.3    The main catastrophe models used by firms are: RMS, AIR and EQECAT.  Their approach to modelling differs and each vendor model has both strengths and weaknesses.

10.1.4    Management's review of catastrophe models and their use has historically focussed on pricing and monitoring the aggregation of claims.  As a working party, we have seen increasing focus on the use of catastrophe models in capital modelling (for example,  consideration adjustments made to model output to reflect a firm's specific risk profile or pricing strategy).

10.1.5    We acknowledge that there have recently been excellent publications released by the ABI[30] and Lloyd's[31] around the use of catastrophe models.  We do not seek to repeat the guidance issued here, but instead explore some of the practical challenges and give some examples to help firms implement their recommendations.

### Comments on our scope

10.1.6    We note that the requirements set out in the directive apply equally to other types of external models (for example, economic scenario generators).  However, the specific issues may vary (e.g. ESG's can often be validated by comparison to market prices).  In this paper we focus on the specific issues of validating catastrophe models as these are often most material to a non-life firm's SCR.

---

[29] This may reflect bias in our survey sample towards the largest insurance and reinsurance firms.
[30] ABI, 6th December 2011, "Industry Good Practice for Catastrophe Modelling"
[31] Lloyd's of London, 9th July 2012, "External Catastrophe Model Validation: Illustrative Example (US WS, High Materiality"

10.1.7    We have focussed our work on companies who license catastrophe models from third-party vendors directly, and on the challenges they face in meeting the Solvency II requirements.  We note many other insurers rely on third-parties to license and run external models on their behalf, often brokers.  We believe however, that many of the principles set out below would apply equally to the approval of other models, including those of companies not licensing directly a catastrophe vendor model but relying on a broker to provide results.

10.1.8    We note that any external model, including catastrophe models, must be subject to the firm's change policy and change control environment.  Major or minor changes to the external model (e.g. release of    RMS version 11) should be monitored within this environment, and reported to the regulator where necessary.    This poses particular challenges for insurers and regulators where upgrades to external models are controlled by the third-party vendor and have a market-wide impact on capital.

## 10.2 Summary of Framework Directive

10.2.1    Under the Level 1 Directive, use of an external model to quantify catastrophe exposures is not an exemption from any of the other tests and standards for model approval (Art. 120 to 125).

10.2.2    Article 121 of the Level 1 text includes the following statements which are particularly relevant to CAT models.

•    "The methods used to calculate the probability distribution forecast shall be based on adequate, applicable and relevant actuarial and statistical techniques, and shall be consistent with the methods used to calculate technical provisions."

•    "Data used for the internal model shall be accurate, complete and appropriate."

•     "Insurance and reinsurance undertakings shall update the data sets used in the calculation of the probability distribution forecast at least annually."

10.2.3    In addition, Article 124 requires firms to have a regular cycle of validation in place over catastrophe models; and Article 125 requires firms to document the design and operational details of external models, including any circumstances where the catastrophe model does not work effectively.

10.2.4    The Level 2 draft includes more detailed requirements:

•    Detailed knowledge of the underlying methodology and structure of the model, including an understanding of the models' capabilities, limitations, and appropriateness for use in deriving the SCR. (Article 213 TSIM3)

•    In relation to the data used in the internal model, insurance and reinsurance undertakings shall establish, implement and maintain a data policy which covers at least...:

a) the definition and the assessment of the quality of data, including specific qualitative and quantitative standards for different data sets, based on the criteria of accuracy, completeness and appropriateness;

b) the use of assumptions made in the collection, processing and application of data; data updates, including the frequency of regular updates and the circumstances that trigger (c) additional updates and recalculations of the probability distribution forecast.  (Art 220 TSIM10)

c) Insurers must be able to explain the role of any external models in their internal model, justify the use of an external model against alternative approaches, and monitor compliance of the internal model with Articles 110, 112 and 120 to 125. (Article 235 TSIM24)

## 10.3 Practical Challenges

10.3.1    As part of the research, the working party asked which areas of the Solvency II tests and standards proved to be most challenging in applying to external CAT models. The following main areas were identified:

- Level of external model documentation to produce

- Demonstrating a detailed understanding of CAT models outside the CAT team, specifically understanding of  the risk officer and the board;

- Model methodology and parameterisation;

- Justification of loadings applied to CAT model results; including allowing for the roll-forward of exposure data, blending results from different vendor models, and adjustments for non-modelled exposures.

- Validation of model results;

- Strengths and limitations of the CAT model;

- Governance of the CAT model

10.3.2    We summarise in the diagram below, the key Solvency II challenges applying to CAT models:

## Catastrophe Models
*Key modelling steps and practical challenges*

| Exposure Data (ie policy data) | External Model (eg RMS, AIR, EqeCat) | Adjusted outputs (eg adjustments for non-modelled risks, blending between models) | Use of results (ie aggregate exposure monitoring, calculation of SCR) |
|---|---|---|---|
| • **Statistical Quality/ Data Quality** demonstrating quality of exposure data, adjustments for known weaknesses in the data | • **Statistical Quality/ Methods & Assumptions** Showing external model meets SQS requirements, justifying selected third-party model<br><br>• **Documentation** what documentation will be produced by the vendor, what is required in-house? | • **Statistical Quality/ Expert Judgement** what adjustments are appropriate? How to justify them?<br><br>• **Validation** of model results e.g. Sensitivity testing, Scenario Testing, Back Testing | • **Governance** Demonstration understanding of third-party model and its limitations |

10.3.3   We found in interviews that there is often a large reliance on the CAT team to produce the documentation around CAT models.  It therefore appears that currently knowledge on CAT models is very much concentrated within the CAT team and internal education will need to be given in the future to propagate this knowledge to the risk team and the board.

## 10.4 Documentation

10.4.1   We found that firms faced the following key challenges in the documentation of CAT models:

• What level of documentation can firms rely on the third-party vendor to produce (either on a confidential basis, e.g. shared with FSA but not with other firms, or with firms as part of their license agreement)?

• What level of documentation should be produced internally?

10.4.2   This is consistent with the ABI report[32], which discusses good practice for two forms of documentation:

• The sections of a company's Internal Model documentation that cover the catastrophe model, in accordance with Article 125, and

• Documentation that catastrophe model vendors may provide to help the company understand and use the catastrophe model.

10.4.3   As well as resourcing and the knowledge gap between firms and CAT model vendors, we found that a key block to firms producing internal documentation has been the lack of clarity around the level of documentation required by regulators compared to that already produced by model vendors.

**Vendor Documentation**

---

[32] Chapter 3, "Catastrophe modelling Documentation", Industry Good Practice for Catastrophe Modelling (December 2011)

10.4.4    Each of the three vendor CAT modelling companies have Solvency II documentation which gives information on key components needed for Solvency II reporting, such as model construction and validation. There is usually separate documentation on the construction of stochastic event catalogues (probabilistic simulations of potential events which can occur in a given year). Validation of the model is also usually provided, sometimes on a component basis, and in other cases on an overall loss basis. This information should be the first point of reference when compiling Solvency II documentation.

10.4.5    However, in the interviews, firms noted they would like to see further detailed information from vendors, including the parameterisation of vulnerability functions and the formulation of secondary uncertainty distributions. We note this information is typically available to licensees on request.

10.4.6    The three vendor modelling companies require some form of non-disclosure agreement for the release of their documentation. This is usually specified in the license agreement, or upon firms requesting the release of information to regulators. Confidentiality of the intellectual property is an important issue for the preparation of the necessary documentation.

10.4.7    The ABI advises firms should request the following information from vendors:

- Version control

- Model history, including changes

- Methodological approach, including geographical, hazard, vulnerability, building type/ construction and financial information (e.g. reinsurance structures), options & settings

- Validation results

- Knowledge of the model strengths and limitations

- Uncertainty

- Systems/software

**Company Documentation**

10.4.8    The ABI suggest[33] that companies document the following:

- Demonstrating understanding, including documents provided by the vendor, training and conferences attended by individuals within the company, records of meetings and correspondence with the vendor

- Demonstrating Operation, license agreement, use and management of data, reasons for model selection, history of model changes, model methodology and limitations, model validation performed by the company on the model results, and finally, governance and use of the model results (including justification of loadings applied to model results).

---

[33] Para 3.3.1 and 3.3.2

- In interviews, we found that some entities were producing CAT model documentation based on their own understanding of the model, as required by Solvency II.

10.4.9    Some were simply referring to specific sections of the vendor's documentation.  As a key  requirement under Solvency II, is for firms to show an understanding of the vendor's CAT model documentation, we felt this approach was insufficient to meet the requirements.

10.4.10  We found firms were mindful of materiality and proportionality. They therefore focused efforts on key lines of business which impacted on their Solvency Capital Requirement.

10.4.11  This said, different regulators require different levels of documentation and information on CAT models.  From talks to regulators from other European countries, not all of them understand the models at the same level, or require the same information to be disclosed.

## 10.5 Validation Practices Overview

10.5.1    The Solvency II Directive does not specify the validation procedures required for external models.  In practice, there are two main types of catastrophe model validation, as mentioned in the ABI paper [34]:

- Vendor model validation to ensure that results are appropriate for the specific peril at a country-wide level

- Individual company validation to ensure that the model is suitable for its actual portfolio

### *Reviewing Vendor Model Validation*

10.5.2    CAT models can be validated **against losses from historical events**.  The first port of call for information on the validation performed should be made by referring to the relevant vendor documentation.

---

[34] Chapter 7, "Catastrophe model validation"

10.5.3    Validation is also often done through **internal review at the vendor company or through an external peer review**.  The vendor company can provide details of the relevant peer reviewers if applicable.  Furthermore, often an existing body of research is used in the creation of the stochastic catalogues and intensity calculations used by vendor modelling companies.  For example the USGS Next Generation Attenuation Equations published in 2008 are used by all vendor companies and studies into Japanese Seismicity by HERP (Headquarters for Earthquake Research Promotion) have also been adopted by the vendor companies.  These two examples of research can be viewed as having been peer reviewed by the scientific community and thus validated.  Finally, the external data used to calibrate the CAT model is often given in the model documentation.  In many cases this data (excluding company claims data) will be public and available to be compared against the vendor assumptions (e.g. using USGS data to compare historical earthquake frequencies against modelled frequencies for a particular region).

## *Individual Company Validation*

10.5.4    Most companies interviewed performed validation on the outputs of their CAT model for business planning, although they might not currently label this as 'validation' for Solvency II. For example they may compare the results to their own loss data for an historical event, or check their exposure data against built in exposure data. They can also check historical footprints or wind speeds with the ones used by the model. The purpose is generally to make sure the model makes sense for their portfolio of exposures.

10.5.5    Some companies delved into details such as looking into catalogue size (described below) and performed detailed checks on model results for example, comparing the clustering of European Extra Tropical size by intensity.

10.5.6    In some cases a company would commission a scientific study to look at a particular region/peril of interest to them, to judge whether the vendor model appropriately represented the peril under consideration, or needed to be adjusted.

10.5.7    Some specific examples of validation activities performed by individual companies are discussed below.

## *Examples of Validation Practices*

**Exposure Data Validation**

10.5.8    Since external model parameters may have been calibrated using external data, it is critical for undertakings to test the performance of external models against their own portfolio of exposures.  Where there is a scarcity of internal data, undertakings' reviews will have to rely more on alternative validation techniques, such as expert judgments. It is important that undertakings periodically challenge external models and data using their own models and data, even if they are not as sophisticated, accurate or exhaustive.

10.5.9   Generally, companies have several levels of exposure data validation checks. Typically these are designed to ensure that all relevant exposure data is captured within the model.

10.5.10  As noted by the ABI, firms should validate compare the quality of their own data to the key data items required by the model[35]. For example, shortcomings in company's data might include:

- no location; and if this is the case, the selected distribution of exposure data, for example, country-wide values

- unknown or limited exposure data, versus vendor model industry data or market data

- unknown secondary modifier; and vendor view on LOB, structure type, and occupancy

10.5.11  The data entered into the model will usually be held in Microsoft Excel, SQL, an internal bespoke tool or an external tool such as Exact. Checks are often made between the database value and the value actually used in the software.  Typical metrics compared before and after include; the total insured value (including a breakdown between modelled and non-modelled), total premium, geo-coding summaries and so on.  This will assist in judging firstly, whether the total insured value is as expected and secondly, how to compare the modelled amount with the results, since adjustments may need to be made depending on the proportion of non-modelled exposure.

10.5.12  In some cases, a company might have an independent review of a subset of the portfolio to judge whether the replacement values and characteristics are representative of the underlying exposures.

10.5.13  In some cases, an external tool such as those produced by vendor modelling companies, e.g. AIR's TruExposure and RMS' Data Quality Toolkit, could be used to perform checks to validate the data and possibly to augment the data with information from a database of characteristics of individual properties. Modellers can provide advice if needed.

**Roll-forward of exposure data**

10.5.14  Often external models are run based on the previous year's exposure data, with some adjustments (if necessary) regarding the change in exposures from one year to the next. Assumptions regarding changes in exposure are made by  wide range of parties (including underwriters, actuaries, data managers) according to the company policy.

---

[35] Para 7.4.2, Model validation

10.5.15  On facultative or primary business, the analysis is usually carried out periodically, maybe on a monthly basis, to cover for variations in exposure. In order for a company to monitor its capacity and accumulations of the current year, the analyses are "frozen" at a point in time and "added" to the company portfolio. On reinsurance and retrocessional books, these adjustments are necessarily more judgemental.

10.5.16  Further validation and/or governance over adjustments made to data, to allow for changing exposures, is likely to be required by regulators under Solvency II.

**Historical Loss Validation**

10.5.17  We found that key validations looked at historical losses and compared these with modelled losses.  Other techniques looked at single event exposures in detail (e.g. Lloyd's Realistic Disaster Scenarios) and judged whether the modelled losses seemed sensible given own expert judgement, policy details and experience of previous historical events.  Other forms of loss validation which could be performed included benchmarking modelled losses with losses reported by centralised claims reporting agencies, such as the Property Claim Services (PCS) in the US and PERILS in Europe.

10.5.18  The validation of CAT models through historical losses is not straight forward due to the hypotheses used to develop the models. For example a model is not created to represent a particular storm (eg, Lothar - a historical European windstorm in 1999 affecting France and Germany), but instead is created to represent European windstorm phenomena.  Thus if we validate the model for Lothar, we will probably not get a 100% match for this particular storm.  Furthermore, real time loss validation can be done using either event sets produced by vendor modelling companies to replicate the actual parameters of the event itself, or similar events in the vendor's stochastic catalogue.  Due to uncertainties in estimating event parameters (like exact focal depth of an earthquake or maximum wind speeds at specific locations) there are often several possible scenarios and this is one source of uncertainty that can feed into loss results.  Each modeller has their own modelling approaches and assumptions, as well as standards for an acceptable representation of  phenomena.  In general when performing historical loss validation, we recommend the undertaking attempts to define a level of acceptable divergence (for example, a +/-20% difference between modelled results and actual losses triggers an investigation of the difference).

10.5.19  There are also other perils/regions where there is no historical data to validate with. A possible form of validation here would be for undertakings  to perform a reasonability check on the footprint of events, with regards to the size of loss, shape of footprint and damage ratios by line of business of affected risks.

10.5.20  Information regarding how model vendors perform validation should be within their documentation.  Model vendors can guide or advise undertakings on how to validate the model, but this will depend on what firms need or want to validate.

**Catalogue Size**

10.5.21   Sometimes catalogue size may be deemed insufficient to cover low frequency regions for some vendor models and hence larger catalogues or other methods of artificially increasing catalogue size are needed.  This is to ensure that the loss statistic used by the CAT model for the Solvency Capital Requirement is stable and not volatile year on year if the underlying exposure geography varies slightly.  This can  be done in several ways, all of which aim to resample from the original catalogue to give a catalogue of sufficient size where the appropriate statistics (e.g. Average Annual Loss and loss at specified exceedance probabilities) are stable. Sometimes statistical methods are appropriate, but in other cases when there are insufficient events in a certain region, then a large event set with more distinct events is needed.  It is worth noting, that some vendor modelling companies employ a constrained sampling technique, to obtain smaller samples (e.g. 10,000 simulated years) from larger catalogues. Under this technique, the catalogue statistics (e.g. AAL, return periods and so forth) are stable when comparing the smaller catalogue to the larger catalogue.

10.5.22   Whichever approach is taken by the Company, this should be justified/explained for Solvency II.

**Non-modelled exposures**

10.5.23   A key challenge in using CAT models is to understand exactly what is modelled and what is not.  To some extent this already exists in modelling companies' documentation.  Nevertheless feedback has been given that an exhaustive list of what is modelled and what isn't should be included for each region/peril.  If this has not yet been published, such information, once requested, should be provided by vendor modelling companies.

10.5.24   Depending on the vendor model, non-modelled exposure could include non-modelled lines of business (e.g. industrial facilities losses to a peril such as European windstorm), non-modelled coverages (e.g. Sue and Labour in modelling tropical cyclone risk to offshore assets), non-modelled perils (e.g. landslide following earthquake) and non-modelled exposure due to insufficient resolution to import into a CAT model (e.g. country resolution, where the country where the risk is located is known but no other geographic information is given).

10.5.25   Furthermore, non-modelled perils then need to be adjusted for, using some other means.   For example, one means is using statistical methods such as a claims-based model to attach the non-modelled loss to a modelled loss (e.g. externally modelled hail loss with vendor modelled European extra-tropical cyclone loss).  Even in cases where no claims data exists, assumptions on the shape of the statistical distribution, average claim size, tail length and so forth can be used to account for losses in regions for which no catastrophe model output is available.  Another method (not necessarily endorsed by this working party) of modelling data for non-modelled lines of business or countries was to map the un-modelled to a modelled line of business, e.g. auto to property.

10.5.26   Another technique was to apply an assumed percentage amount to the total gross exposure in a region. This would represent the proportion of exposure damaged at a specified return period.

10.5.27   Non-modelled aggregate exposure, which was of an insufficient resolution to model, was found to sometimes be redistributed countrywide such that the total value was modelled, e.g. redistributing 5% of UK exposure to European Wind across the country-wide exposure, since some data could not be imported due to insufficient or incorrect geocoding information.

10.5.28   Again, whichever approach is taken, these judgements will needs to be rationalised and justified to the regulator.

**Sensitivity Testing**

10.5.29   Sensitivity tests in general are carried out in the CAT model, and then fed into the capital model to observe the overall effect on the company's SCR.

10.5.30   The specific shape of the vulnerability functions can be obtained either through documentation or testing of the vendor modelling software.  This information can be used to assess the sensitivity of model output to changes in construction/occupancy classes, as well as model over/under estimation of the hazard (intensity) characteristic.

10.5.31   Apart from this, firms often perform sensitivity testing via changing construction/occupancy characteristics and replacing values of the exposure in their portfolio to observe the effect on modelled output.

10.5.32   Further sensitivity tests used are; switching options on and off in the software (e.g. Demand Surge), testing the results of alternative event catalogues (e.g. Hurricanes in the North Atlantic) and granularity of data (e.g. aggregate data vs disaggregating the data to an underlying grid).

10.5.33   Another sensitivity test found was to output results by event and region/peril basis. Then apply a factor chosen randomly from a set (for example 50%, 75%, 100%, 150%, 200%) to one region and peril.  Repeat for all regions and perils and then combine the losses to give the overall EP curve.  Repeat this multiple times to give several different realisations of the combined EP curve to understand the impact of "model miss" on the loss results.

10.5.34   Sensitivity to different model results (e.g. provided by brokers) can also be investigated.

10.5.35   For Solvency II, the scope and results of the sensitivity testing should be summarised and explained to management; to enable them to fulfil their responsibilities for oversight of the model and key underlying assumptions.

## 10.6 Use and Governance of Results

10.6.1   As mentioned before, it is important for Senior Management to have a good understanding of catastrophe models, key sensitivities and what the results mean for their organisation.

10.6.2   In our opinion, this implies that they need to understand several issues, some of which are given below:

- Basic Model components;

- Modelled perils and exposure;

- Quality of the data used for the analysis;

- Limitations and weaknesses of the models;

- Position relative to other vendor models (more or less conservative results);

- Adequacy of the model to company's portfolio;

- Expert judgement applied in loading vendor results;

- Model output;

- Validation and sensitivity test performed to evaluate the model;

- Understanding of any movements in model results year-on-year (analysis of change).

10.6.3   In addition, the ABI[36] note that senior management's responsibilities extend to actively seeking the levels of information and detail needed to make decisions with regard to the catastrophe model, and ensuring proper processes and procedures are in place for operation, governance and use of model results (including validation).

---

[36] Para 1.3, Senior Management

## 10.7 Multiple Models

10.7.1    We found that often when multiple model results existed, they were blended in some form. Either the results from one vendor model were adjusted to mimic the other vendor model or the output was combined at event level.  In the latter method, weights were applied either in equal amounts (e.g. 50/50 or 33/33/33) or using some expert judgement from scientific investigations or professionals such as actuaries.  Such weights were typically applied at event level to losses post location (but pre-layer) terms.  Loss output would need to put into the same format (i.e. rate or year) and in some cases for one model a different weighting was applied to lower layers than higher layers indicating an increased reliance for one model at low return periods.

10.7.2    Though challenging, any blending method used will need to be justified as part of the internal model application and kept consistent year-on-year.

## 10.8 Loadings applied to model results

10.8.1    Though challenging, Solvency II requires justification of the loadings applied for capital calculation on a 'best-estimate' basis.  Companies have their own pricing strategies and these usually dictate how they use the models, and the level of loadings applied.  We note loadings may vary between pricing, aggregation monitoring and capital calculation purposes.  We have not seen insurers make as much progress as might be required expected in this area.

10.8.2    Companies that make the choice of adding a factor or using a specific method to cover the non-modelled perils/LOB should explain why and how they did it.  They may also make reference to the results of other validation (e.g. results of validation against historic experience or knowledge of their own portfolio).  If the company has asked their model vendor or an external expert for guidance, this should be added to the documentation and any loading factors used must be justified when submitting an application for model approval.  If the model has been changed due to the loading, a recalibration or validation of this model may be necessary after the loading is applied to make sure the model still makes sense.

# A   Appendix A :  Output from 2011 Industry Survey

In this appendix we summarise the results of the 2011 ESurvey.

**Please note we received 44 responses to the survey.  This is unlikely to be a representative sample of the UK market's preparedness for Solvency II.**

| Question  1 | Type of Company |
|---|---|
| **Response Options** | <ul><li>Lloyd's/London Market</li><li>UK General Insurance Company</li><li>Multinational General Insurance Company</li><li>Bancassurer</li><li>Composite Insurer (i.e. Life and Non-Life)</li><li>Reinsurer</li><li>Other</li></ul> |

Below is a pie chart indicating the breakdown of the types of companies involved in the survey. Note that 'Other' responses included an overseas general insurance company and a catastrophe risk consultancy.
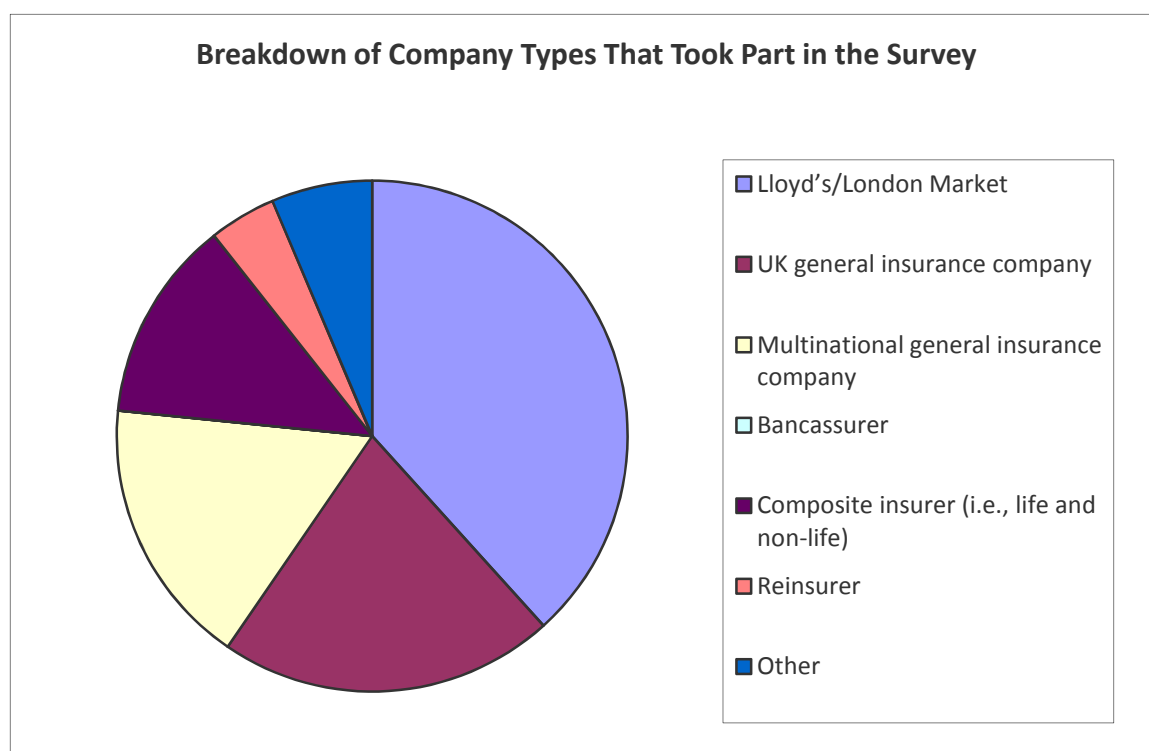
| **Question 2** | **Annual premium income (gross, after commissions) – approximate figures only. If you are a group/multinational firm, please enter the amount likely to fall under the Solvency II remit.** |
|---|---|
| **Response Options** | <ul><li>£0 to £50 million</li><li>£50million to £200million</li><li>£200million to £500million</li><li>£500million to £1billion</li><li>£1billion +</li></ul> |

Annual premium income of the respondents is given in the pie chart below.
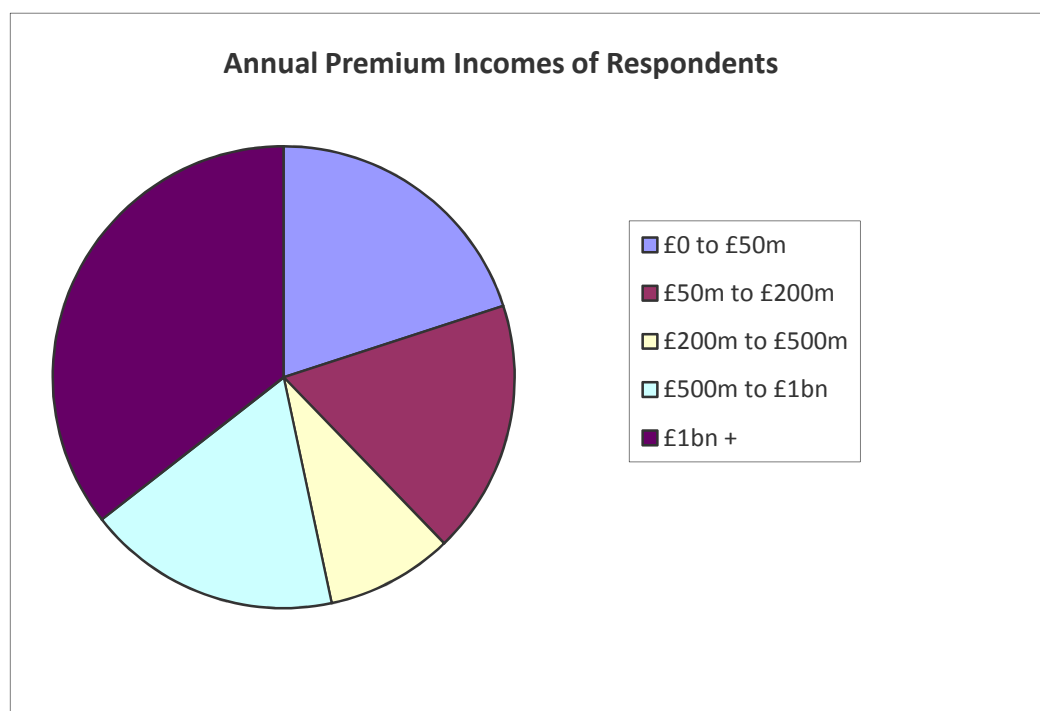


The survey covered a wide range of areas relating to Solvency II implementation.

| **Question 3** | **Resourcing (current approx headcount, including contractors). Total (all projects for the company).** |
|---|---|
| **Response Options** | <ul><li>Risk Function (Including Actuaries)</li><li>Actuarial Function (Including Students)</li><li>Other</li></ul> |

In terms of resourcing, the majority of respondents will make use of the Actuarial and Risk function to a great extent, for all projects and Solvency II in particular, as well as a smaller percentage of other function areas (see charts below).

**Total Resource Usage of Respondents**
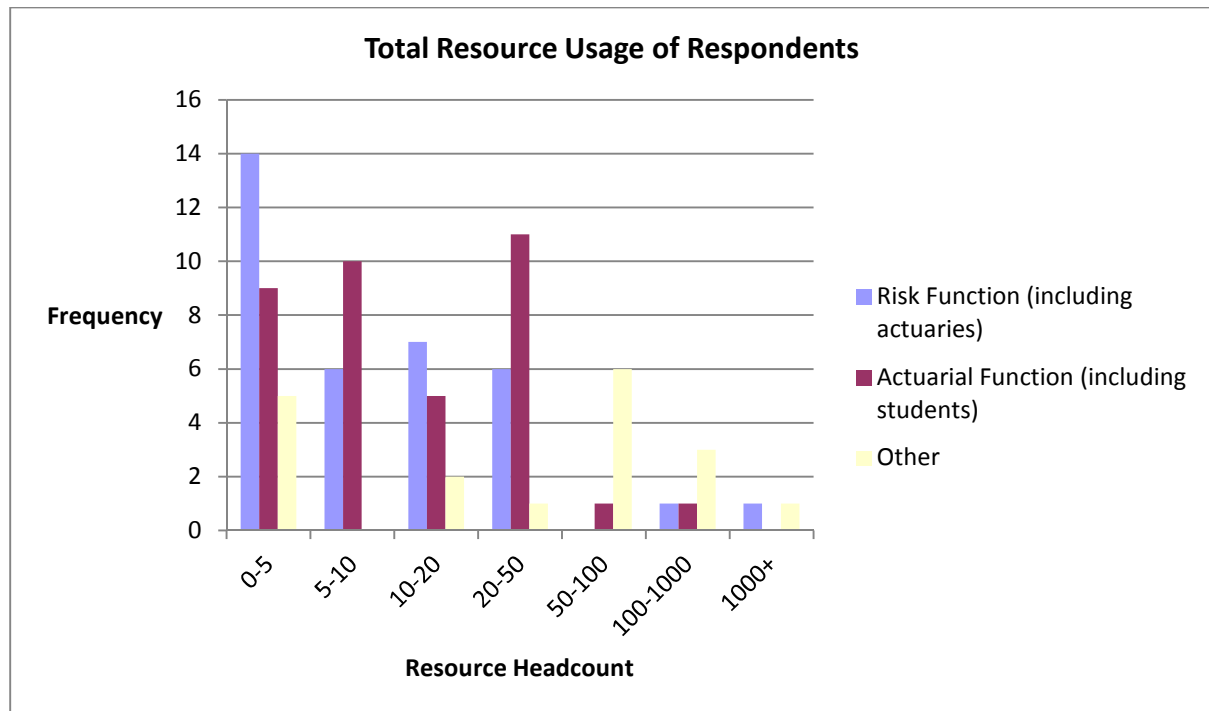
| Question 4 | Resourcing (current approx headcount, including contractors). Total for Solvency II. |
| --- | --- |
| Response Options | • Risk Function (Including Actuaries)<br>• Actuarial Function (Including Students)<br>• Other |

**Total Solvency II Resource Usage of Respondents**
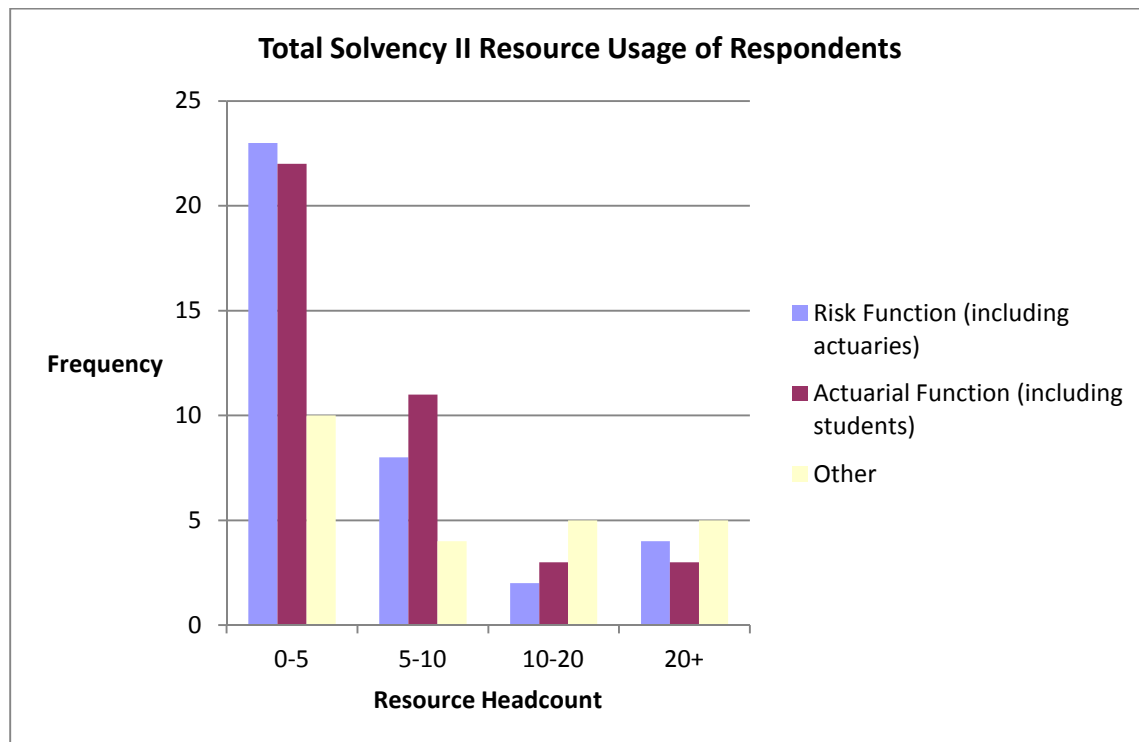
| Question 5 | Do you intend to apply for full or partial model approval? |
| --- | --- |
| **Response Options** | • Full<br>• Partial<br>• Standard Formula |

79% of respondents intend to apply for full model approval (which will include the approval of the standard formula) as shown in the pie chart below.

**Type of Model Approval Sought**



Legend:
- Full
- Partial
- Standard Formula
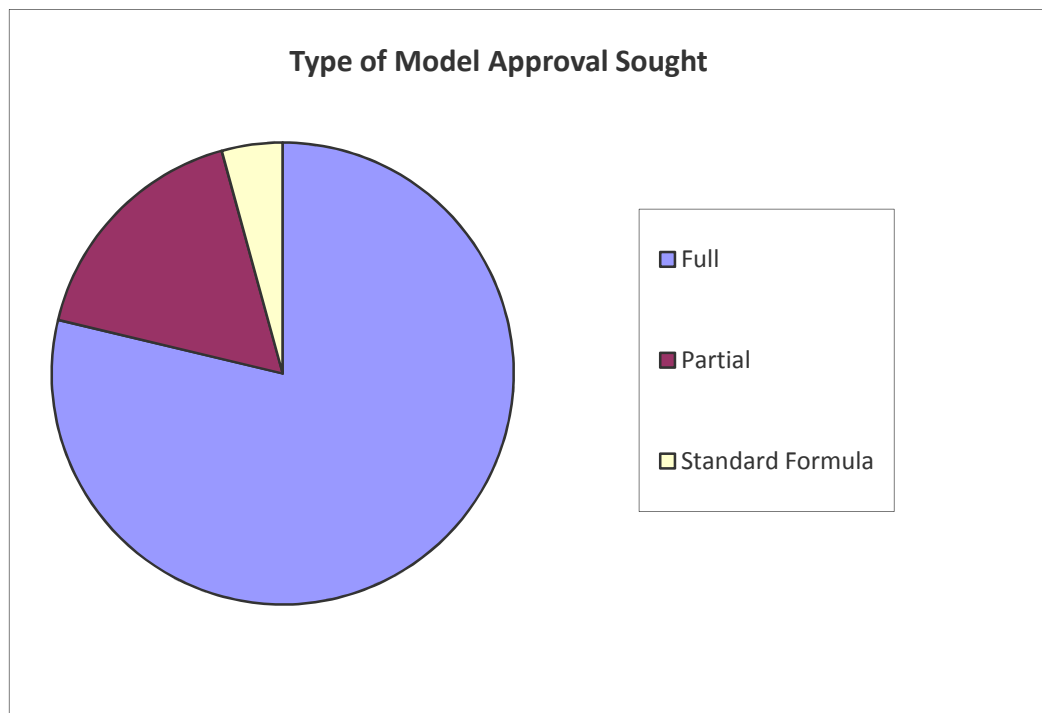
| **Question 6** | | **How onerous will your validation be compared to ICA?** |
|---|---|---|
| **Response Options** | Year One or Going Forward (e.g. Every Quarter) | • Same<br>• 1.5 times as extensive<br>• 2 times as extensive<br>• 3 times as extensive<br>• More than 4 times as extensive |

The onerous of Solvency II compared to ICA is summarised below in terms of year one as well as going forward.

| Question 7 | Who produces the validation report? |
|---|---|
| **Response Options** | • Risk Function |
| | • Actuarial Function |
| | • Other (please specify) |

The validation report will mostly be prepared by the Risk Function (48%), with slightly fewer (38%) using the Actuarial function, and a small proportion making use of External advisors (see pie chart below).

**Who Produces the Validation Report?**

- Risk Function
- Actuarial Function
- Other (please specify)

| Question 8 | Who is the report addressed to? |
|---|---|
| **Response Options** | • Board or delegated Board Sub-Committee |
| | • Senior Management |
| | • Chief Risk Officer |
| | • Chief Approved/Appointed Actuary |
| | • Internal Audit |
| | • Other (please specify) |

The validation report would be addressed to the board in 76% of the cases, with some companies addressing the report to other entities such as senior management and the Chief Risk Officer.

**Who is the Report Addressed To?**

■ Board or delegated Board Sub-Committee

■ Senior Management

□ Chief Risk Officer

□ Other (please specify)
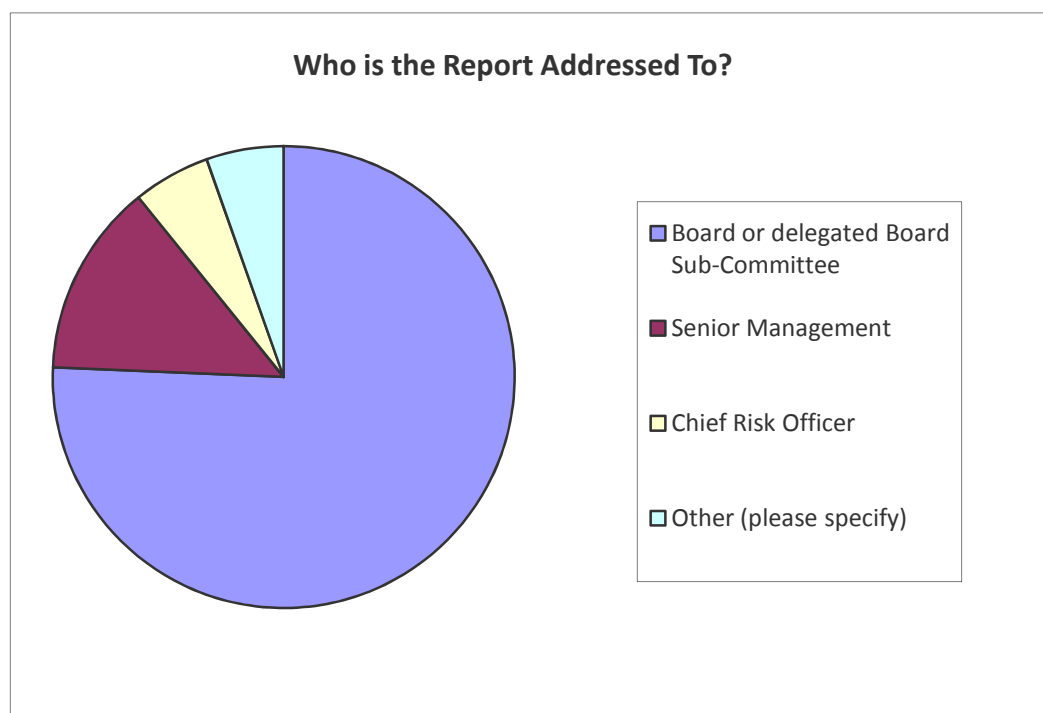
| Question 9 | Who will carry out the independent review? |
|---|---|
| **Response Options** | • Independent internal actuaries<br>• Risk managers<br>• Internal audit<br>• External audit<br>• External consultants<br>• Other (please specify) |

As per the chart below, 'External consultants' will be used in the majority for the Independent Model Review with 'Other' including the Parent company, underwriters and sub-committees of the board.

**The Independent Review**
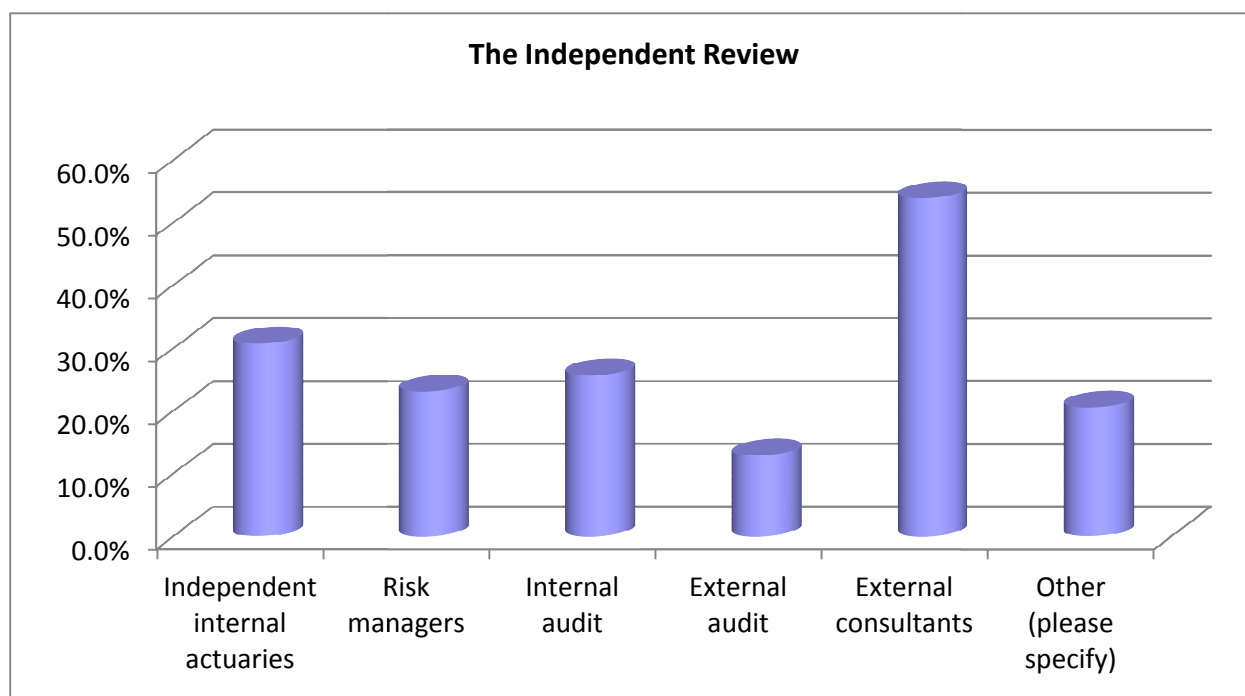
| Question 10 | What is the nature of the independent review? |
|---|---|
| **Response Options** | • High level/benchmarking |
| | • Peer review |
| | • Audit |
| | • Technical Review |
| | • Other (please specify) |

The nature of the independent review is broken down below, with 'Other' including all 4 options in the majority of the responses.



**The Nature of the Independent Review**

Legend:
- High level/benchmarking
- Peer review
- Audit
- Technical review
- Other
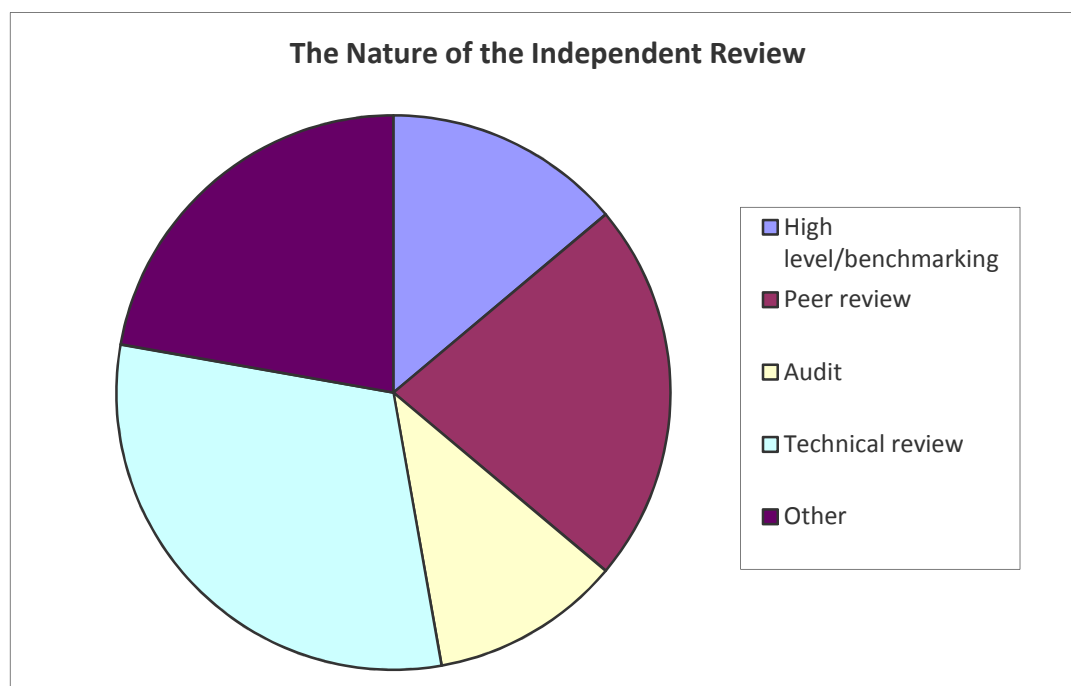
| Question 11 | Which of the following third-party catastrophe modelling software do you use, if any? |
|---|---|
| **Response Options** | <ul><li>RMS</li><li>AIR</li><li>EqeCat</li><li>Other (please specify)</li></ul> |

With respect to third party catastrophe modelling software, 83% of respondents will make use of RMS; 45% of AIR; 28% of EqeCat and 14% will use no software due to no, or limited, exposure to such catastrophes (see chart below).
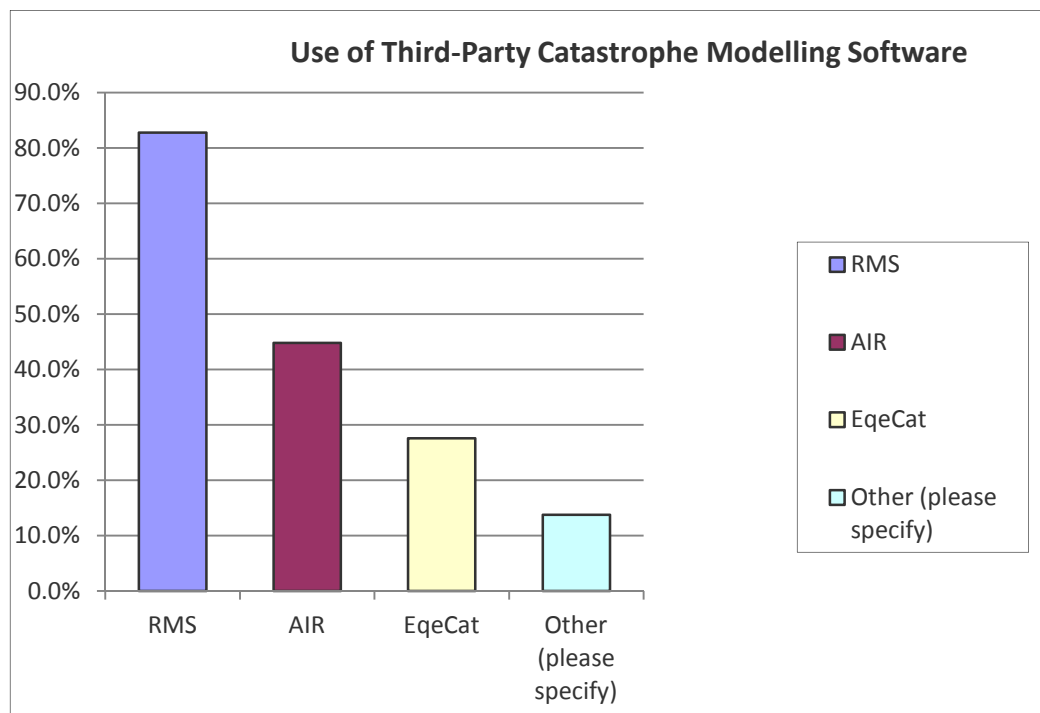
| Question 12 | Where do you see the key challenges in applying the Solvency II tests and standards to catastrophe models? Please prioritise on scale of 1 (low) to 8 (high). |
| --- | --- |
| **Response Options** | <ul><li>Use (1-8)</li><li>Validation (1-8)</li><li>Calibration (1-8)</li><li>Statistical quality (1-8)</li><li>Documentation (1-8)</li><li>Model Governance (1-8)</li><li>Profit and Loss Attribution (1-8)</li></ul> |

Below is a bar graph indicating the key challenges in applying the Solvency II test and standards to catastrophe models, rated by area of Model Validation with 1 being low and 8 being high.



**Key Challenges in Applying Solvency II Tests and Standards to Catastrophe Models**
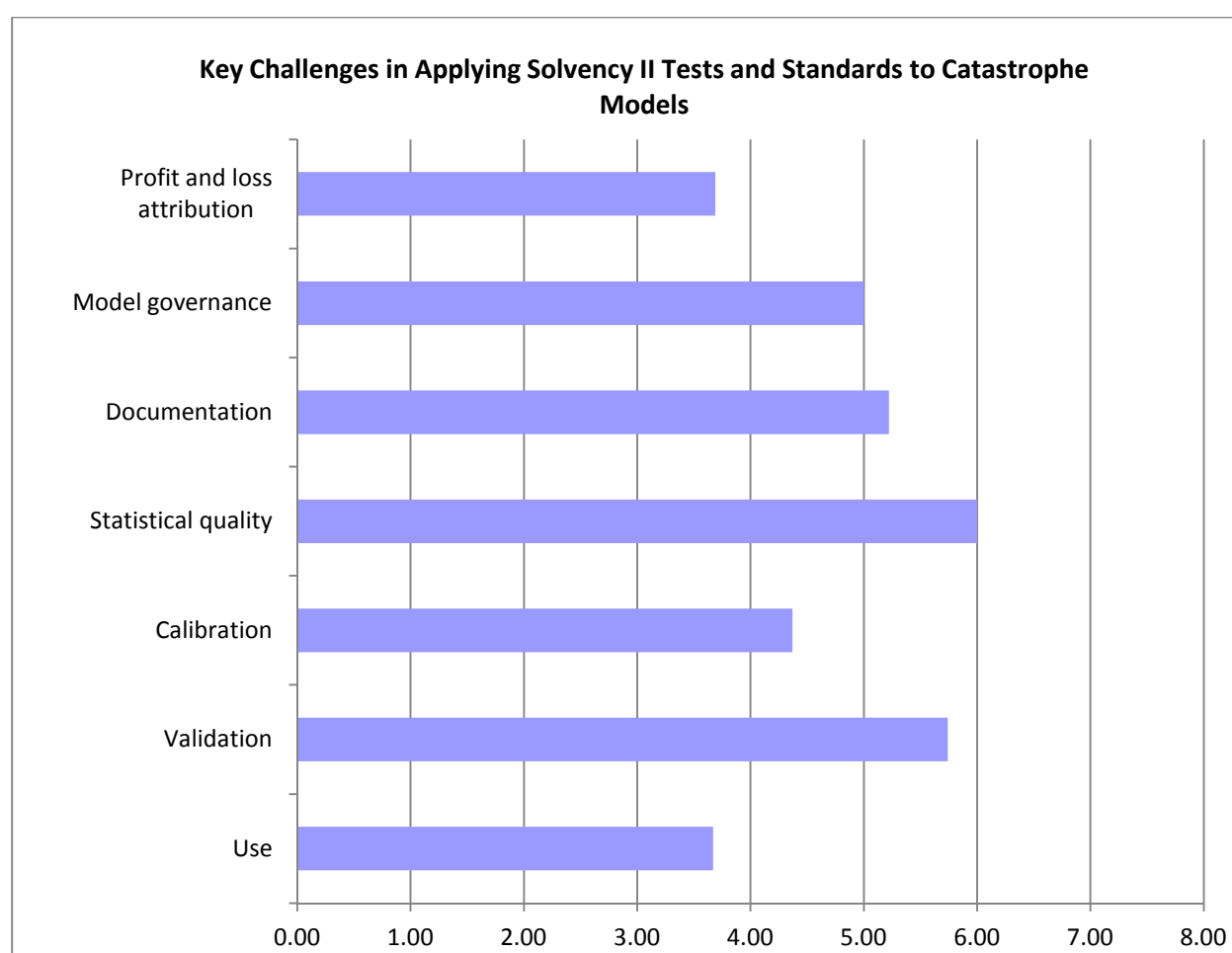
| Question 13 | What are the reasons for this? Please prioritise on a scale of 1 (low) to 8 (high). |
|---|---|
| **Response Options** | • Demonstrating a detailed understanding of the external model<br>• Quality of policy/exposure data<br>• Validation of outputs of the model<br>• Aggregation of the results other modelled losses<br>• Consistency of assumptions in capital model and exposure management/pricing<br>• Other - please describe in the box below |

The reasons for these key challenges above are rated below by category, with 'Other' including reasons such as Multi Model implementation challenges.
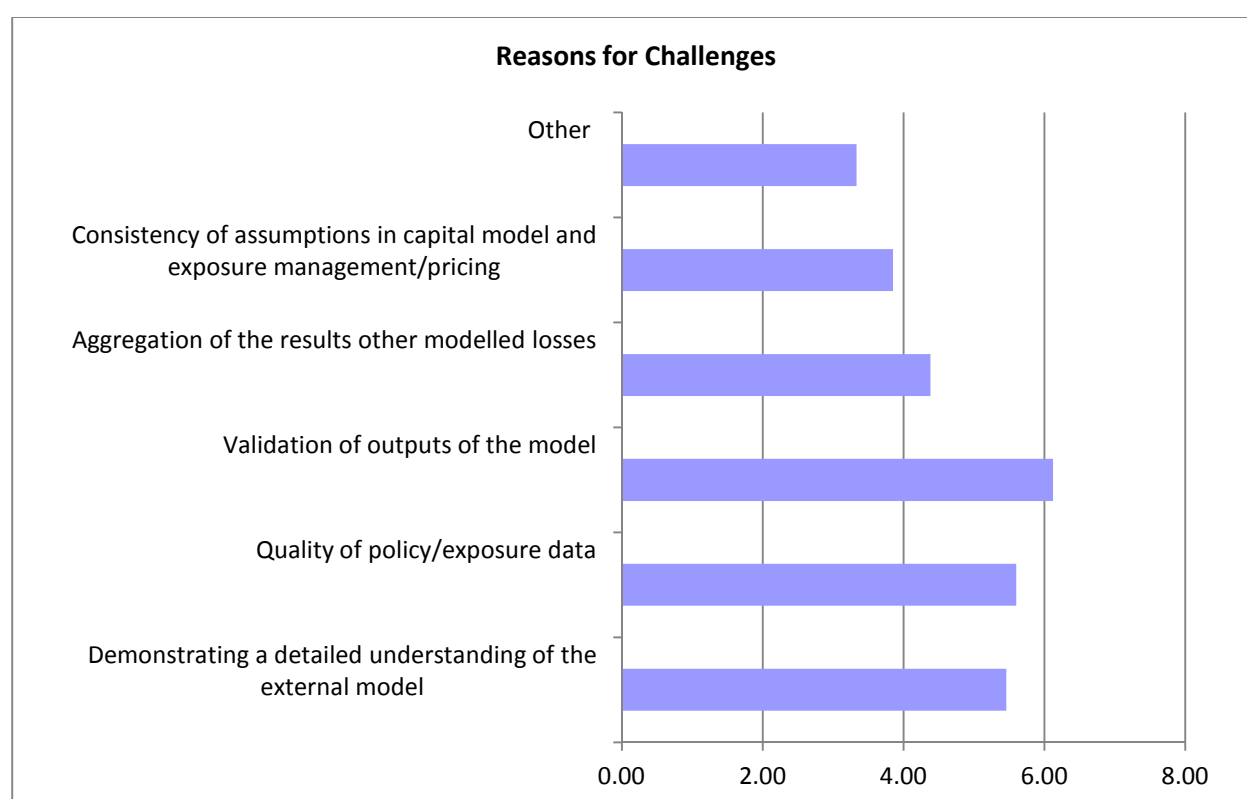
| **Question 14** | **Who is responsible for Solvency II approval of the use of the catastrophe model in your company?** |
|---|---|
| **Response Options** | • Capital Modelling Team<br>• Underwriters<br>• Catastrophe Modellers<br>• Other (please specify) |

The responsibility of Solvency II approval of the use of the catastrophe model relies heavily on the Capital Modelling team, with slightly less responsibility in respect of Catastrophe modellers and the Risk Management of companies (see pie chart below). 'Other' responses included, risk management, risk committee, internal model and governance committee and the reinsurance team.



Responsibility for Catastrophe Model Approval
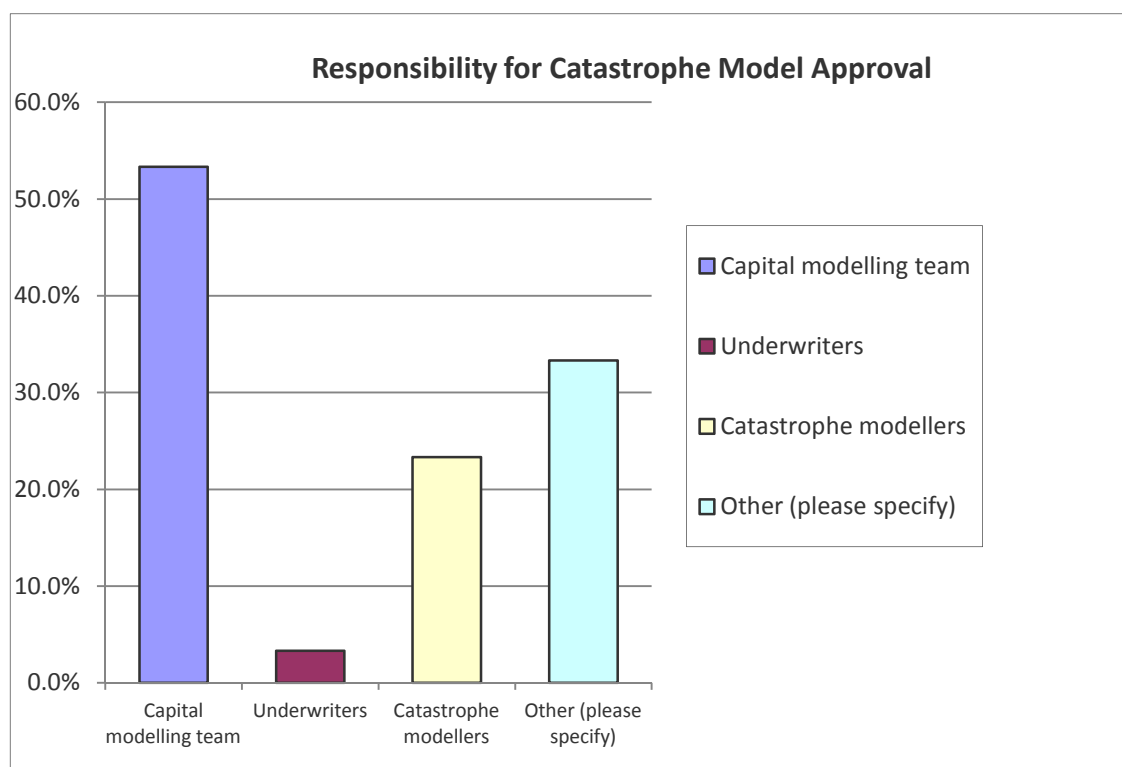
| Question 15 | What measurement basis do you plan to use for value the SCR? |
|---|---|
| **Response Options** | • 99.5% VaR over 1 year<br>• Other 'at least equivalent' measure - please briefly describe basis in the box below |

92% of respondents indicated that they will be using the 99.5% VaR over 1 year measurement basis to value the SCR (see pie chart below). 'Other' responses included, average over 99.25 to 99.75 percentile and 99% TVAR.

**Measurement Basis for Valuing the SCR**

■ 99.5% VaR over 1 year

■ Other 'at least equivalent' measure
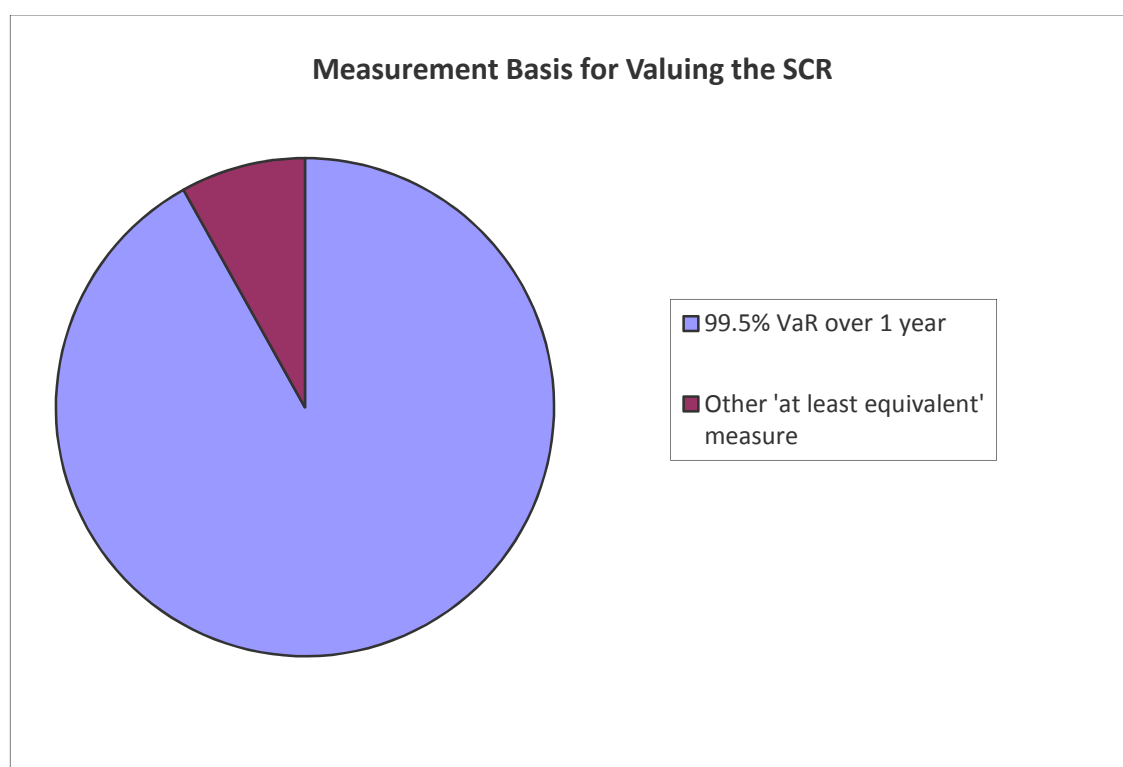
| **Question 16** | **How do you intend to define economic capital?** |
| --- | --- |
| Response Options | • Respondents own definition |

There were a variety of responses for this question. Most respondents' definitions included some reference to a multiple of SCR.

| Question 17 | What method(s) best describe your approach to the 1-year time horizon? |
|---|---|
| Response Options | • Actuary in the Box/Re-reserving (definition – 1 year of claim development is simulated, and an algorithm is applied to the results of this simulation to provide a distribution of reserve estimates) |
| | • Proportionate emergence (definition – a fixed proportion of the ultimate reserve deterioration is assumed to occur in the first year) |
| | • Merz-Wüthrich (definition – a standard formula is used to generate a 1-year standard deviation) |
| | • Other (please specify) |

In terms of Calibration, 44% of respondents will be using the Actuary in a Box / Re-reserving approach, 28% Proportionate Emergence, 3% Merz-Wüthrich and 25% taking alternative approaches, such as reserve to ultimate and non-proportionate emergence to come up with the 1 year time horizon.
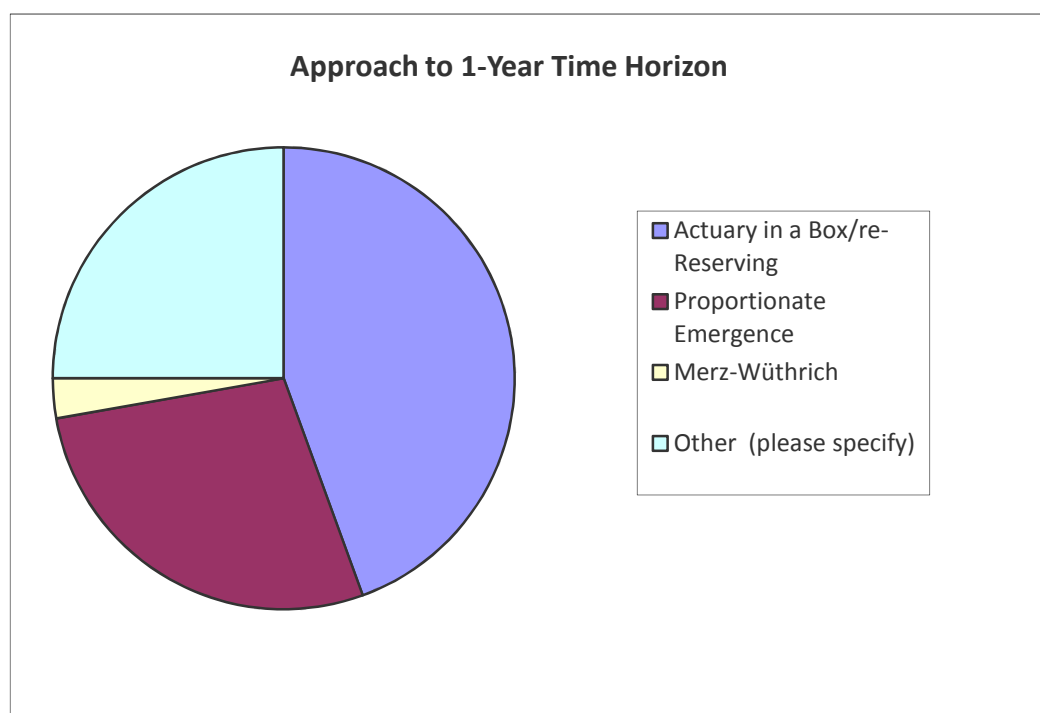
| Question 18 | Will the above methods vary by class? |
|---|---|
| **Response Options** | • Actuary in the Box/Re-reserving<br>• Proportionate emergence<br>• Merz-Wüthrich<br>• Other (please specify) |

In some companies these methods will vary by class. Below is a summary of the 3 main methods, and shows the proportion of respondents whom will be able to vary these methods by class.
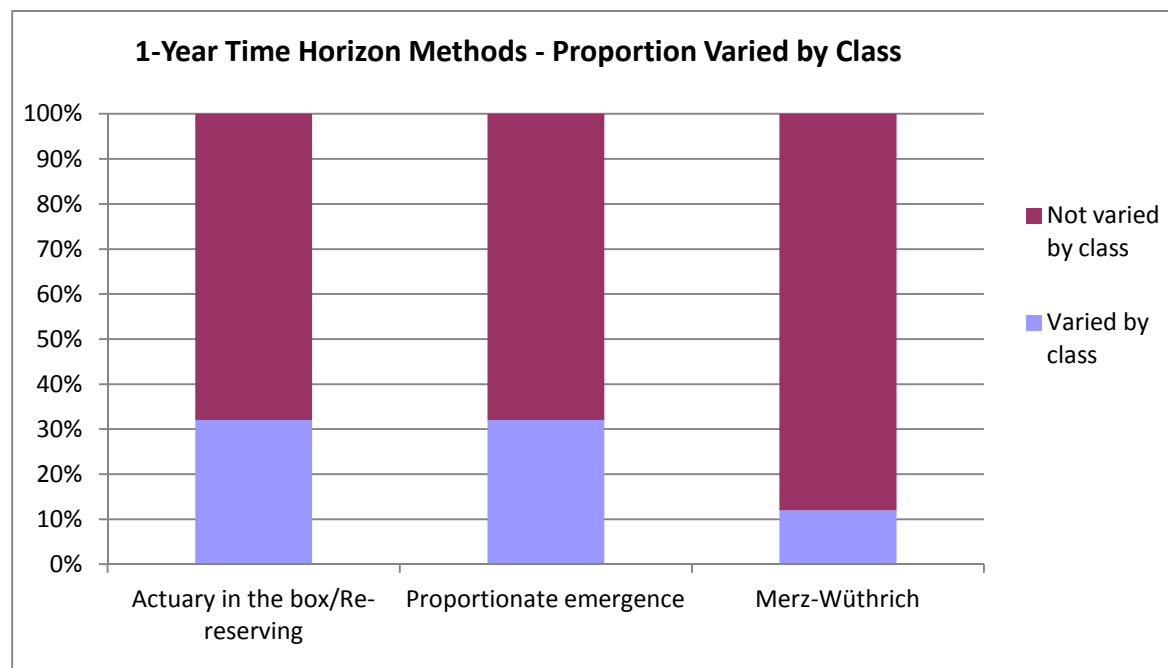
| Question 19 | How are you planning to estimate the risk margin at the end of 1 year? |
|---|---|
| **Response Options** | • Fixed proportion of best estimate <br> • Stochastic <br> • Other |

Regarding the estimation of the risk margin at the end of 1 year, 53% of respondents will be using a fixed proportionate basis, 36% a stochastic approach and 11% are unsure (see pie chart below).

**Estimating the Risk Margin**

■ Fixed proportion of best estimate

■ Stochastic
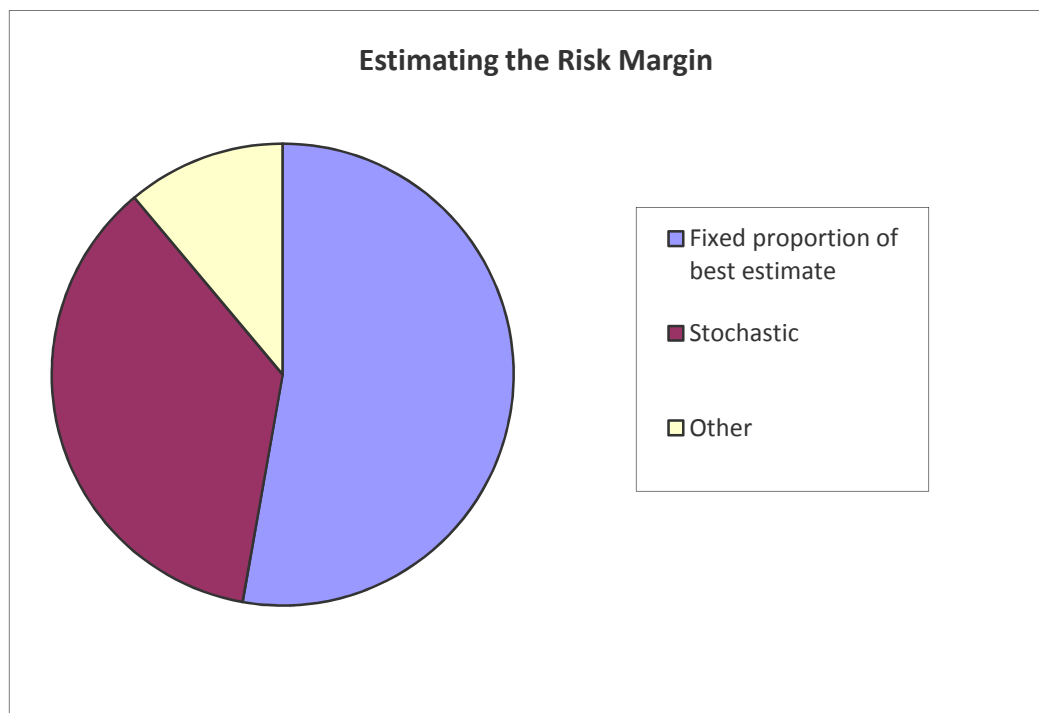
□ Other

| Question 20 | Which platform or system do you intend to use for your Solvency II documentation? |
|---|---|
| **Response Options** | • Sharepoint<br>• Quality workstream<br>• Conductor<br>• Nimbus<br>• Other – please describe in the box below |

As the pie chart below shows, in terms of the platform to be used for Solvency II Documentation, 60% of respondents will be using Sharepoint, 6.3% a Quality Workstream and 'Other' responses included K Center, Documentum, ARIS and BPM One..
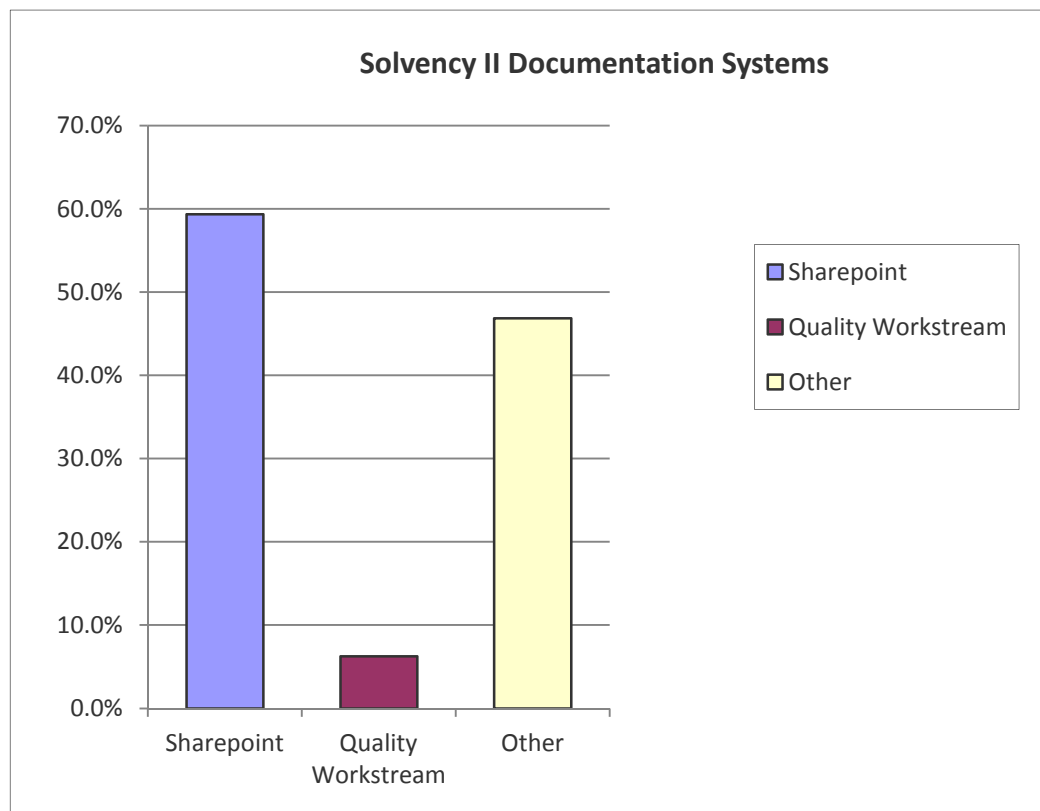
| Question 21 | Please give an indication of the level of enhancements required to EXISTING documentation (e.g. ICA) that you are planning to carry out to meet the Solvency II documentation standard for each of the following categories of documentation. Select from no change/minor change/substantial rewrite/not producing such documentation. |
|---|---|
| Response Options | • Methodology documentation<br>• Underlying mathematical theory<br>• Detailed model code documentation<br>• Internal controls documentation<br>• Executive summaries for senior management and users<br>• Outputs/management information<br>• Validation documentation<br>• Data quality documentation<br>• Other data documentation<br>• Process/mapping flowcharts<br>• Process documentation<br>• Evidence of use<br>• Audit trails |

The chart below gives an indication of the level of enhancements required to EXISTING documentation (e.g. ICA) that companies are planning to carry out to meet the Solvency II documentation standard for each of the following categories of documentation. In most cases 'Other' included the fact that the level of enhancements has not been determined yet.
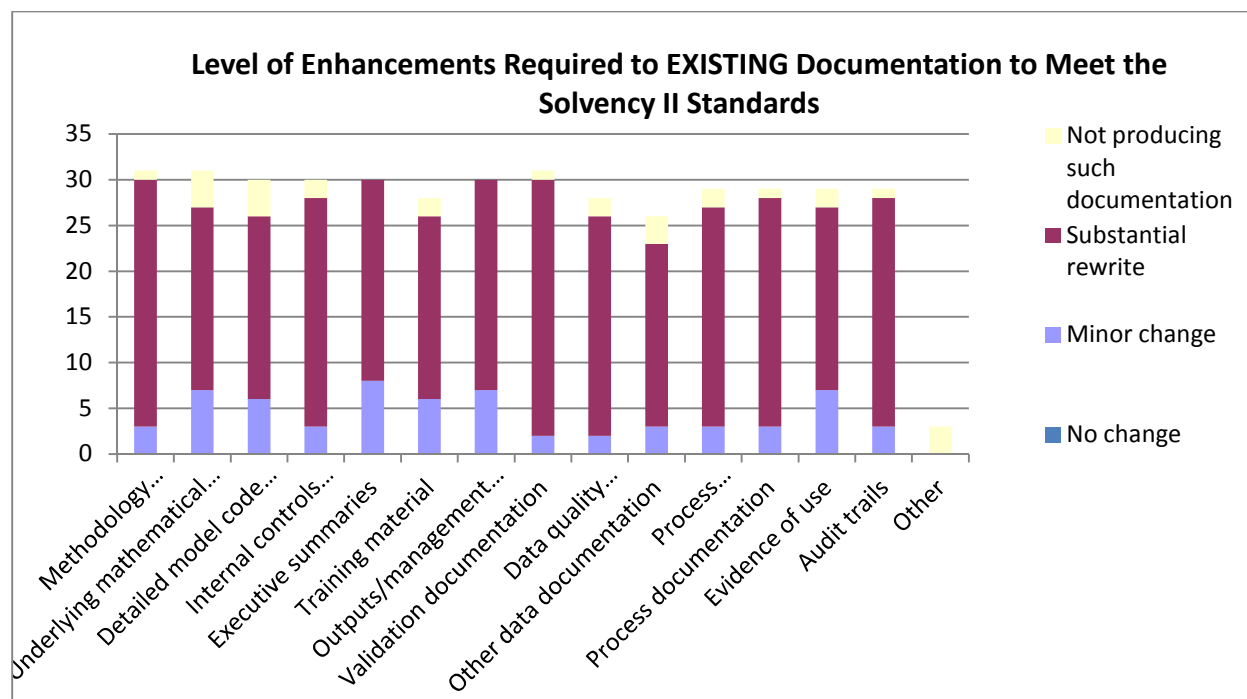
| **Question 22** | **Please give an indication of the level of NEW documentation that you are planning to produce out to meet the Solvency II documentation standard for each of the following categories of documentation. Select from none (i.e. using existing)/some/lots/not producing such documentation.** |
|---|---|
| **Response Options** | • Methodology documentation<br>• Underlying mathematical theory<br>• Detailed model code documentation<br>• Internal controls documentation<br>• Executive summaries for senior management and users<br>• Outputs/management information<br>• Validation documentation<br>• Data quality documentation<br>• Other data documentation<br>• Process/mapping flowcharts<br>• Process documentation<br>• Evidence of use<br>• Audit trails |

The chart below gives an indication of the level of NEW documentation (e.g. ICA) that companies are planning to prepare to meet the Solvency II documentation standard for each of the following categories of documentation. In the case of 'Other', it mostly included the fact that the level of new documentation has not been determined yet.
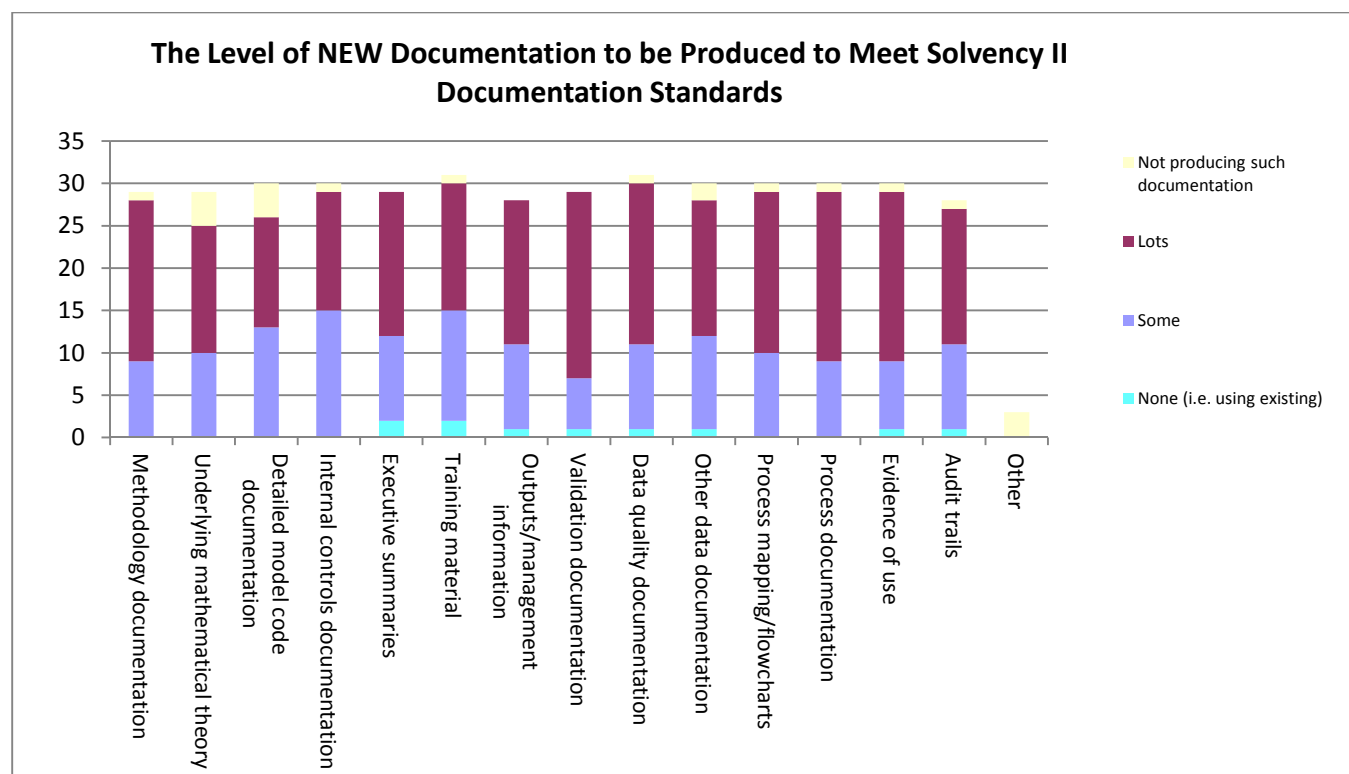
| Question 23 | How much resource (full time equivalent months) does this require? For 'other', please describe the function and full-time equivalent months in brackets, e.g. Other (5). |
|---|---|
| Response Options | • Actuarial<br>• Technical writer<br>• Risk management<br>• Other (please describe, and give time required) |

In terms of resourcing, time wise IMAP will require 97% of Actuarial, 63% of Technical Writers, 90% of Risk Management and 30% of other areas, which includes IT, Finance and Senior Management.

**Level of Resourcing Required**
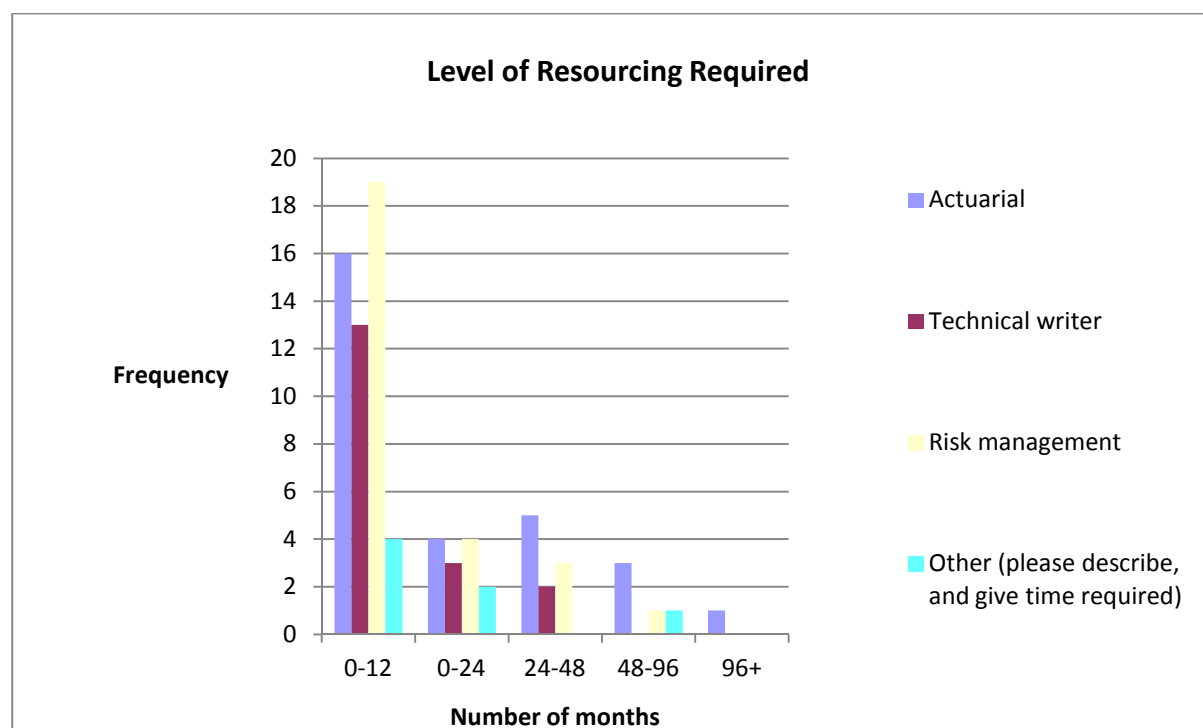
| **Question 24** | **Do you plan to produce a mapping from the Solvency II Directive, Lloyd's and/or CEIOPS guidance to your own documentation to provide evidence of compliance?** |
|---|---|
| **Response Options** | • Yes<br>• No |

69% of respondents prefer to use a mapping from the Solvency II Directive, Lloyd's and/or CEIOPS guidance to their own documentation to provide evidence of compliance (see pie chart below):

**Use of Mapping**



Legend: Yes, No
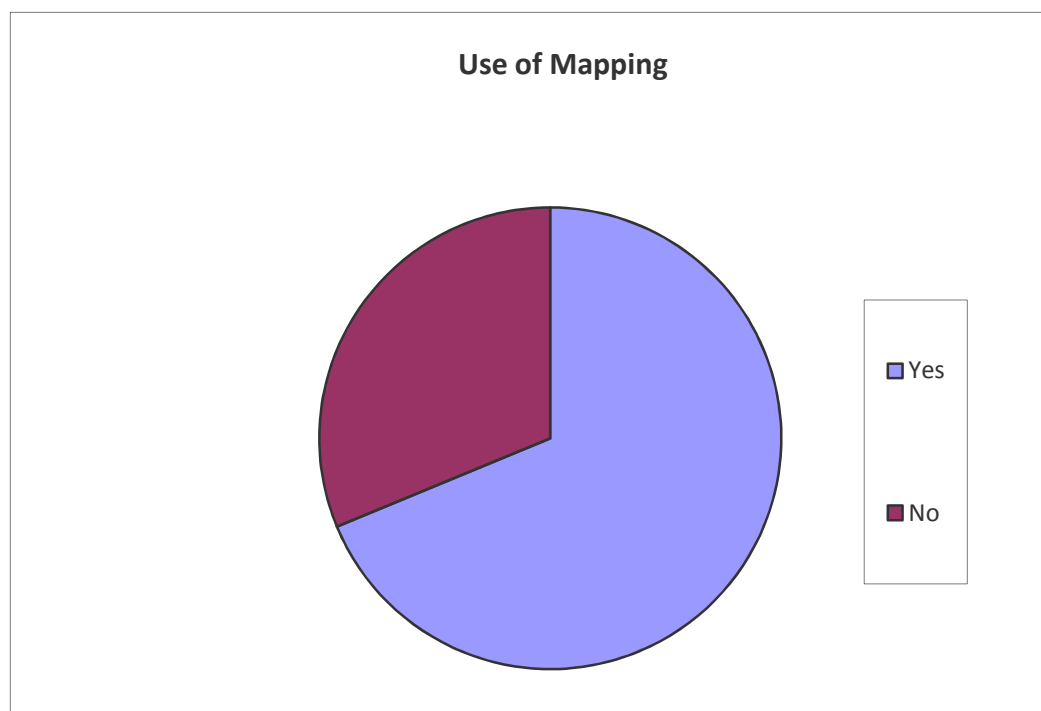
| Question 25 | **Do you have good examples of Solvency II documents you would be prepared to share with the working party? If Yes, please e-mail document that you are willing to share to peter.stirling@actuaries.org.uk, Secretary, General Insurance Practice Executive Committee** |
|---|---|
| **Response Options** | • Yes<br>• No |

N/A

| Question 26 | How do you intend to use the model? Select from core use/non-core use |
|---|---|

| **Response Options** | • Adequate pricing |
|---|---|
| | • Assessing customer benefits, for example, bonus setting |
| | • Asset/liability management |
| | • Business planning/strategy |
| | • Capital management |
| | • Development and monitoring of risk appetite |
| | • Development of risk strategies |
| | • Efficient use of capital/Portfolio optimisation |
| | • Efficient use of capital/Performance measurement |
| | • Exposure management and limit setting |
| | • External risk reporting |
| | • Financial reporting - internal model provides market valuations for IFRS |
| | • Incentive/target setting |
| | • Internal risk monitoring (through MI) |
| | • Investment decisions e.g. strategic, tactical and operational decisions |
| | • Mergers and acquisitions |
| | • Measurement of material risks |
| | • ORSA |
| | • Other risk mitigation |
| | • Portfolio transfer pricing |
| | • Reinsurance |

The uses of the model are set out below.
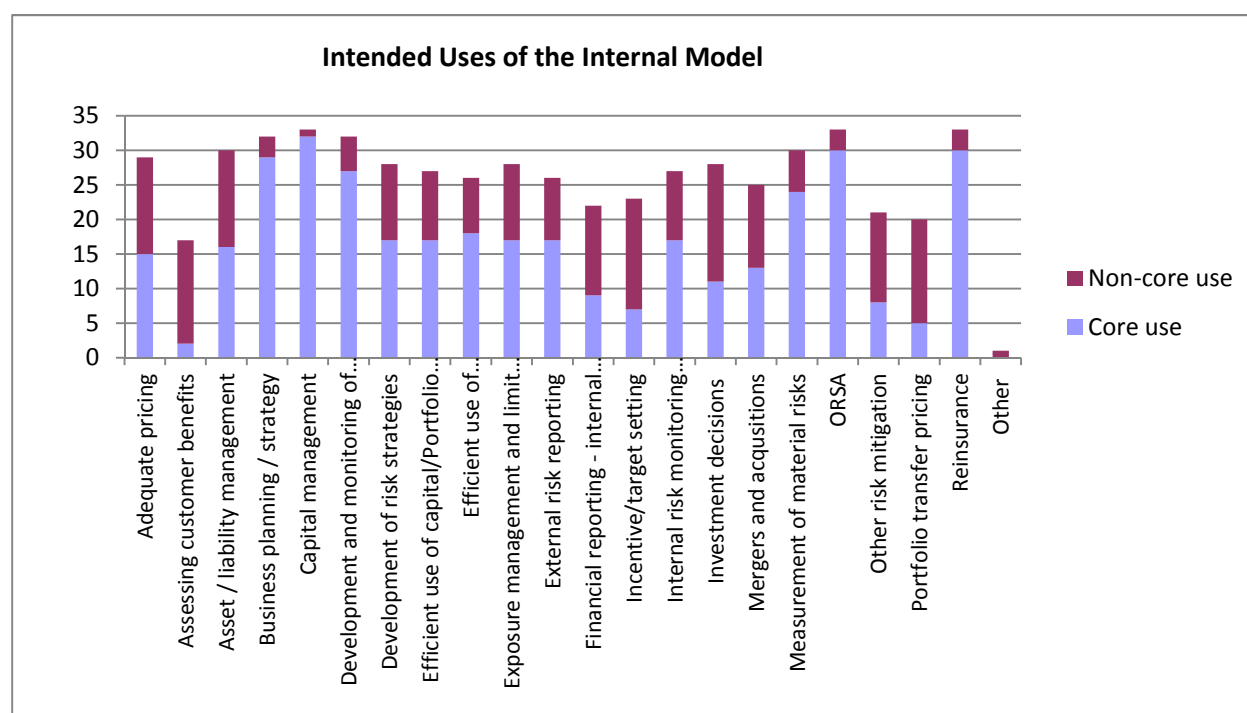
| **Question 27** | **At what level are you looking to apply the use test?** |
|---|---|
| **Response Options** | • Group level |
| | • Legal entity level |
| | • Business unit level |
| | • Other (please specify) |

The level at which companies plan on using the Used Test is outlined below. 'Other' includes all of the levels and in some cases it has not been decided yet.

**The Level at Which the Use test Will be Applied**



- Group level
- Legal entity level
- Business unit level
- Other
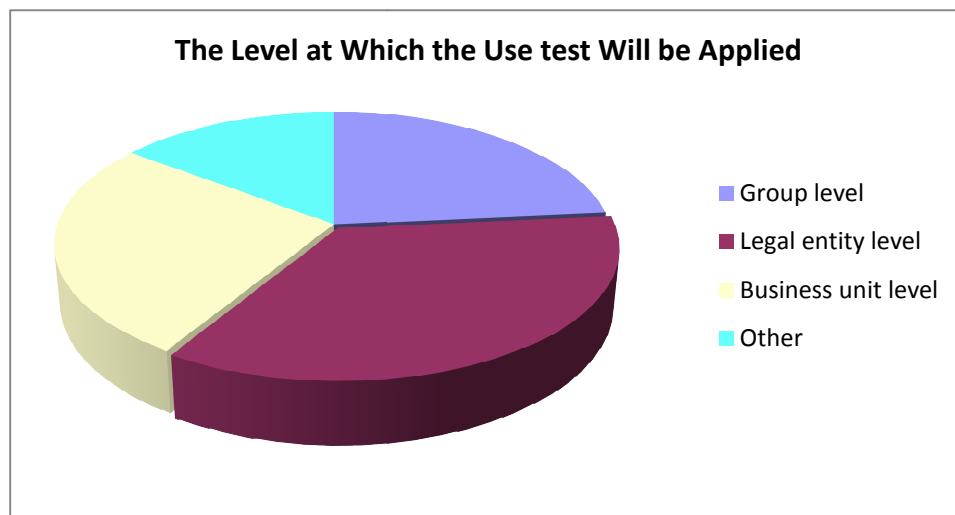
| **Question 28** | **What training programme are you providing to the Board and senior management to enable them to effectively challenge the model?** |
|---|---|
| **Response Options** | • Statistics<br>• High-level summary of the model<br>• Detailed training regarding the model<br>• Other (please specify) |

Outlined below are the training programmes that companies are providing to the Board and senior management to enable them to effectively challenge the model. 'Other' includes responses such as none and still to be decided.

**Training Programmes to be Provided to the Board and Senior Management to Enable them to Effectively Challenge the Model**
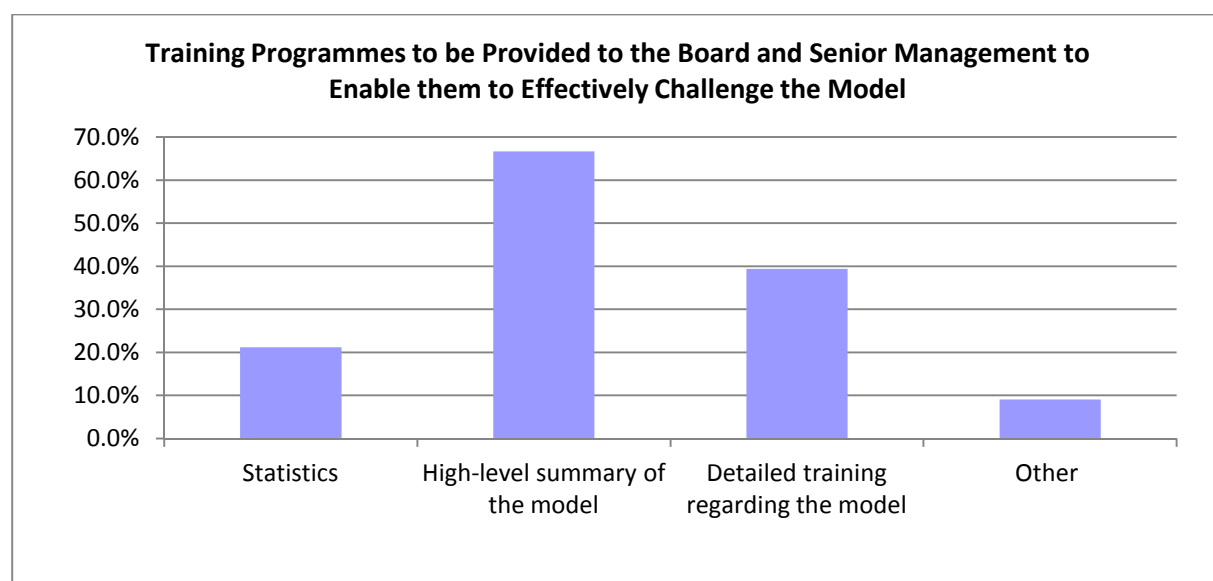
# B   Appendix B :  Overview and background of Solvency II

## B.1   Solvency II

B.1.1     Solvency II is the new regulatory regime for the insurance industry due to be implemented EEA (European Economic Area) wide on 1 Jan 2014[37].  At the time of writing this paper, we expect Solvency II to become "fully operational" by 1 Jan 2015, followed by a period of transitional measures over which specific areas of the regulation will be implemented (e.g. public disclosure implemented over a period of up to 3 years).

B.1.2     Solvency II encompasses the supervision of all insurance undertakings domiciled in the EEA. It includes provisions relating to the authorisation and winding up of insurance companies, as well as requirements for enhanced risk and capital management.  Its key objectives are to promote an efficient marketplace in insurance services in the EEA, and the protection of policyholders.

B.1.3     The current solvency requirements for insurers were in large developed in the 1970s and 1980s, with some minor changes in the 1990s.  In 2002 they underwent a limited revision, and became known as Solvency I.

B.1.4     The new regime is being created through a standard legislative process, termed the "Lamfalussy Process", which is described below.

B.1.5     Solvency II represents a fundamental change in the risk management and reporting standards for the industry.

## B.2    Lamfalussy process - the four "levels"

B.2.1     The Lamfalussy Process is an approach to the development of financial service industry regulations used by the European Union.  It was developed in 2001, and named after the chair of the EEA advisory committee that created it, Alexandre Lamfalussy.

B.2.2     The process is composed of four "levels", each focusing on a stage of the implementation of legislation.

### *Level 1 - Framework Principles*

B.2.3     At the first level of the Lamfalussy process, the European Parliament and Council of the European Union adopt a Directive, establishing the core principles of a law and outlining guidelines on its implementation.

B.2.4     The Solvency II Framework Directive (Directive 2009/138/EC) forms the Level 1 part of Solvency II, and is formally entitled the "Directive on the taking up and pursuit of the business of insurance and reinsurance".

---

[37] 3.1.2   The Solvency II Directive states that the new regime will go live on 1 November 2012, however this has now effectively been extended to 1/1/2014 by the Ominbus II Directive and further delays remain a possibility.

B.2.5    The Omnibus II Directive proposes amendments to certain areas of the Solvency II Framework Directive (described below in Section 3.9). Amendments were voted on in March 2012 and the final Directive is expected to be adopted later in 2012, after the plenary vote in September.

B.2.6    In this paper, we have referred to this Solvency II Directive as either "The Framework Directive", "The Directive", or the "Level 1/2/3 text".

## *Level 2 - Implementing Measures*

B.2.7    Once the Level 1 Directive was adopted, the process moved on to the second level.

B.2.8    Solvency II's implementing measures (Level 2) provide the detailed requirements of what will be required of insurers.

B.2.9    It is currently being developed by the European Commission, based on advice from EIOPA (formerly CEIOPS).  In this report, we refer to the draft implementing measures released by the European Commission for private consultation in October 2011.

B.2.10   For internal model purposes, the key advice was contained in CEIOPS' paper to the EC entitled "CEIOPS' Advice for Level 2 Implementing Measures on Solvency II Tests & Standards for Internal Model Approval (former Consultation Paper 56)".

B.2.11   We refer in this paper to the implementing measures being developed by the EC as the "Level 2 Text", and the CEIOPS advice as "Level 2 Advice".

## *Level 3 - Guidance*

B.2.12   At the third level, national regulators work on coordinating new regulations with other nations.

B.2.13   In this Level, EIOPA will adopt 'comply or explain' guidelines and recommendations, carry out peer review, mediate and settle disagreements, take action in emergency situations, facilitate delegation of tasks and responsibilities, monitor and assess market developments, undertake economic analyses and foster investor protection.  A new concept has recently been introduced into Level 3 called "Binding Technical Standards" under which EOIPA can set standards with which firms must comply, extending the concept of Level 3 beyond the original remit of guidance.

B.2.14   Final guidelines were released in July 2012. We expect EIOPA to begin public consultation of the entire Level 3 package shortly.

## *Level 4 - Enforcement*

B.2.15   The fourth level in the Solvency II Lamfalussy process involves the EC checking that each Member State has complied with the Solvency II Directive in its national implementing legislation.  The EC has the power to take enforcement action against any Member States which have not fully complied.

B.2.16   The Level 4 part of the process will not commence for a number of years.

## B.3   The Three Pillars

B.3.1   Solvency II encompasses three distinct areas, known as Pillars 1, 2 and 3.  Pillars 2 and 3 are often referred together as Pillar 5 due to the synergies between them.

- Pillar 1 consists of the quantitative requirements.  This pillar aims to ensure firms are adequately capitalised with risk-based capital. This pillar includes the requirements for the use of internal models.

- Pillar 2 sets out requirements for the governance and risk management of insurers, as well as for the effective supervision of insurers.

- Pillar 3 focuses on supervisory reporting, public disclosure and transparency requirements.

B.3.2   The focus in this paper is on the Pillar 1 quantitative requirements, and in particular, the requirements for the approval of an internal model.

## B.4   The Solvency Capital Requirement

B.4.1   Under the Solvency II regime, as with the current regime, insurers will have to establish technical provisions to cover expected future claims from policyholders.

B.4.2   In addition, insurers must have available resources sufficient to cover both a Minimum Capital Requirement (MCR) and a Solvency Capital Requirement (SCR).  Note the MCR is subject to a minimum floor (in general EUR2.2m for non-life companies; or EUR3.2m for non-life reinsurers, or insurers writing liability, credit or suretyship.).

B.4.3   The MCR is calculated by using a prescribed formula, and it is intended to be calibrated to ensure an 85% probability for the insurer to meet its financial obligations over a one year time horizon.

B.4.4   The SCR is the amount of capital needed to ensure a 99.5% probability for the insurer to meet its financial obligations over a one year time horizon using a Value-at-Risk (VaR[38]) measure.  It is the target level of capital which firms are required to hold.  The SCR can be calculated using an internal model, a standard formula, or a partial internal model.  This paper does not cover the standard formula or partial internal models.

B.4.5   Assets and liabilities underlying the SCR calculation should be valued at the amount for which they could be exchanged (or in the case of liabilities, settled) between knowledgeable willing parties in an arm's length transaction.

B.4.6   If an insurer's available resources fall below the SCR, then supervisors are required to take action with the aim of restoring the insurer's finances back into the level of the SCR as soon as possible.  If, however, the financial situation of the insurer continues to deteriorate, then the level of supervisory intervention will be progressively intensified.  The aim of this 'supervisory ladder' of intervention is to capture any ailing insurers before they pose a serious threat to policyholders' interests.

---

[38] http://en.wikipedia.org/wiki/Value_at_risk

B.4.7 If, despite supervisory intervention, the available resources of the insurer fall below the MCR, then 'ultimate supervisory action' may be triggered. In other words, the insurer's liabilities will be transferred to another insurer and the license of the insurer will be withdrawn or the insurer will be closed to new business and its in-force business will be liquidated.

B.4.8 This paper is concerned with the calculation of the SCR using an internal model, and does not cover the standard formula.

## B.5 Definition of Internal Model

B.5.1 There is no definition of an "internal model" within the Solvency II Directive. Instead, the onus is on firms to define their internal model and the scope of their IMAP submission.

B.5.2 The International Association of Insurance Supervisors (IAIS) Standard includes the following definition of an internal model, based on the definition in the Group Consultatif/CEA Glossary:

B.5.3 "[An Internal model is] a risk measurement system developed by an insurer to analyse its overall risk position, to qualify risks and to determine the economic capital required to meet those risks. Internal models may also include partial models which capture a subset of the risks borne by the insurer using an internally developed measurement system which is used in determining the insurer's economic capital (IAIS, 2008)".

## B.6 The Internal Model Approval Process

B.6.1 The Level 2 draft requirements (Article 203 IM1, refer also CEIOPS' CP37) set out how the internal model approval process is expected to work in practice. CEIOPS also proposes using a pre-application process to bring forward supervisory review, but not pre-approval of models.

B.6.2 The FSA required firms seeking a 'day one' decision on their internal model to submit a Pre-Application Qualifying Criteria (PAQC) document by 28 February 2011. They accepted 78 firms into pre-application at the end of March 2011 and closed the process to new entrants.

B.6.3 In April 2011[39], the FSA announced that they have decided to concentrate on a smaller population of firms representing a significant market share or having complex risk structures and which they regard as having the highest potential impact on their objectives, with all other firms receiving a reduced level of engagement.

B.6.4 The FSA's original intention was that firms would have a window to submit a formal application for model approval between March and May 2012. This window was extended to mid-2013 , reflecting likely 'go live' date of 1 Jan 2014. This includes a full self-assessment against the Level 1 and Level 2 draft requirements.

---

[39] http://www.fsa.gov.uk/pages/Library/Communication/Speeches/2011/0319_ja.shtml

B.6.5    Given that Solvency II has yet to be transposed into national law, the FSA does not currently have the powers to approve internal models. Therefore the current applications are strictly "submissions" which will need to be converted to formal applications for regulatory consideration once the FSA has the relevant powers.

B.6.6    In light of delays resulting from implementation of Omnibus II, it is likely that FSA would review its implementation schedule. However, currently FSA will continue to work with an implementation date on 1 Jan 2014.

## B.7    The Six Tests of the Internal Model

B.7.1    The Level 1 framework sets out six tests or standards that the internal model must meet for the supervisor to give approval.  These are:

- The Use Test - firms need to demonstrate that the internal model is widely used in, and plays an important role in, their system of governance;

- Statistical quality standards - Appropriate statistical quality standards should be applied across all aspects of the internal model;

- Calibration standards - The internal model should be calibrated to calculate the SCR at a 99.5% confidence level over one year, or an equivalent measure;

- Profit and loss attribution - A profit and loss attribution exercise must be performed each year;

- Validation standards - All aspects of the internal model should be validated, including a comparison with emerging experience;

- Documentation standards - The internal model should be documented to a certain standard.

B.7.2    In addition to these tests, the framework provides guidance on External Models and Data.

B.7.3    This paper is organised into sections based on these 6 tests (plus a section on External Models and Data), and further information on each test is given in the relevant section of this paper.

## B.8    Omnibus II

B.8.1    On 19 Jan 2011, the European Commission published its proposal for Omnibus II (OII).  OII will amend the framework directive, to bring it into line with the Lisbon Treaty and to take account of the EU's new supervisory structure. It also pushes Solvency II's implemention date back and gives the EC powers to introduce many parts of Solvency II over a 10 year period ('transitional measures').

B.8.2    Omnibus II is an EEA Directive, with the same legal basis as the two Directives it amends – the Prospectus Directive and the Solvency II Directive.  Putting it into force requires agreement from the European Parliament and Council, and the process has been subject to many delays.  Postponement of Omnibus II is significant because until it is in place, the European Commission cannot formally adopt Level 2 measures and Level 3 guidance cannot be produced.

## B.9   Solvency I

10.8.4    The current solvency requirements for insurers were in large developed in the 1970s and 1980s, with some minor changes in the 1990s.  In 2002 they underwent a further limited revision, and became known as Solvency I.

B.9.1    For non-life insurers, the Solvency I Directive is set out in Directive 2002/13/EC.  Member States were required to adopt their laws to comply with the 2002 Directive by 20 September 2003 and its measures were first applied to the supervision of accounts for financial years beginning on 1 January 2004.

B.9.2    Solvency I did not fundamentally change the requirements, and in the process of making the changes it became clear that a more wide-ranging reform was required – hence Solvency II.

B.9.3    In outline, each Member State must require each non-life insurer whose head office is situated in its territory to have an adequate solvency margin, equal to the requirements of the Directive. Those requirements are:

- The insurer's solvency margin is its assets free of foreseeable liabilities, less intangible items.

- The solvency margin must be determined on the basis of either the annual amount of premiums or the average burden of claims for the past three financial years.

- The solvency margin is set as to the higher of two results: the premium basis and the claims basis. Calculations of these are detailed but straightforward. Provision is made for some classes, viewed as volatile: for example, for classes 11, 12 and 13 (aircraft liability, ships liability and general liability), the premiums and claims used must be increased by 50%.

- One-third of the required solvency margin constitutes the guarantee fund. This may not be less than EUR 2m, and if risks in classes 10 - 15 are covered, it must be EUR 3m.

B.9.4    These requirements apply to UK insurers through provisions set out in the FSA's GENPRU handbook.  UK insurers are also subject to the FSA's Individual Capital Adequacy Standards (ICAS) Framework, which is close to Solvency II in its approach.

# C   Appendix C: Glossary of Terms

| Acronym | Term |
|---------|------|
| AAL | Average Annual Loss |
| BAU | Business As Usual |
| CaDeT | Capital Deployment Tool |
| CAT | Catastrophe |
| CEA | Council of Economic Advisors |
| CEIOPS | The Committee of European Insurance and Occupational Pensions Supervisors |
| CRO | Chief Risk Officer |
| EEA | European Economic Area |
| EIOPA | The European Insurance and Occupational Pensions Authority |
| EP | Excedence Probability curve |
| EPD | Expected Shortfall |
| ESG | Economic Scenario Generator |
| FCAS | Financial Claims Advisory Service |
| FIA/FFA | Fellow of the Institute of Actuaries/ Fellow of the Faculty of Actuaries |
| FSA | Financial Services Authority |
| GAAP | Generally Accepted Accounting Principles |
| GIRO | General Insurance Research Organising committee |
| HERP | Headquarters for Earthquake Research Promotion |
| IAIS | International Association of Insurance Supervisors |
| ICA | Individual Capital Assessment |
| ICAS | Individual Capital Adequacy Standards |
| IFRS | International Financial Reporting Standards |
| IMAP | Internal Model Approval Process |
| LGI | Legal and General Insurance |
| LOB | Line of Business |
| MCR | Minimum Capital Requirement |
| ORSA | Own Risk and Solvency Assessment |
| P&L | Profit & Loss |
| PAQC | Pre-Application Qualifying Criteria |
| PCS | Property Claim Services |
| PDF | Probability Distribution Forecast |
| QIS 5 | Quantitative Impact Study |
| QQ | Quantile-Quantile |
| RMS | Risk Management Solutions |
| S&P | Standard & Poors |
| SCR | Solvency Capital Requirement |
| TVaR | Tail Value at Risk |
| ULR | Ultimate loss ratios |
| USGS | U.S. Geological Survey |
| USP | Undertaking Specific Parameters |
| VaR | Value at Risk |

# D  Appendix D:  Bibliography and Further Reading

Council Directive (EC) 2009/138/EC 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II)

European Commission, 31st October 2011, "Draft Implementing measures Solvency II"

Board for Actuarial Standards, November 2009, "Technical Actuarial Standard R: Reporting Actuarial Information"

SCOR, November 2010, Enterprise Risk Management (ERM): A risk-based approach to the management of a (re)insurance company, Focus series

FSA, February 2012, Internal Model: Self Assessment Template

Lloyd's of London, March 2010, Detailed Guidance Notes Section 7 – Technical Provisions

CEA Insurers of Europe, 2007, Solvency II Glossary

CEIOPS, 2009, CEIOPs' Advice for Level 2 Implementing Measures on Solvency II: Articles 120 to 26 Test and Standards for Internal Model Approval

http://www.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all

Lloyd's Market Association "External catastrophe Model validation Conceptual framework"

Lloyd's Old Library, 9th of July 2012, "Validating External Catastrophe Models under Solvency II"

Lloyd's Market Association, 6th of July 2012, "External catastrophe Model validation ILLUSTRATIVE validation document No. 1 US Windstorm, high materiality"

Lloyd's Market Association, "External catastrophe Model validation Examples of model limitations & weaknesses"

EIOPA- FinReq – 11/023, 21 December 2011, Pre- Consultation Paper, "Draft proposal for Implementing Technical Standard on Undertaking Specific Parameters: Methods"

FSA, Oct 2007, Insurance Sector Briefing, " ICAS –Lessons learned and looking ahead to Solvency II"

Faulty of Actuaries Solvency & Capital Management Research Group, 20 October 2009:" Practical Implementation Challenges of Internal Models under Solvency II" (http://www.sias.org.uk/siaspapers/pastmeetings/view_meeting?id=October2009)

# E   Appendix E:  Worked Example of Risk Ranking of Operational Risk

E.1.1   We've built a case study for ranking different operational risks to illustrate the principles of risk ranking.

E.1.2   This is an example only.

## *Background*

E.1.3   Relevant business leader rated about 50 risks in the risk register.

E.1.4   Key inputs per risk are:

- Frequency (i.e. number/yr)  -  assume Poisson distributions.

- Severity (i.e. likelihood of loss within specified bands, say £0-1k, £1k – 10k, £10k - £100k, etc.) - assume uniform distributions.

- Upper cap on monetary loss possible.

E.1.5   Selections are judgement-based, but validated against data if available.

E.1.6   No dependency applied but stresses/sensitivities considered.

E.1.7   Simulated results used for capital setting.

E.1.8   Options considered: Stress tests and stochastic methods.

E.1.9   Most companies will need to review the results on several bases to form a judgement.

## *Approach*

E.1.10   Analysis – Stress Test Methods.

E.1.11   These require development of scenarios and a deterministic measure of cost in each case.

E.1.12   Some fictional operational examples:

- Scenario 1: Contract written outside underwriting limits and "limits loss" occurs:

- Typical policy limit: $1m.

- Breached policy has limit of $2m.

- Operational "loss": $1m.

- Scenario 2: Reputational event causes brokers to direct business elsewhere:

- Annual premium from broker $12m (i.e. $1m/month).

- Loss of confidence lasts:  8 months (judgement).

- Operational "loss": $8m .

E.1.13   Then order the scenarios to build the ranking.

## Analysis – Stochastic Methods

E.1.14   We ranked the simulated results for each approach – note: consistent colours are used to identify risks appearing more than once.

| Option | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Rank | Mean | SD | CV | Var 97.5 | Capital Alloc |
| 1 | UW Auth | RDS | RI | Del UW | Del UW |
| 2 | Staff Ret | Del UW | RDS | Claim systems | RDS |
| 3 | Emp Law Risk | Staff Ret | Phys Inv | Staff Ret | Staff Ret |
| 4 | Claim systems | UW Auth | Other | UW Auth | Claim systems |
| 5 | Del UW | Claim systems | Other | UW Review | UW Auth |
| 6 | IT | Phys Inv | Mkt Change | Emp Law Risk | Emp Law Risk |
| 7 | UW Review | Emp Law Risk | Other | Proj 1 | Phys Inv |
| 8 | Fraud Claim | UW Review | Other | Syst Exploit | Mkt Change |
| 9 | Syst Exploit | Mkt Change | Other | Fraud Claim | UW Review |
| 10 | Dis Rec Plan | Systemic claims issue | Other | Dis Rec Plan | Systemic claims issue |

E.1.15   The top risks reflect ranking method BUT,

E.1.16   The risks coming out on top are generally similar, which gives comfort.

E.1.17   Set against this, similarity is partly a result of no correlation in this model.

E.1.18   'Other' in Coefficient of Variation method – this refers to risks outside top 20 by capital allocation.  CV method is insensitive to size of risk. Use with care.

## Capital allocation – a pragmatic approach…

E.1.19   The first four methods are straightforward.  For capital allocation, there are many methods available – e.g. see references.  Some methods are built in the capital modelling software already.

E.1.20   In this example, we use a pragmatic approach:

1) Simulate model and sample component "risks".

2) Calculate capital based on total loss VaR – say 1 in 200 result.

3) Sort output data based on total operational loss.

4) Calculate the average of the total loss column using just the 1st row, then the top 2, 3, 4, Keep going until you've got enough rows ("n") such that the average is ~ the capital figure calc in step 2.

5) Now extract the top "n" rows from the full results set in order to derive the "tail dataset".

6) Calculate the average loss for each component risk in the tail dataset – the calculated average is taken to be the allocated capital.

## Communication

E.1.21   A picture is better than a thousand words, and thus it may be easier to present the results of the risk ranking graphically than in a essay.

E.1.22   We have demonstrated the use of treemaps, stacked plots and impact vs likelihood plots in this example.

## *Treemaps*

E.1.23   Pie charts may be quite commonly used but we believe treemaps may be better in this case.
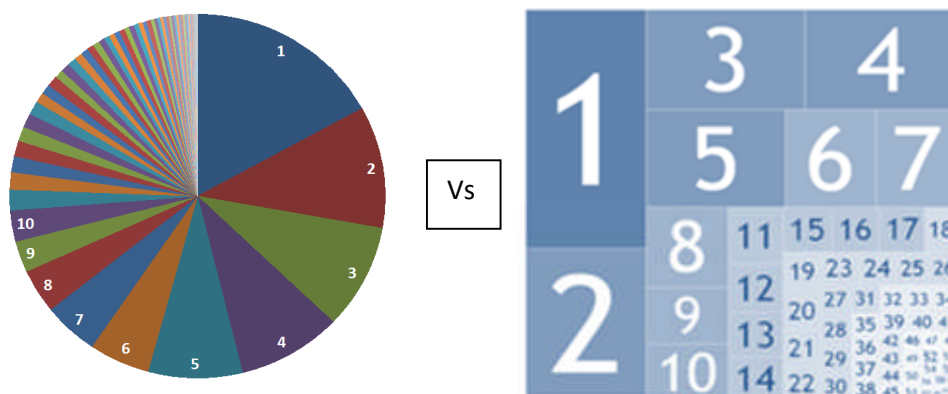


E.1.24   Both use area to represent amount of capital allocated to an area.

E.1.25   Numbers here indicate rank – e.g.  Number 1 = delegated underwriting using capital allocation method in our case study.

E.1.26   Treemaps are better because:

- Relative size of risk is better represented, especially for the contribution of small risks – e.g. 11 vs 18 etc.

- Risk names e.g. "Deleg UW" won't fit on pie but can fit on a treemap.

- It's easier to link colours as a third dimension – e.g. could use colour to reflect mean loss or line of business.

E.1.27   Stacked plots in the tail.

E.1.28   These charts identify how risks interact in the tail...
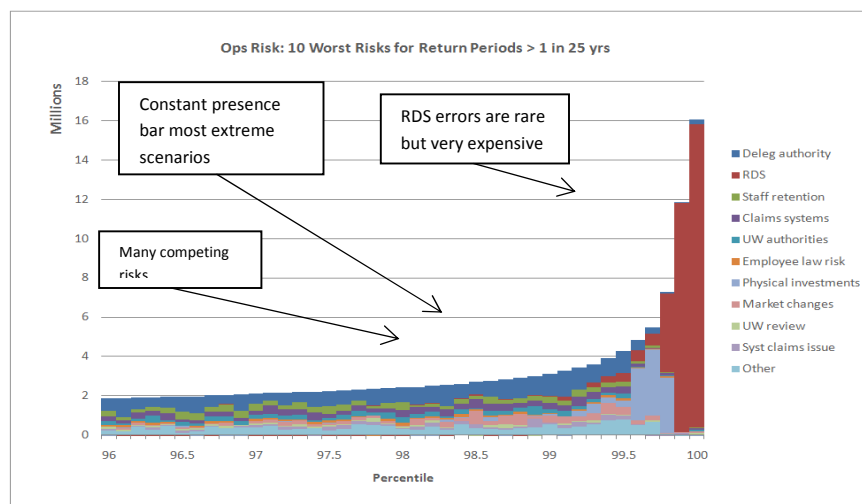
E.1.29   Clear away the "noise" and only three risks are of interest...
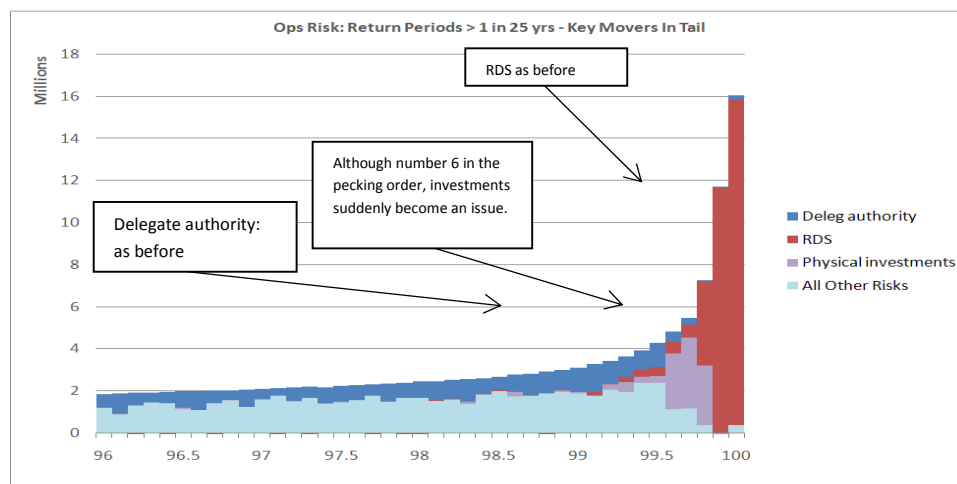


*Impact vs likelihood plot*

E.1.30   The positions of dots based on input parameters, no modelling needed.

E.1.31   The "blue" dots reflect top 10 risks by capital allocation (modelling needed here!).

E.1.32   Some blue dots appear in the bottom left corner. May be counter-intuitive initially, but the severity curve for these risks is very skewed hence they carry very low average losses and are very rare. This is one downside to plotting average severities.

E.1.33   Chart below considers input parameters for likelihood and severity for operational risks considered (if loss occurs):
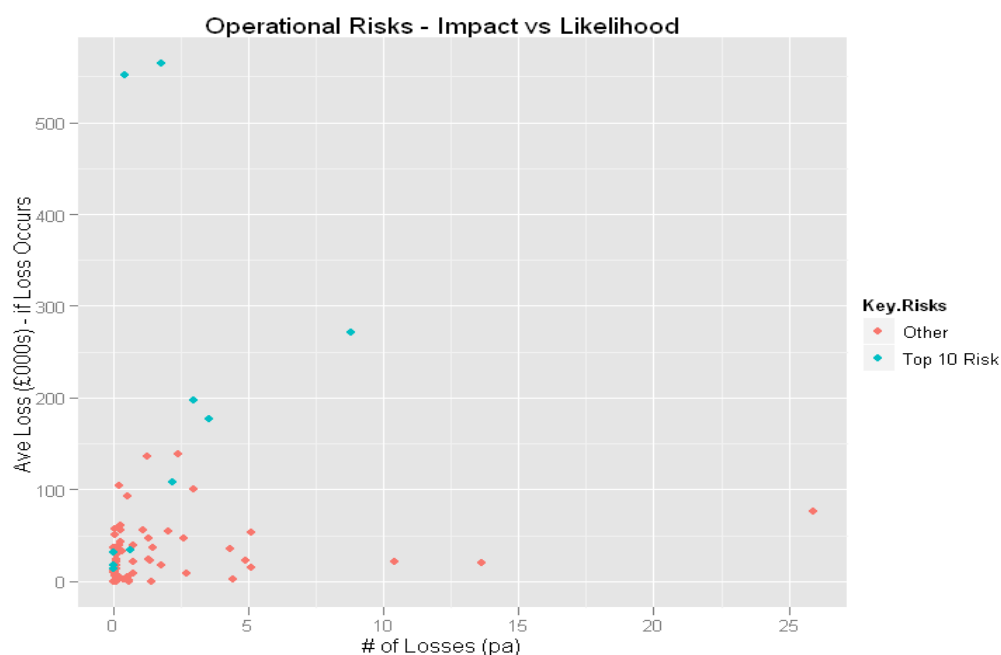


10.8.5   Charts below are similar to the above, but it is split by business activity driving the risks.

Operational Risks - Impact vs Likelihood
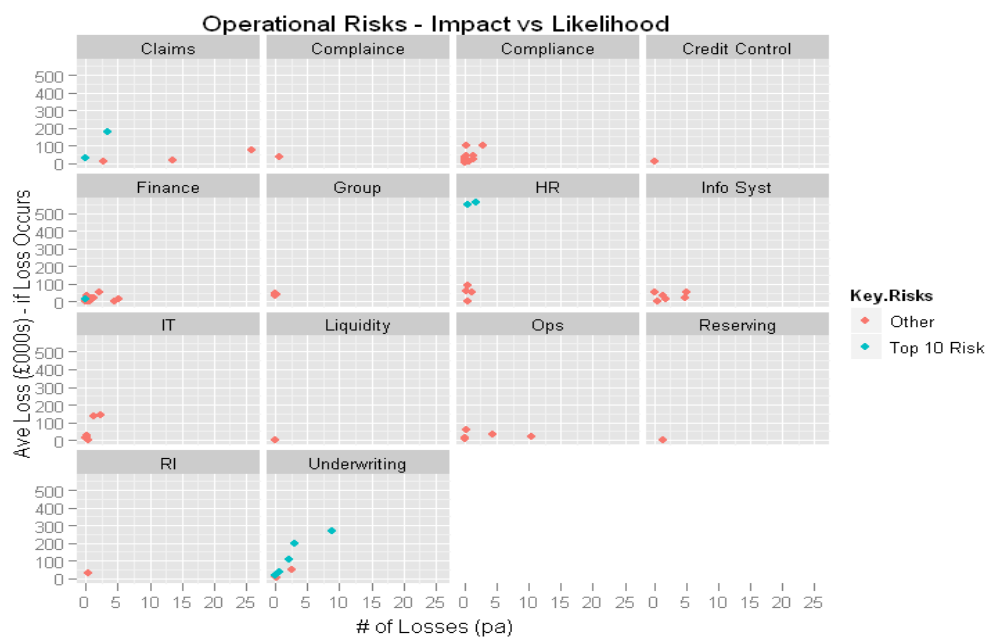
E.1.34   Key risks are where they were expected; in underwriting and claims related activities.

E.1.35   HR risks refers to the income lost if key staff members depart.

E.1.36   Impact vs likelihood – drill down by individual.

E.1.37   The charts can also be split by person responsible for these risks internally.



Operational Risks - Impact vs Likelihood

# F    Appendix F:  Validation Tools

F.1.1    Benchmarking

Benchmarking is a very common tool used widely across all actuarial fields to compare an internal position against industry or industry results. The principles for an internal model are the same – it provides useful validation of the model to compare it to other benchmarks.

Benchmarking can be used at many levels within the internal model. Examples include:

- Comparing distributions fitted to own data, to external data from other peers.

- Comparing model methodologies and documentation.

- Comparing model results to those from peers.

Sources of appropriate benchmarks may often appear limited, mainly due to competitor confidentiality. However, some examples of sources include:

- Comparisons between different subsidiaries in a group.

- Comparisons to publically available regulatory returns such as pillar 3 reports.

- Benchmarks provided by consultancies, reinsurers or professional bodies that see a range of market information from clients or members. Confidentiality can be maintained by anonymising data.

F.1.2    Stress and Scenario Testing

Stress and scenarios testing is a requirement of the ORSA. In this context, it is intended to identify possible single events (stress) or a combination of events (scenario) that could unfavourably affect the overall financial standing.

Stress testing can be used as a validation tool for the internal model, by comparing assessed stresses and scenarios to the internal model for consistency. For example, this might identify scenarios of risks that are not considered in the internal model at all or are seen to have a smaller effect than assessed.

F.1.3    Reverse Stress Testing

The objective of reverse stress testing is to identify hypothetical scenarios which are most likely to cause the current business model to become unviable. These scenarios may be identified through discussions with management or based on outputs of the model. The model can be used to highlight the key drivers of these scenarios and quantify the financial impact of these material risks.

F.1.4    'As if' Losses

The test is used to assess whether a model generates losses which are consistent with actual experience based on the replication of the footprint of actual past events when applied to current exposure. Catastrophe losses are typical used for this type of analysis. For consistency when

comparing against past events, an allowance is made for changes in risk profile, reinsurance arrangement etc.

F.1.5    Sensitivity Testing

Sensitivity testing involves adjusting model assumptions, structure or choice of methodology and reviewing the resulting changes in the overall model results. A common example is to test correlation assumptions as these are often set based on expert judgement and considered to have large uncertainties. Another common example is to assess simulation error, by running the model repeatedly with a number of different seeds and reviewing the resultant variability it results.

Unlike other tests, its main use is often not to demonstrate particular weaknesses, but more so to identify the key parameters and models which have most material influence on the overall model results. The areas seen to be most sensitive can then be given more focus.

It can also serve as a useful quality assurance tool in its own right, by trying to explain the features seen. If changing parameters leads to unintuitive changes in results, then it may be a sign of an area of model weakness. It is also useful to highlight model errors - for example, changing some correlations and seeing that there is no change in the model results may highlight a simple model error (maybe that the correlations are not actually being applied as intended!).

F.1.6    Profit and Loss Attribution

Profit and loss attribution is considered in more detail in its own section of this paper. It is similar to traditional actuarial actual versus expected testing as it involves reviewing the actual emergence of profits and losses as documented in the financial statements and assessing how these stand compared to the model. The objective is to show that the internal model is able to explain all sources and causes of profit and loss, with any significant differences highlighting that the model may not capture the full risk profile of the business.

It is important to note how profit and loss attribution differs from backtesting. Backtesting will typically be more focussed on comparing specific model output distributions to a number of historic datapoints. Profit and loss attribution in turn considers how recent results compare to the model expectations typically at a very high level across the whole model, with the aim of identifying consistency between the profits and losses seen and the risks modelled

F.1.7    Backtesting

This is often referred to in EIOPA texts as testing results against experience. The principle is to test high level results against historic experience, with appropriate adjustments made to bring historic results onto a consistent basis to be representative of the period being modelled (for example inflation adjustments).

Backtesting can be used at various levels of aggregation within the model. This testing is particularly useful at higher levels to ensure the consolidated picture across a number of model components is accurate. For example, backtesting a directly fitted large loss distribution may not add much value as it may be just repeating the fitting process (i.e. fitting the distribution to historic adjusted data). But comparing the resulting high level loss ratio distribution to adjusted historic loss ratios might be

more informative. The modelled loss ratio distribution will be a consolidated model of the combined underlying models for large losses, attritional losses, catastrophes and premiums. Hence, demonstrating the accuracy of the resulting loss ratio distribution will in turn give confidence in the underlying models and their interactions (dependencies). This is analogous in a way to a simple check-total spreadsheet test.

### F.1.8    Model Robustness

Model robustness tests are used to assess the stability of the stochastic simulations used in the calculation kernel. They can be used to show the convergence of the model at various key points of the distribution, and the stability of the model results based on different runs or starting seeds.

### F.1.9    Analysis of change

This is another common tool employed widely across the range of actuarial fields. It involves comparing and explaining movements (or lack of movements) over time, for example following a model change.

In keeping with one of the themes of Solvency II internal model approval – use of this tool is not limited just to model results. For example, it can also be deployed as a tool to validate data as it may quickly identify missing data for example.

### F.1.10   Hypothetical portfolio

This is a specialist tool for validating appropriateness of model methodology. It involves running the model with a standard or benchmark set of parameters – a hypothetical portfolio. The model output can then be compared to different models run with the same hypothetical portfolio. It may be a requirement of regulators to see a model run with a pre-specified hypothetical portfolio as part of the approval process.

It is most commonly used for investment and economic models. An investment model and input ESG can be benchmarked against others by running it with a standard set of investment instruments (the hypothetical portfolio) and comparing the output to the benchmark models run with the same hypothetical portfolio.

It can also be employed for non life insurance risk models. For example, regulators might want to see the model run with a standardised mix of business. However, this is not expected to be widely used for this as the nature of insurance risks is often perceived to be such that standardised models do not work and that companies will develop their own models to be very specific to the nature of their own business (such that the model will not be appropriate for other portfolios).

### F.1.11   Standard SCR Formula Comparison

The objective of this test is to understand and explain the deviations in internal model results from the standard formula results. Typically results from each main standard formula risk module will be compared to corresponding results from the internal model, and areas where the internal model produces figures significantly below the standard formula value should be investigated further.

### F.1.12   Statistical Tests

A variety of statistical tests such as goodness of fit tests, quantile-quantile (QQ) plots and formula tests (e.g. sum of squares) can be used to assess the suitability of historical data for parameterisation and assumption setting.  Statistical tests can be performed manually via Excel or by using statistical packages.

F.1.13   Inspection of outputs/reasonableness checks (sense checks)

This validation tool can be used to check the high level reasonableness of outputs in order to see whether they are in line with expectations.  Back of the envelope calculations and widely accepted rules of thumb may be used as part of the checking process.

F.1.14   Reconciliation of outputs

This check involves comparing model outputs to model inputs in order to ensure that information flows correctly through the model.

F.1.15    Re-performance testing and review of model calculation

The objective of re-performance testing is to provide an independent check of the accuracy of calculations performed in the model by replicating particular model calculations outside of the model (e.g. in a spreadsheet) and comparing these external results to the modelled output. Any unexpected model results should be highlighted. Thresholds and tolerances for differences between the two methods can be set to guide the analysis.

F.1.16   Validating data inputs

The tool is used to check the accuracy of input data being used in the model, for example by reconciling current period data to prior period data.

F.1.17   Qualitative reviews and Use Test

Qualitative reviews and the use test are important for ensuring the model is optimised for its purpose – to provide useful business information about risks.

The development of the model should be focussed on the specific risks and characteristics of the business. The qualitative review can ensure this is the case by comparing the modelling to the top-down management view of risks faced by the business.

The use test is generally regarded as the most important test for model approval. The use test is explored in detail in its own section of this paper.

F.1.18   Review and benchmarking by business

The objective of a review by the business is to ensure the appropriateness and accuracy of the model design, assumptions and methodology by consulting those individuals and departments with the most knowledge of a specific area. The review could be performed by, for example, the finance department or the underwriters. The experts could also use benchmarking to inform their review and recommendations, for example by reviewing a particular assumption for consistency across the business.

F.1.19   Review and benchmarking by Experts

The objective of a review by experts is to provide an independent review and challenge of key model design issues, assumptions and methodologies to ensure their appropriateness and accuracy. The reviewer should be independent of the immediate business teams or units responsible for setting the assumption and could be performed by either internal or external experts with specialist knowledge in the given area, for example, actuaries, engineers. The experts could use benchmarking to inform their review and recommendations.

F.1.20   Methodology review against market practice

The objective of this tool is to ensure that the modelling methodology remains in line with market best practice and takes into account new or evolving methods. The process should involve a review of the key methodology by a suitable reviewer and provide an overview of where the model sits in the market. Recommendations for areas where the methodology requires updating or improving should be documented and justified.

F.1.21   Review of model documentation and evidence

This tool is designed to ensure that there is appropriate documentation in place to support the

Modelling approach taken. Documentation should exist in many areas, including the following:

- Model structure and framework
- Methodology and assumptions including parameterisation
- Outputs
- Interpretation of results
- Validation tests

There should be evidence supporting the model including justifications for simplifications made and details of any judgements, weaknesses and limitations.

F.1.22   Assessment of model controls and governance

The objective of model governance is to ensure that the internal model complies with the Solvency II requirements and that a framework is in place throughout the business to support this compliance.

F.1.23   Audit Trail

The aim of the audit trail is to ensure that all developments of the model (including input data changes) are logged appropriately. This will also include documentation supporting and justifying the parameter choices.

In addition, it is a requirement of the guidance that all data used in the calculation of the SCR is time and date stamped.

# G   Appendix G: Calibration

We note the discussion below was presented at GIRO in 2010.  Views on the different calibration methods have moved on in this time.

## *A mathematical diversion*

Let CF be the sum of the all cash-flows relating to a particular cohort of claims (no discounting here!), covering both the paid at time 0 and future payments to ultimate.

So, assuming that reserves are set as the probability weighted average of future cash-flows (mean of the distribution), given the information we know at the moment, D0, the opening best estimate ultimate, U, is set according the Solvency II requirements as the mean of the cash-flows,

$$U = E[CF | D0]$$

Conditioning on the total information known at time t, Dt, we get

$$Ut = E[CF | Dt].$$

That is, the best estimate ultimate at time t is the expected value of the full cash-flow series, given the information at time t (which certainly includes the cash-flows to that point).  Note that U=U0.

We also consider the uncertainty remaining in CF at time t (i.e. the ultimo variability that we would be capturing in a study at time t, for the cohorts being studied at time 0).

$$Let\ Vt = Var(CF | Dt)$$

V0 is variance of the ultimate at time 0 - this is the "ultimo" view of reserving risk for the existing study.

Now, taking inspiration from a Deloitte's LMAG presentation of 2009, we recognise that the volatility of CF now, includes an element for the degree of uncertainty that will still be there at time t, and use the law of total variance (which applies with no further assumptions):

$$Law\ of\ Total\ Variance:\qquad Var(X) = E(Var(X|Y)) + Var(E(X|Y))$$

(Deloitte's worked with reserves as at time 0, I have switched to ultimates as this is equivalent and, to my mind, clearer, however the arguments are the same)

Using this at time 0 for X=CF, conditioning on the information at time t, they got:

$$V0 = Var(CF|D0) = E(Var(CF|D0|Dt)) + Var(E(CF|D0|Dt))$$

$$= E(Var(CF|Dt)) + Var(E(CF|Dt)) = E(Vt) + Var(Ut)$$

$$And\ specifically\ V0 = E(V1) + Var(U1)$$

Deloitte's then noted that the one year view was the change between U0 (which is known at time 0) and U1, and so the variance of the one year CDR was equal to the variance of U1, and re-arranging,

$$\text{Var(1 year CDR)} = \text{Var(U1)} = V0 - E(V1)$$

In words the 1 year variance is the ultimo variance now minus the current expectation of the ultimo variance in one year's time.

One major significance of this is that the LHS is the one year calibration, while the RHS is the ultimo volatility less something else which is positive. Thus the one-year calibration must be less than the ultimo volatility (at least in variance – this result doesn't necessarily apply to tail percentiles).

Now deviating from Deloitte's work, what happens if you add the ultimate CDR, starting at time 1 to this result. The left hand side is obvious:

$$E[\text{Var(1 Year CDR + ultimate CDR from one year)}|D0]$$

$$= E[\text{Var(Ultimate CDR starting now)}|D0] \text{ (as the intermediate ultimate cancels)}$$

$$= V0$$

So next consider:

$$E[\text{Var\{ultimate CDR from one year\}}|D0] = E[\text{Var\{CF - U1| D1\}}|D0]$$

$$= E[\text{Var\{CF - E(CF|D1) | D1\}}|D0]$$

$$= E[\text{Var(CF|D1)}|D0] \text{ (the expected ultimate at t=1 is known within D1)} = E[V1|D0]$$

So

$$E[\text{Var(1 Year CDR + ultimate CDR from one year)}|D0] = V0 = V0 - E[V1|D0] + E[V1|D0]$$

$$= E[\text{Var(1 Year CDR)}|D0] + E[\text{Var(ultimate CDR from one year)}|D0]$$

$$= E[\text{Var(1 Year CDR)}|D0] + E[\text{Var(ultimate CDR from one year)}|D0]$$

$$E[\text{Var(1 Year CDR + ultimate CDR from one year)}|D0]$$

$$= E[\text{Var(1 Year CDR)}|D0] + E[\text{Var(ultimate CDR from one year)}|D0]$$

This looks like its saying that the one year CDR is uncorrelated of the future deterioration. Note that this makes no claim that the CDRs are independent, just uncorrelated.

Remember, the only assumption we made in this process was that we reserved at the mean of the distribution of future cashflows (given the current knowledge)

Actually this result seems intuitive. It states that, assuming we take into account all information available at time t in setting our reserves as the mean expectation, then a deterioration having happened in the current period should not influence our expectation of what will unfold in the next

period.  If we expect a year which has just deteriorated to continue deteriorating, we should have reflected that within our ultimate selection and reserved more strongly (until we no longer expected that).

(This is basically a result from the fact the time series of ultimates form a Martingale.)

So in conclusion, three (non-exhaustive!) theoretical tests for a one-year calibration method are:

- It produces lower variance than the ultimo volatility

- which is a weak case of,

- The variance of the CDRs in different periods (generated by a single consistent method) should sum to the variance of the CDR over the total period

- or equivalently

- The CDRs in different periods (generated by a single consistent method) should be uncorrelated

## *One year Calibration Methods*

**Perfect foresight:**

- Project ultimate volatility using an existing method

    o   In our empirical tests we used a Mack bootstrap based on Incurred data

- Set the ultimate which will be recognised next year as the output of this ultimate volatility method

**Key Points**

- Should be prudent (we would expect ultimate volatility to be higher than one-year volatility)

- Easy to calculate – no new methods required

- Not directly in compliance with Solvency II standards (not a one year view)

- Consistent with ultimate volatility measures (still likely to be used in ORSAs, Lloyds etc)

**Simulated Re-Reserving:**

- Use an existing method to project the next periods emergence of the data triangle

    o   Commonly this output would be part of an ultimate volatility study (e.g. bootstrapping)  so would be extracted from here

    o   In our empirical tests we used a Mack bootstrap based on Incurred data, taking only the simulated next diagonal in the Incurred triangle (not the ultimate)

- On each simulation get an augmented triangle

- Apply a reserving method to this triangle

    o   This should be consistent with the method used to project the next periods emergence.  E.g. if bootstrapping was used the same underlying Chain Ladder assumptions should be used in the reserving method

- This generates a set of ultimates for this simulation (a deterministic calculation given the stochastic input)

- Repeat for each simulation's triangle – get a distribution of projected ultimates

**Key Points**

- It is a true one-year method

- Generally consistent with ultimate volatility measures (still likely to be used in ORSAs, Lloyds etc)

- No new assumptions/parameters required

- Slightly more complex to carry out

- Will rely on a standardised reserving approach – how appropriate is this for reserving the triangle?

- Re-reserving step mechanical

    o Will not adjust method for very large claims etc.
    o No expert judgement in step between ultimate volatility and one-year calibration
    o Black box?

- Satisfies the theoretical test above

*Ref: England - The Ultimate and One-Year Views of Reserving Risk with Respect to Solvency and Risk Margins Casualty Actuaries of Europe Fall Meeting 2009*

***Proportion Emergence:***

- Project ultimate volatility using an existing method

    o In our empirical tests we used a Mack bootstrap based on Incurred data

- Set the ultimate in one year's time as the weighted average of the current ultimate and the perfect foresight ultimate (we simulated the weights to use from a Normal distribution)

**Key Points**

- How do we parameterise the weight?

- Consistent with ultimate volatility measures (still likely to be used in ORSAs, Lloyds etc)

- It is a one-year measure

- Highly practical to implement

- Fails the theoretical test above if weights are deterministic (future CDRs will be perfectly correlated)

- Assumes that the intermediate projected ultimates between now and fully run-off don't become overstated and come back in (at least if deterministic weights are used).

How we parameterise the weights

- In order to perform empirical tests we needed to parameterise the weights. The methodology we used is below (this is not necessarily recommended)

- The key formula is $Ult_{y,t+1} = Alpha_t * Ult_{y,t} + (1 - Alpha_t) * PF\ Ult$

    o Where $Ult_y, t+1$ is the ultimate for cohort y at development period t
    o $Alpha_t$ is the weight at development period t
    o And PF Ult is the (simulated) perfect foresight ultimate – the true value that we will know once the business is fully run-off

- We assumed current booked ultimates are correct (and hence we know the perfect foresight)

- Look at historical ultimates, and the one-year time step movements in these

- Get a triangulation of the Alpha values which would be required to have correctly predicted each movement

- Average these by development period

- (if simulating the weights, also take the standard deviation of this column)

**Merz-Wüthrich:**

- Use the Merz-Wüthrich formulae to generate a one-year standard error

- Use this to parameterise a distribution

  o In our empirical tests, we assumed a lognormal distribution, with expected value at the current level of the ultimate, and standard error from Merz-Wüthrich

**Key Points**

- Complex formula, but can be implemented in Excel

- New approach, new strengths and weaknesses to be understood

- Should produce same standard error as the simulated re-reserving (if using Mack bootstrapping and a standard Chain Ladder)

- Merz-Wüthrich formula relies on a standard formula, which in turn relies on a standard chain ladder methodology for reserving, with no tail

  o Is this appropriate for reserving the triangle?

- How is it made consistent with the simulated ultimate volatility?

- Re-reserving step is mechanical (relies on chain-ladder)

  o Will not adjust method for very large claims etc.
  o No expert judgement
  o Black box?

- Satisfies the theoretical test above

*Ref: Merz & Wüthrich - Modelling The Claims Development Result For Solvency Purposes*

**Hindsight Re-Estimation:**

- Takes a triangulation of historic best estimate ultimate

- Looks at historical adjustments from period to period

  o In our empirical tests, we took one step adjustments

- Adjusts these for exposure (taken to be the booked reserves at time 1)

- Re-applies these for future periods

**Key Points**

- Requires a triangulation of best estimate ultimates

- New approach, new strengths and weaknesses to be understood

- Empirical distribution – will only consider results as extreme as history

- Unrelated to reserving process

- May be made consistent with ultimo view – but this will require a change to this methodology for ultimo as well

*Ref: Houltram - An empirical approach to insurance liability prediction error assessment - The Institute of Actuaries of Australia XVth General Insurance Seminar 2005*

**QIS 5 USP Method 1:**

- Takes a triangulation of historic best estimate ultimate claims, plus core data (we used incurred)

- The method essentially looks at how much movement there has been historically between:

  o The outstanding claims in one period
  o The sum of the following periods incremental claims and its closing outstanding claims

- It is assumed that, on average, these should be the same

- Volatility is measured based on how much movement there has been (looking at calendar periods in aggregate)

- This gives a standard deviation – we then simulated the reserve deterioration from a Lognormal distribution

**Key Points**

- Requires a triangulation of best estimate ultimates

- New approach, new strengths and weaknesses to be understood

- A prudent reserving policy will show as an increase to the volatility in this method

- Unrelated to reserving process

- May be made consistent with ultimo view – but this will require a change to this methodology for ultimo as well

*Ref: European Commission - QIS 5 Technical Specifications - 2010*

*Empirical Testing*

## The Test

Data triangle, Paid or Incurred
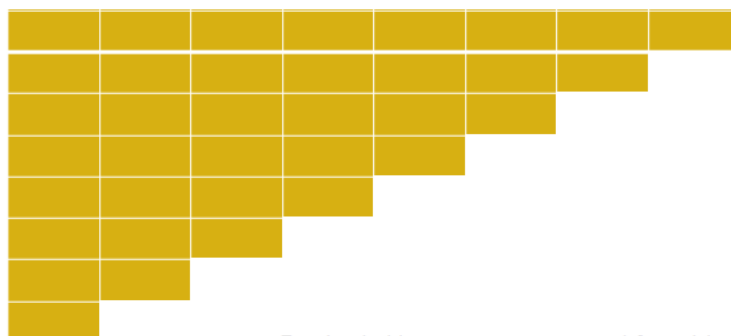
Booked Ultimates

Booked ultimates are extracted from historical returns for each of the past 10 years, providing a triangulation of historic ultimates
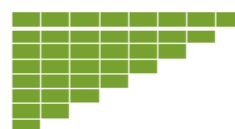
## The test: step 1
## Exclude latest position

Data triangle, Paid or Incurred
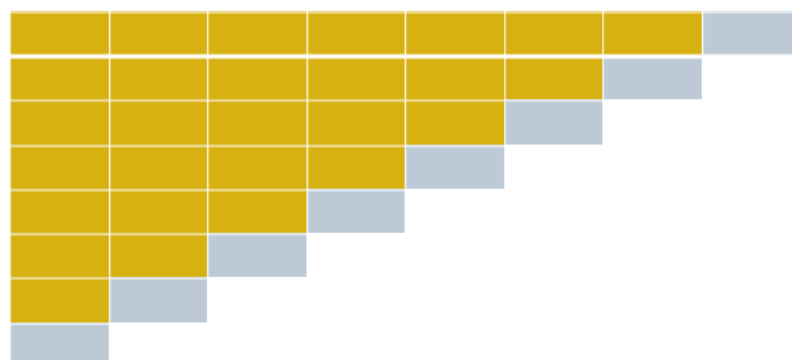
Booked Ultimates
one period ago
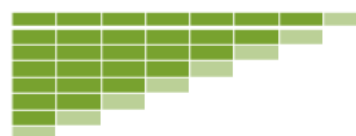
## The test: step 2
Use projection method to predict ultimate one year after data cut



Data triangle, Paid or Incurred

Projected one-year ultimate (distribution)

Total
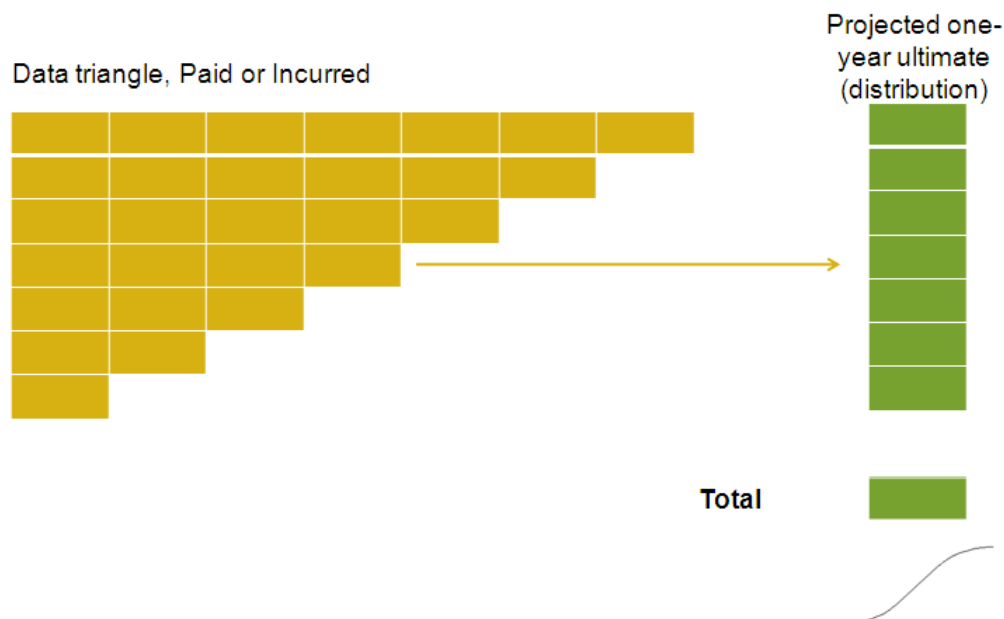
## The test: step 3
Bring back actual booked ultimate in latest period



Projected one-year ultimate (distribution)

80th Percentile

Booked Ultimate
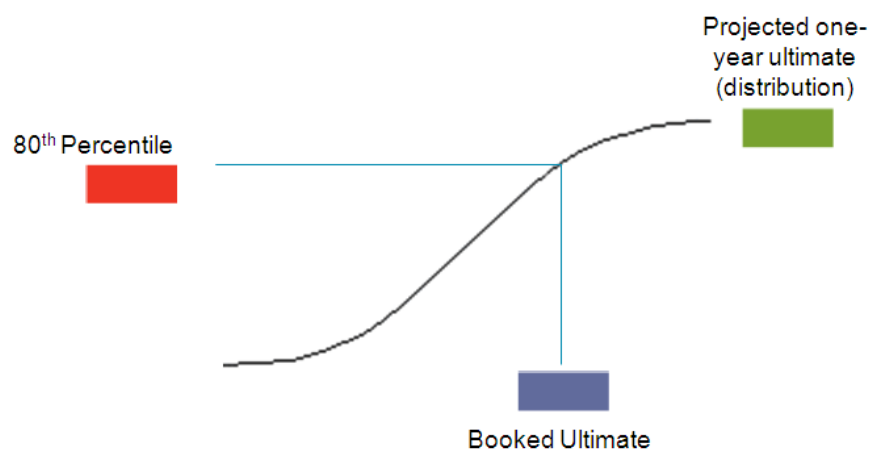
Interpret this as, given the method, the company was 80% unlucky over the year

## The test: Step 4
## Repeat for multiple companies (209)
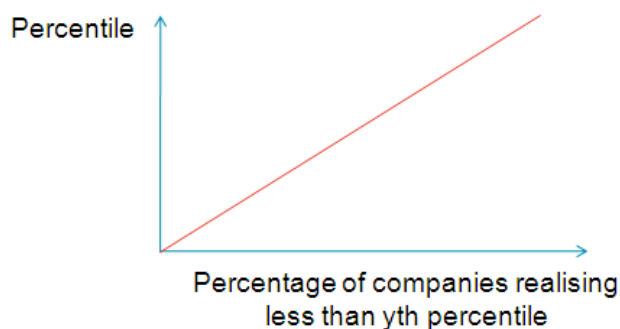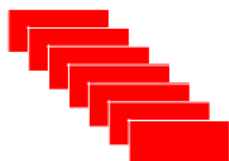


If the method is a good predictor, would expect roughly half to be (up to) 50% unlucky etc.

Sorting and plotting these "realised" percentiles gives us a plot of predictiveness
The closer the line is to a 45° line, the more predictive the method is

A note on this test:

- This test is similar to a test performed by the reserve uncertainty working party in their 2007 and 2008 papers.

- In their test they simulated both past histories and future emergence based on the same models and parameters

- They then applied reserve uncertainty methods to the simulated histories and generated projections of the 99th percentile for each past history

- Next they compared the projected 99th percentile to the simulated future emergence

- They found the proportion of cases where the future emergence was greater than the simulated 99th percentile was generally above 1%

- This was interpreted as weaknesses in the reserve uncertainty methods

- However England and Cairns showed that this behaviour happened even in simple cases, and that the test itself may be more complex and not as easily interpreted

- As such the following results should not necessarily be taken as implying that the methods are flawed even if the outcome is not a straight line. However we believe that this test may still be useful as a comparative basis.

*Ref: Gibson et al - ROC/GIRO Working Party Best Estimates and Reserving Uncertainty Paper for GIRO 2007*

*Ref: Bruce et al - ROC Working Party Reserving Uncertainty Paper for GIRO 2008*

*Ref: England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009*

Data:

- In order to apply these tests we have taken data from FSA returns

- Data was extracted where there was at least 5 years of returns

- The data covered multiple companies and classes

- Each dataset was then reviewed, and where the data appeared erroneous it was excluded

- Bias was to keep in data though – we ended up with 209 datasets

- Though there were a few issues…



Data – some triangles would be difficult to predict any future uncertainty from
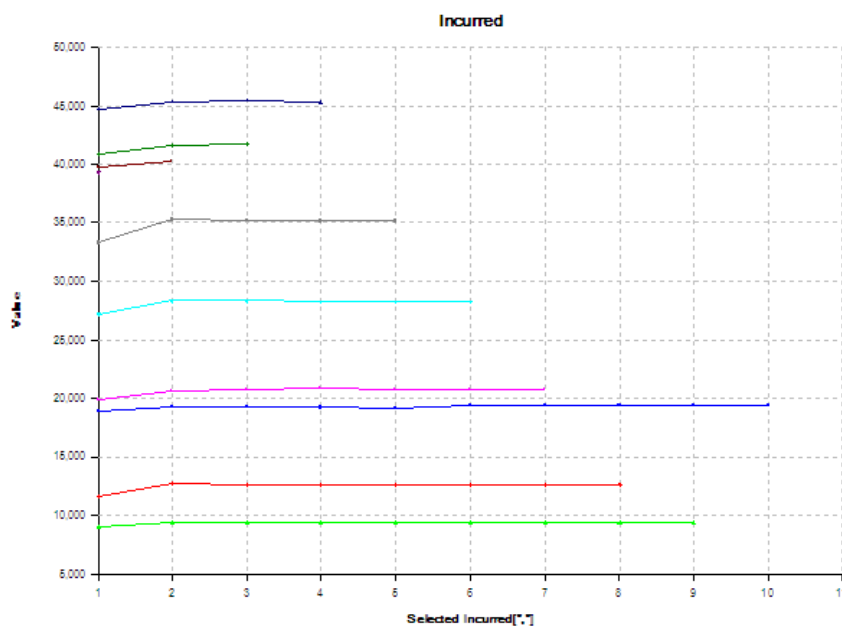
## Data – some triangles which are just difficult to predict were also included
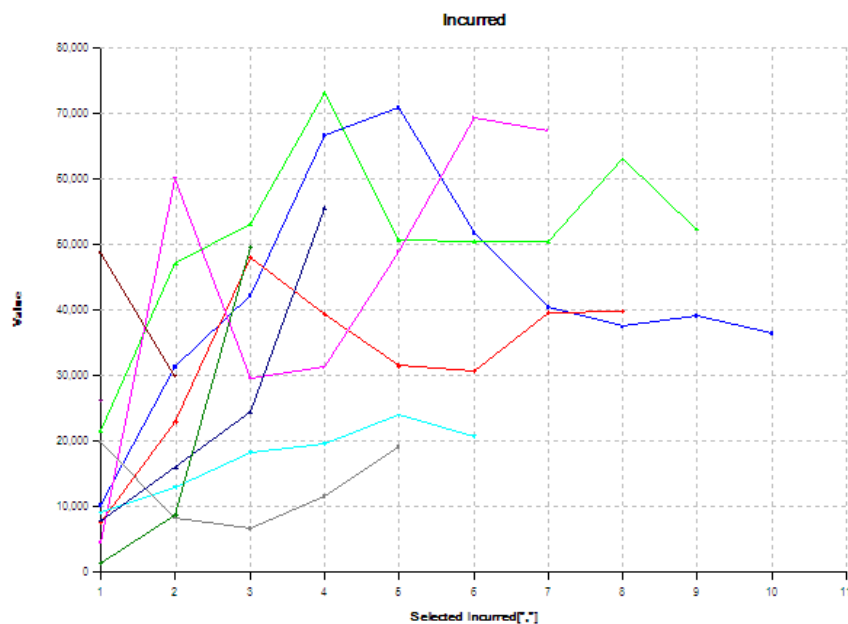


## Data – certain extreme movements were excluded
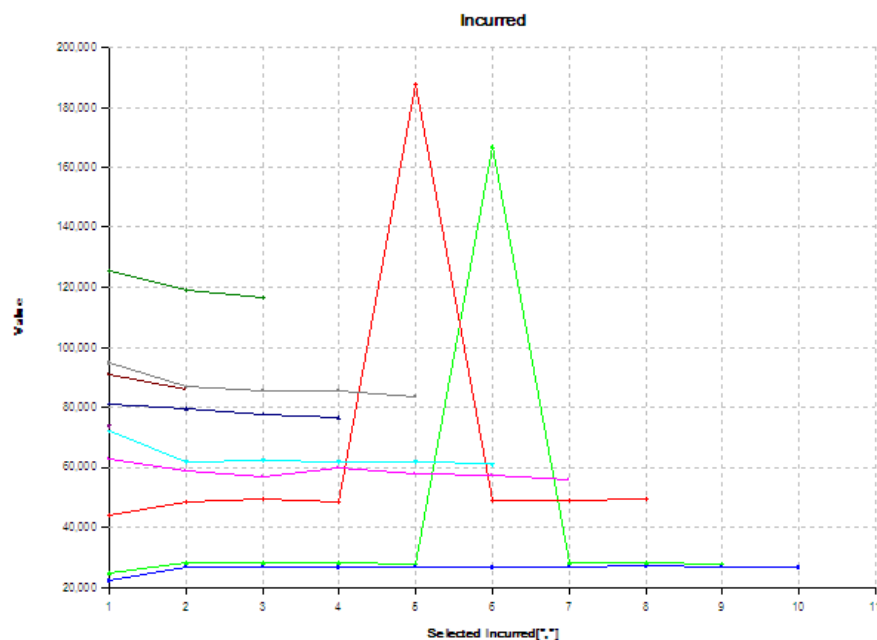
## Data – as were some calendar effects
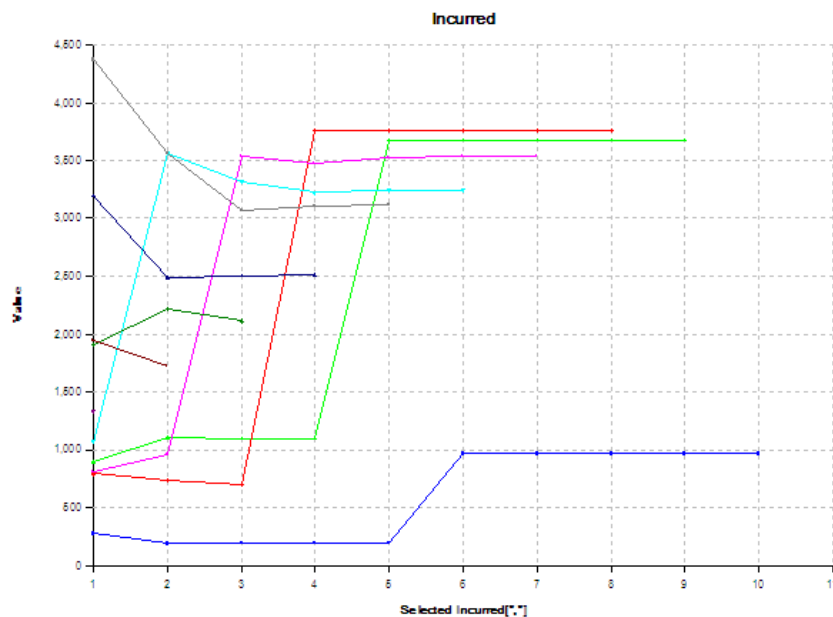


Simulation error in this test:

- This test has only been performed for approximately 200 companies.

- We are assuming these are 200 independent trials

- 200 companies / classes is a lot to process!, but in the context of stochastic modelling, 200 simulations is very little

- As such, we should expect the results to be subject to potentially significant simulation error

- We have tested the extent of this by simulating 200 draws from a Normal distribution, and then converting these back into the true percentiles.  Again plotting these should give a 45$^o$ line.

- We've tried this for 10,000 simulations – the range around the 45$^o$ line should give us an idea of the simulation error we are exposed to

## Simulation error results - Normal

Simulation error in assessing realised percentiles with 200 data points
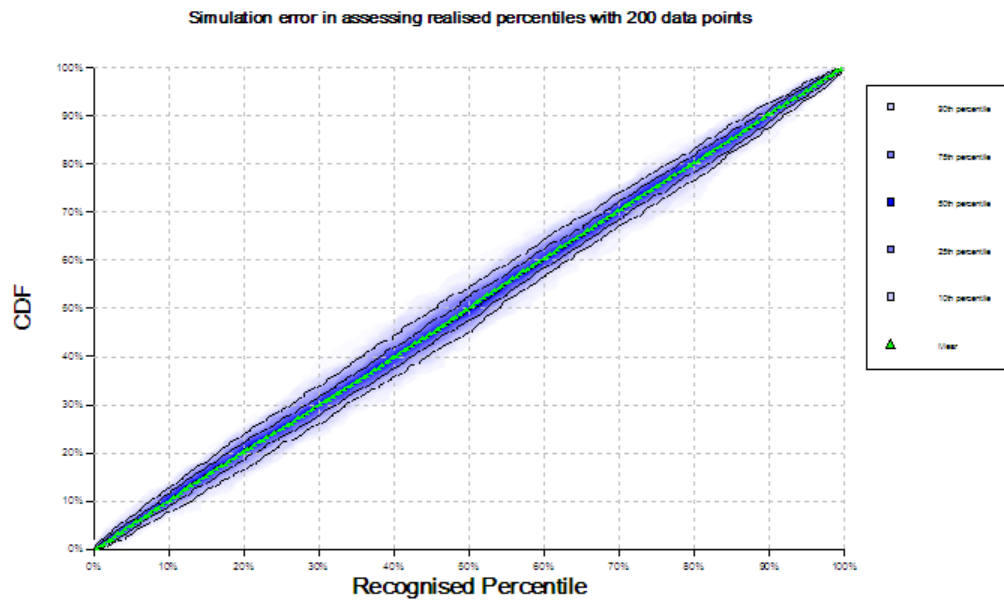


How to read the test:

- Taken from England and Cairns, the following slides give an indication on what we would expect to see under certain violations of the test

- For example – if the methods under-predict the volatility we will not see a 45$^o$ line – what will we see?
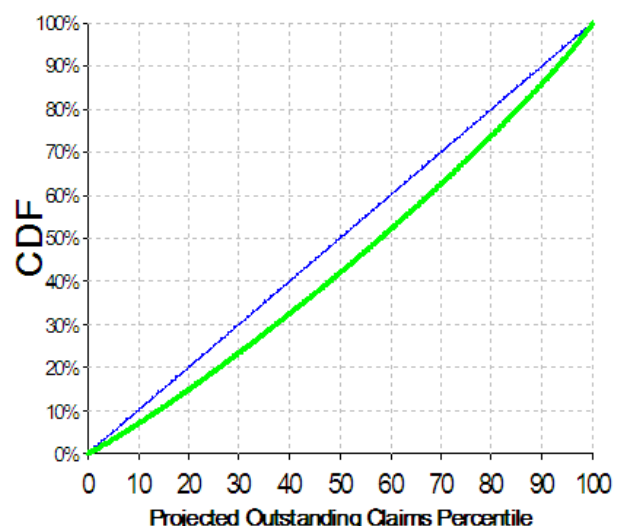
## Some background on what we would expect to see

- Let's assume the test is valid for the moment.
- The test looks at the proportion of simulated data sets for which the (bootstrapped) projected $x^{th}$ percentile is exceeded by the (simulated) true future outstanding
- Plotting this for all percentiles, with the $x^{th}$ percentile on the $x$-axis, and the (equivalent) proportion of simulations where the "true" outstanding is less than the projected $x^{th}$ percentile should give a uniform distribution, $Y=X$
- Effectively this is plotting the CDF of the distribution function Pr("true outstanding"$< x^{th}$ percentile) for $0<x<100\%$
- But how does the $Y=X$ line change as we violate the uniformity assumption?
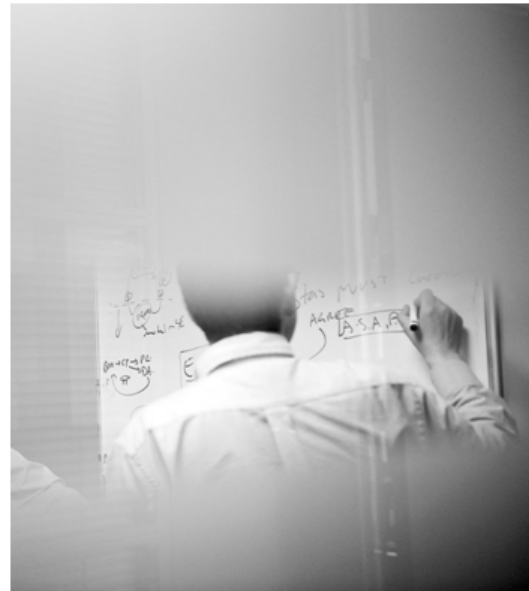


Slide taken from England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009
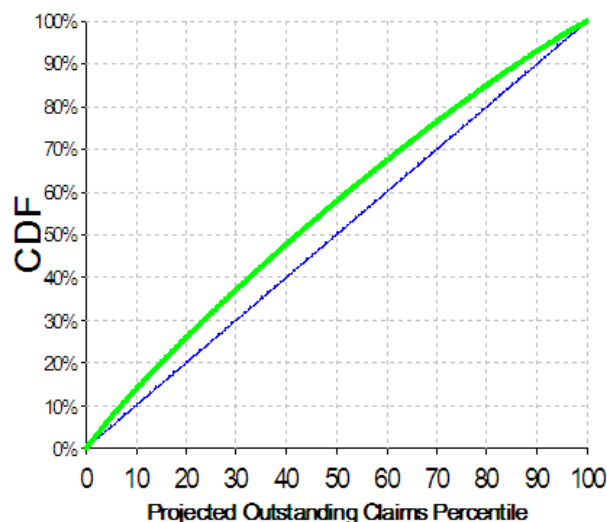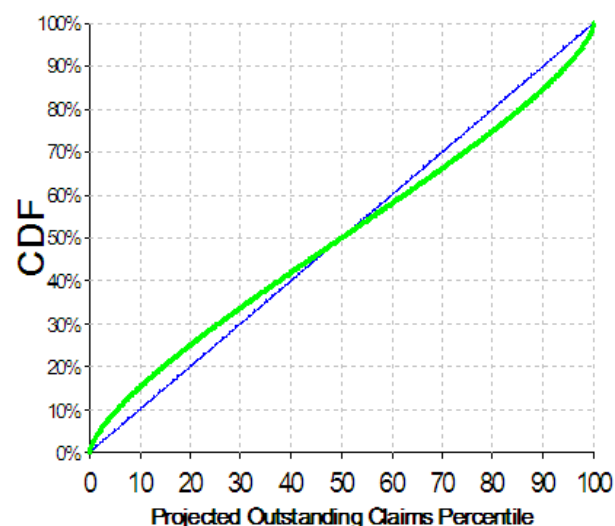
## Some background on what we would expect to see

- According to the test the CDF of the proportion of "true" outstandings below the projected percentile should be uniform, $Y=X$ (shown in blue)
- The green line shows the impact of changing from the true distribution (assumed to be Normal(8,1) for this simple example, to one with a biased mean below the true mean, with the correct standard deviation, (assumed to be Normal(7.8,1) for this simple example)



Slide taken from England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009

## Some background on what we would expect to see

- According to the test the CDF of the proportion of "true" outstandings below the projected percentile should be uniform, *Y=X* (shown in blue)
- The green line shows the impact of changing from the true distribution (assumed to be Normal(8,1) for this simple example, to one with a biased mean above the true mean, with the correct standard deviation, (assumed to be Normal(8.2,1) for this simple example)

Slide taken from England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009
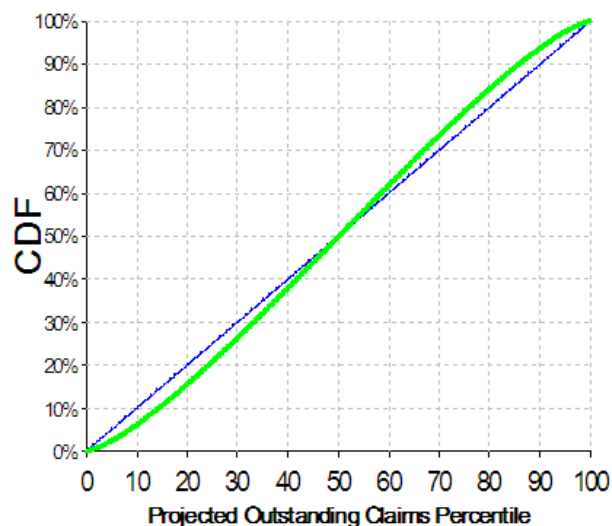
## Some background on what we would expect to see

- According to the test the CDF of the proportion of "true" outstandings below the projected percentile should be uniform, *Y=X* (shown in blue)
- The green line shows the impact of changing from the true distribution (assumed to be Normal(8,1) for this simple example), to one with the correct mean, but a biased standard deviation below the true value, (assumed to be Normal(8,0.8) for this simple example)

Slide taken from England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009

## Some background on what we would expect to see

- According to the test the CDF of the proportion of "true" outstandings below the projected percentile should be uniform, *Y=X* (shown in blue)

- The green line shows the impact of changing from the true distribution (assumed to be Normal(8,1) for this simple example), to one with the correct mean, but a biased standard deviation above the true value, (assumed to be Normal(8,1.2) for this simple example)



Slide taken from England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009

## Some background on what we would expect to see

- According to the test the CDF of the proportion of "true" outstandings below the projected percentile should be uniform, *Y=X* (shown in blue)

- The green line shows the impact of changing from the true distribution (assumed to be Normal(8,1) for this simple example), to one with the correct mean, and the correct standard deviation, but too much skewness, (assumed to be a Lognormal distribution with a mean of 8 and a standard deviation of 1 for this simple example)
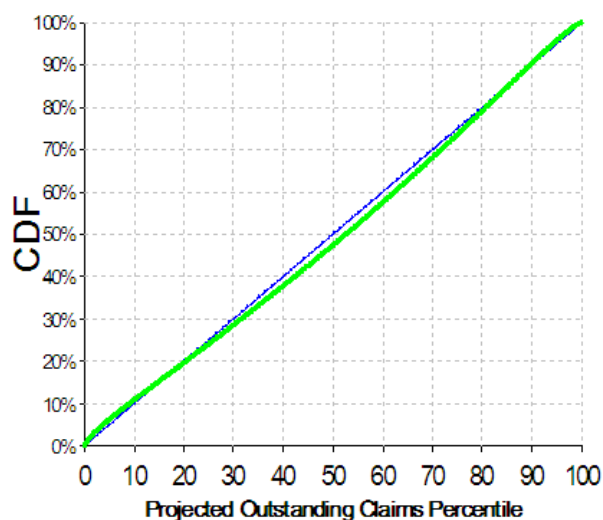


Slide taken from England and Cairns - Are the Upper Tails of Predictive Distributions of Outstanding Liabilities Underestimated when using Bootstrapping? – GIRO 2009

Other issues:

- Best estimate?

- o Have companies historically reserved as the mean of the distribution?  If not it could bias the results

- No review of methods

  - o Some methods may not be appropriate in some cases – this has not been tested

- Mechanical methods – no user judgement

  - o This may reduce the credibility of the results.  It may be the case that an experienced user can improve the results of some methods more than others.

- Bugs in the coding

  - o There could be errors in our logic or implementation

## Results – incurred



One Year Calibration Predictiveness Test Results - Incurred
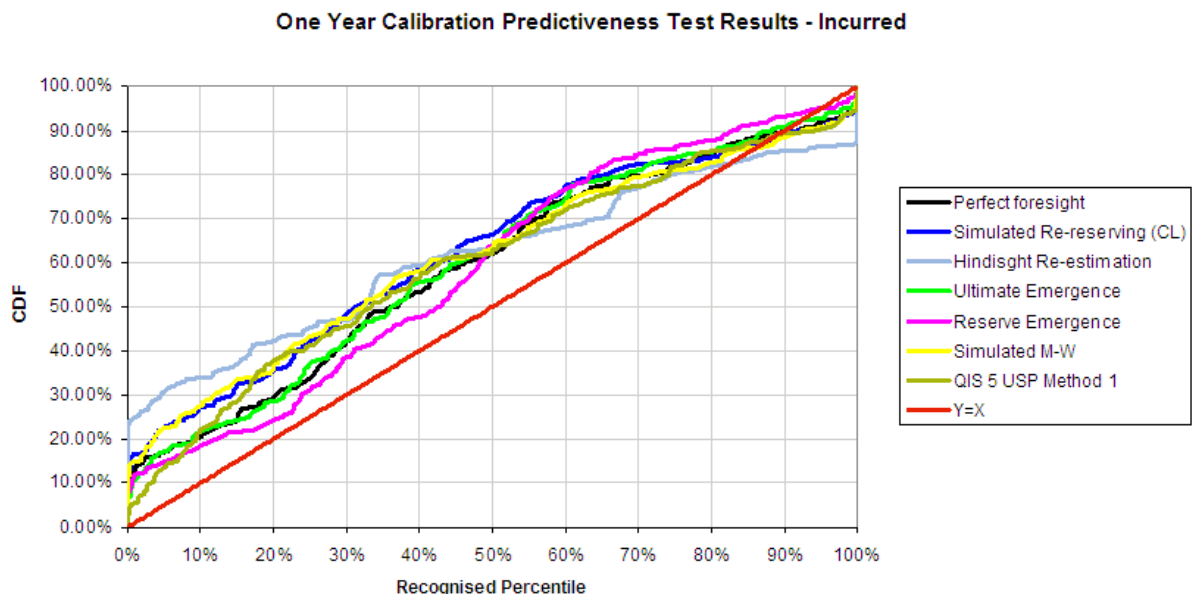
Which method fits bests?

- We have a variety of methods, none of which fit perfectly

- In order to test which fits best we have calculated two test statistics:

  - o Squared error (observed percentile minus theoretical percentile)$^2$ : The best method is assumed to be the method which minimises this

  - o Chi-squared test (with 25 equally sized bins): The best method is assumed to be the method which minimises this
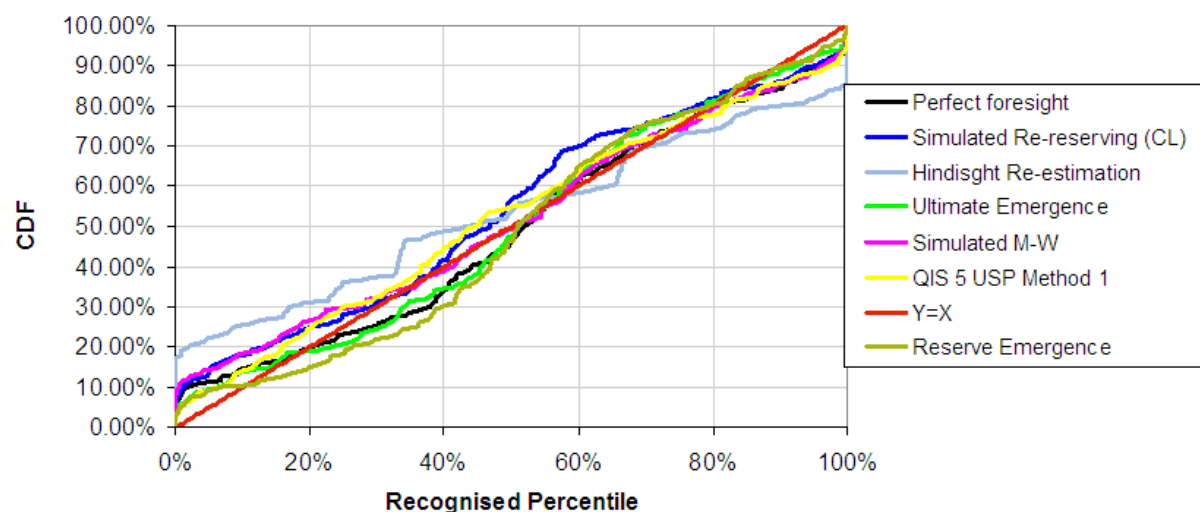
## Results – Incurred

|  | Total Squared Error | χ² test statistic |
|---|---|---|
| Perfect foresight | 234% | 113 |
| Simulated Re-reserving (CL) | 420% | 190 |
| Hindsight re-estimation | 515% | 420 |
| Ultimate emergence | 252% | 109 |
| *Reserve emergence* | *203%* | *93* |
| Simulated MW | 384% | 187 |
| QIS 5 USP Method 1 | 206% | 94 |

## Results – based on Paid data

### One Year Calibration Predictiveness Test Results - Paid

## Results – Paid

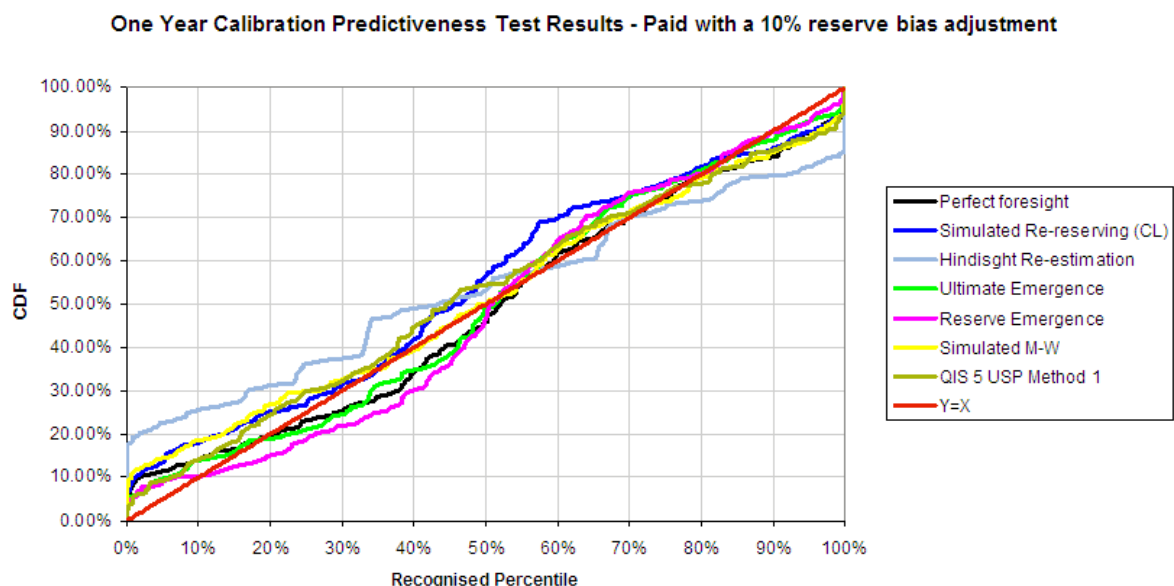| | Total Squared Error | χ² test statistic |
|---|---|---|
| Perfect foresight | 33% | 70 |
| Simulated Re-reserving (CL) | 66% | 83 |
| Hindsight re-estimation | 192% | 320 |
| *Ultimate emergence* | *27%* | *49* |
| Reserve emergence | 55% | 59 |
| Simulated MW | 45% | 99 |
| QIS 5 USP Method 1 | 34% | 70 |

Historical bias in the booked ultimates:

- Best estimate?

  o Have companies historically reserved as the mean of the distribution? If not it could bias the results

- There appears to be some evidence from our empirical tests that the reserve setting has been prudent

- To identify what impact this may have had, we have removed a 10% prudency margin from the booked reserves, and re-run the tests

## Results – Incurred with a 10% reserve bias adjustment



One Year Calibration Predictiveness Test Results - Incurred with a 10% reserve bias adjustment

## Results – Paid with a 10% reserve bias adjustment
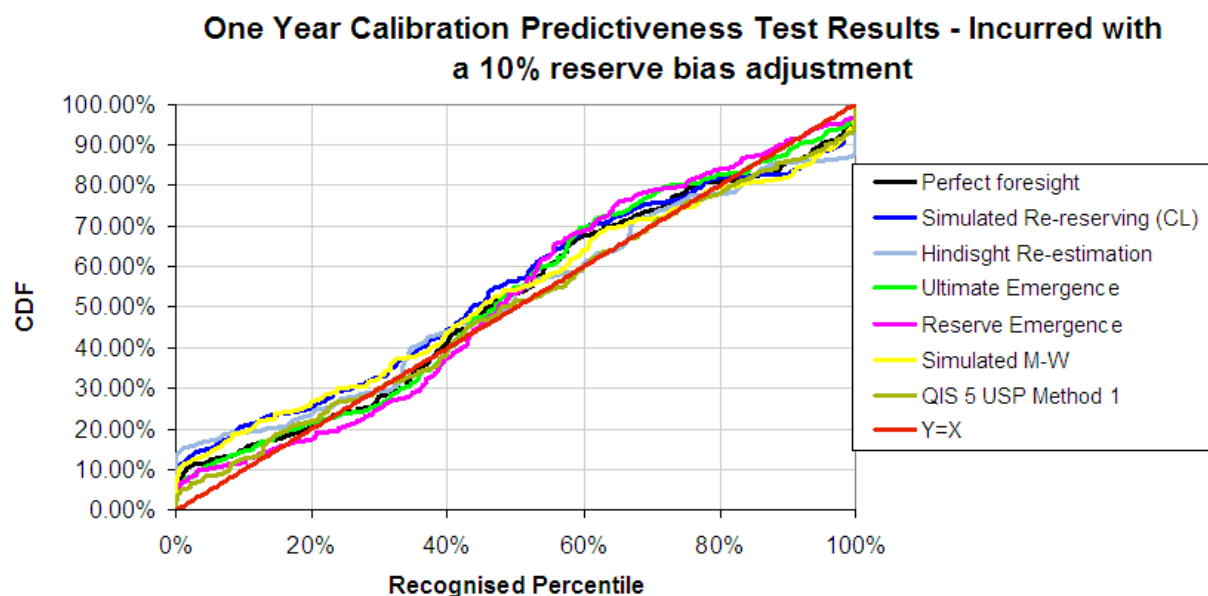


One Year Calibration Predictiveness Test Results - Paid with a 10% reserve bias adjustment

## Results –Incurred with a 10% reserve bias adjustment

| | Total Squared Error | χ² test statistic |
|---|---|---|
| Perfect foresight | 39% | 87 |
| Simulated Re-reserving (CL) | 93% | 123 |
| Hindsight re-estimation | 66% | 207 |
| Ultimate emergence | 45% | 70 |
| Reserve emergence | 55% | 73 |
| Simulated MW | 66% | 110 |
| *QIS 5 USP Method 1* | *15%* | *46* |

## Results – paid with a 10% reserve bias adjustment

| | Total Squared Error | χ² test statistic |
|---|---|---|
| Perfect foresight | 34% | 74 |
| Simulated Re-reserving (CL) | 69% | 81 |
| Hindsight re-estimation | 194% | 327 |
| *Ultimate emergence* | *27%* | *54* |
| Reserve emergence | 56% | 64 |
| Simulated MW | 46% | 100 |
| QIS 5 USP Method 1 | 32% | 58 |