

Model Validation Best Practices

June 2024



Jeremy Levitt, FSA, MAAA

David Schraub, FSA, CERA, MAAA, AQ

Vincent Chen, ASA

Background

Inaccurate or incomplete models can impair decision-making by key stakeholders, lead to significant losses by the organizations who use them¹, and ultimately erode consumer confidence in financial institutions. Regulatory changes over the last decade have also led to an increased need for actuarial models to be validated². Compounding all of this, actuarial transformation efforts including model conversions, mergers and acquisitions activity, and model upgrades trigger the need for validation.

Hence, ensuring the accuracy, comprehensiveness, and actuarial soundness of models should be a primary priority for institutions who rely upon their output. This is also consistent with numerous industry guidelines, regulations and standards of practice, such as ASOP 56. In this whitepaper, we set out best practice model validation techniques for actuaries, and address how AI can be used to support these techniques.

Model validation

Model validation³ forms part of an organization's model risk management framework and is the independent challenge and thorough review of an actuarial model for accuracy, completeness, compliance, ease of use, theoretical accuracy and goodness of fit.

An organization's model governance framework⁴ will typically define the manner in which model validation should be performed. A best-practice validation program will clearly specify the validation scope and the series of steps applicable to the model under review. Unless exclusions are explicitly spelt out, all aspects of the actuarial model should be validated including its accompanying tools, topside adjustments, and documentation.

Model validation findings should also be clearly categorized, with clear distinctions for unintentional flaws, simplifications, lack of conceptual soundness, methodology misinterpretations, and maintenance and documentation shortcomings. Certain companies allow the validation team to propose recommendations for remediation of findings.

In this whitepaper, we set out numerous best-practice validation procedures. All validation procedures highlighted in this whitepaper could be applied. However, the actuary should

¹There are numerous examples of this. In 1998, Long-Term Capital Management was liquidated due to losses arising in part from reliance on inaccurate correlation assumptions used in trading models. A decade later, the result of insufficient extreme testing analysis applied to mortgage backed security models led in part to the 2008 financial crisis. A few years later, JPMorgan Chase adopted an inaccurate mathematical model that led to a nearly \$6 billion loss.

² For example, in 2015 the NAIC Risk Management and Own Risk and Solvency Assessment Model Act went into effect, requiring large and medium-sized US insurers to prepare and file an ORSA report with their state regulator as part of their broader ERM framework. The ORSA summary report must also contain a general description of the insurer's model validation process. Other regulations, such as IFRS17 or LDTI, have led to significant changes to actuarial models, which in turn increase the need to validate those changes. Additionally, the NAIC issued the Model Audit Rule effective for financial years post January 1, 2010, which requires additional certifications, assertions on internal controls, and audit committee requirements.

³ Also termed model certification, affirmation, or review.

⁴ Model governance is not covered in detail in this whitepaper, but it should be noted that the governance program will dictate how model validation procedures are carried out, including determining the order in which models are validated, frequency of validations, the risk rating assigned to each model, peer review of validation procedures, maintenance of the model and model stewardship.

exercise judgment on the combination of validation procedures that should be applied, since this should depend on the nature, size and context of the validation.

Input review best practices

Best-in-class model validation procedures will place significant emphasis on input validation, particularly ensuring these are accurate, compliant and fit for purpose⁵. The nature of inputs into an actuarial model will vary depending on the model use case and product type, but might include product or service features, valuation, economic or projection assumptions, and inforce or new business data.

Inputs should be reviewed against independent and updated company source data, such as signed-off assumptions memos, experience study conclusions, treaty schedules, admin extracts, product illustrations, credit score cards or rate grids. Inputs should also be reviewed against industry and regulatory source data, such as valuation manuals and incidence tables published by trusted institutions. The actuary needs to confirm that the difference for each cell, entry, or table populating the model nets out to 0 unless intended. Various data analytic and comparative tools should be leveraged to execute on this exercise efficiently, subject to peer review and randomly assigned spot checks.

Inputs should also be reviewed for intuitive reasonableness. Identification of outliers using scatterplots and comparisons against prior period inputs are examples of basic ways in order to assess reasonableness. Reverse stress testing, which back-solves for the inputs needed to reach a pre-defined outcome, provides valuable insight to the model validator who will be able to assess the reasonableness of the back-solved inputs. Special attention should be paid to input data that has changed since the prior period due to a higher probability of unintentional error. Existing input review procedures performed by the first line of defense should be peer-reviewed by the validator for completeness.

The choice of input review method should be appropriate within the context of the input type. For example, the change in model point file population since the prior period should be split by driver (lapses, deaths, new business etc.) each of which should then be reviewed for reasonableness. Experience studies used to inform mortality assumptions should be reviewed to ensure these are based on credible data or are sufficiently blended, claims lags are allowed for, and mortality rates allow for mortality improvement where appropriate. Data repository facilities that store the inputs should be consistent across models and time, a central source of truth drawn upon by all intended downstream users, and securely stored. Compression methods, plan code mapping, data transformation, and extraction processes should all be carefully scrutinized to ensure reliance can be placed on the inputs flowing into the actuarial models.

⁵ This is consistent with [ASOP 56](#) Section 3.1.3, which states that the model validation exercise needs to be consistent with the intended purpose of the model

The actuary should exercise judgment to ensure special consideration is given to unique cases or model input types. Stochastic input data could be validated by performing martingale tests and ensuring the stochastic scenario generator output converges to the mean and standard deviation of the assumed distribution. Economic inputs should be double checked to confirm their start date aligns with the project or valuation start date. Compressed or grouped input data should be checked to ensure there has not been a loss of integrity due to grouping.

Use of AI to improve upon best practice procedures

Certain model validation tasks can be performed by leveraging GenAI tools. Documentation version comparison, ad-hoc queries, model comparisons and data visualization dashboard analyses are practical examples. Regardless of the tool used, the responsibility for the model validation remains with the model validation team.

Validation of AI or Machine Learning tools should follow the same model validation principles, with a specific emphasis on model transparency, discrimination testing and model drift. The scope of the model validation should be validated by the ethics committee, in addition to the risk management function. The model validation process should also leverage insight from the audit function, whether internal or external.

Model calculation and output validation best practices

Best-in-class model validation procedures will closely scrutinize model calculations and the resulting output extracts⁶, ensuring these are accurate, actuarially sound and consistent with regulatory or intended methodology.

Calculation Review

An independently built first-principles model is the gold standard for calculation validation. The independent model should be run on a representative sample of policies or policy groupings⁷, and the output from each policy or policy grouping compared against the model undergoing validation. If any of the policy or policy group-level outputs differs in excess of threshold⁸, the drivers need to be traced back to source and understood.

⁶ This is consistent with [ASOP 56](#) Section 3.6.2, which states that the model validation must confirm that the output is representative of what is intended to be modeled.

⁷ The representative sample should be determined such that the model points cover all material combinations of demographic characteristics, product features, and assumptions. Statistical sampling software, pivot tables, or a multivariate distribution which outputs the most material intersection of cells could be deployed to obtain the best possible sample. Spot checks using out-of-sample model points should be run through the first-principles model too. Care needs to be taken that the total number of model points in the sample is statistically significant.

⁸ Determining the optimal threshold will be based on the form of the underlying metric (e.g. cover amount vs. policy count), cell granularity, absolute size of the outputs being compared, use case of the model (e.g. statutory versus pricing), complexity of the comparative model used to review the target model, and governance policy.

A variation on the above approach is the use of a challenger model⁹, which could be used to compare results against the model undergoing validation. Using a challenger model, full regression testing can be performed, assessing differences at plan code, product, subsidiary or aggregate level. It should be noted that the threshold used to determine differences should take into account the granularity of the output being compared. An analysis of change procedure could be performed, where each change is successively updated in the challenger model until no major calculation or formulaic differences exist between the two models, and the impact assessed step by step for reasonableness.

A less sophisticated option would be to utilize a simplified model that captures the main features of the model undergoing validation, but can be developed with less resources and more efficiently. Here, the differences are expected to be larger than those from an independent model, and hence the threshold for comparison should take this into account.

Output Review

The actuary should review the stability of the output being produced by the model - across both time periods and varying input data. For example, a small change in expenses should not lead to a major discontinuity in PV Profit arising from a block of new business being modeled for pricing purposes. Stability of projected output should also be validated using dynamic validation procedures, which is described below.

Various forms of stress testing should be performed. Extreme value testing should be performed to confirm the behavior and stability of the output is reasonable - and the model does not malfunction - under extreme scenarios. Sensitivity testing should be performed to assess reasonableness of output following adjustments to individual assumptions feeding into the model. Scenario testing, which involves varying multiple assumptions at a time in order to create a potential future scenario, enables management to understand how the model performs under these particular scenarios.

Both static and dynamic validation procedures should be performed to ensure the output can be relied upon. Well executed static validation procedures ensure that all Time 0 balance sheet items¹⁰ input into the model tie out to the output of the model as well as to an independent source. Dynamic validation procedures will plot the trend in historical data against projections being produced by the actuarial model to ensure there are no major unexpected kinks in the smoothness of the trend lines. For example, the actuary should ensure that historical premiums trend smoothly into projected premiums, split by product line, plan code, subsidiary and/or in aggregate. An additional essential procedure is back-testing; here, the model should be run

⁹ A challenger model is a different model to the model undergoing validation, and may be contained within a separate software program, be a prior period model, or have a different use case that is subsequently manipulated for comparative purposes against the target model. This is a form of parallel testing. In general, a culture of comparative model validation is strongly encouraged. The model output should be compared against similar models to confirm reasonableness of output. Similarly, internal model results should be compared against each other to confirm consistency across different types of code, such as how the results change for different aggregated groupings.

¹⁰ These include account values, in-force count, cash values, total face amount, asset values against management accounting balance sheet.

using prior period data, and compared against actual historical outcomes, to assess how well the model performs retrospectively.

Targeted, manual checks should be layered onto the approaches set out above. Model code that is expected to have the most significant impact on results should be reviewed line by line¹¹. An experienced actuary should assess the model output by eye, ensuring that it looks intuitively correct. Random spot checks on output should be performed, for example checking that a cell reflective of higher mortality on term insurance leads to a larger reserve than another with lower mortality, all else being equal.

The actuary should also take account of the specific type of model and use case. Stochastic models should undergo stability testing to ensure that the model output is stable while varying the random seed or number of simulations. However, this may lead to a significant increase in runtime. Regulatory models will need to have a different validation approach to internal management models; the focus in the former is to ensure compliance, whereas the focus of the latter is to ensure congruence with internal assumptions and methodologic intent.

Conceptual soundness

The actuarial model needs to be assessed for theoretical accuracy, actuarial soundness, and alignment with intent. The calculations and implied probability distributions used to project cash flows need to be based on sound statistical theory. Timing of cash flows, such as reinvestment of cash flows, needs to be realistic. The method used by the model should be checked against methodology documentation, to ensure that the approach is sound. Additional value will be derived from obtaining an independent expert's opinion on conceptual soundness of the model, which is consistent with ASOP 56 Section 3.5

Model performance

Given the multitude of calculations demanded by the actuarial models by financial institutions today, efficiency has emerged as a significant concern. Addressing efficiency issues often involves enhancements in coding and optimally utilizing calculation servers. Although the calculation engine may vary depending on the software used, there are several best practices to confirm are implemented.

In particular, the model validator needs to ensure that unnecessary calculations are removed from the model. This represents the most direct approach to enhancing model efficiency. Given that the model often encompasses calculations for diverse reporting bases or requirements, not all calculations are pertinent to a particular reporting run. Thus, generating a calculation report elucidating the execution of calculations within the model is typically invaluable. This report enables the actuary to identify any extraneous calculations and subsequently tailor the coding or leverage software functionality to circumvent unnecessary computations.

¹¹ It is typically unrealistic to review all the code in an actuarial model, but code that is most impactful should be reviewed for accuracy, completeness, compliance and format. The code, and the formula prescribed by the code, should reflect intended regulatory or internal methodology and sound actuarial theory.

The validator should understand the computing flow and software functionality: Understanding the inner workings of the model calculation process within the machine is crucial for implementing efficient coding or configuration adjustments to optimize performance. Performance optimization typically hinges on factors such as data granularity, the frequency of calculations, and the utilization of CPU cores. Certain software may offer features to partition data, aggregate results across various dimensions, or distribute calculation tasks among multiple CPU cores. Armed with insights into the software's calculation engine and functionalities, tailored codes and configurations can be used to optimize performance.

The validator should also perform a benchmarking analysis. Insurance companies can conduct a thorough benchmarking analysis by comparing the performance of their models during regulatory reporting runs. Model performance is contingent upon factors such as the volume of data and the intricacies of insurance or banking product features. Through benchmarking, companies can gauge the scalability of their models and establish benchmarks to evaluate the level of efficiency deemed acceptable.

Aesthetics and documentation

Best practice model validation also takes the aesthetic and presentation of the model into account. There should be a clear and detailed model design document which outlines exactly how the model is designed and how future modifications should be designed. It is the validator's role to ensure this has been adequately implemented in the model. Here are examples of key features that a best practice model design should adhere to:

- Consistency: the model should be developed in a consistent manner, both within the model (e.g. different sections of code) and across other models that the organization deploys.
- Formatting: the labeling of the model, color scheme, presentation, and spacing of the code, should be well formatted and easy to follow. Model point files should have the correct format for ingestion by the model.
- Coding efficiency: the structural integrity of the model should be efficient, e.g. the same code should never be duplicated
- Model flow: the model should flow logically from the feed of inputs into the modeling infrastructure, all the way through the downstream population of reports or databases using model output. Sections should flow in a common sense manner with an intuitive user interface which easily allows the actuary to navigate through the model.

Additionally, it is essential for model documentation to be comprehensive, accessible and up to date. This will also enable the model to be replicated by a third party, minimize key person risk, facilitate better understanding of the model's intended use, limitations and results, and improve overall model transparency. The actuary performing the validation should ensure that there's comprehensive model documentation covering:

- Model user guide or implementation guide (including how to update/execute model runs, hardware and software environment, first line validation protocols, model change/access control, and system maintenance)
- Coding standards, which will typically apply to other models too
- Methodology, assumptions, and model purpose documentation, which should include references to external sources where applicable (e.g. regulation or academic sources)
- Model & process maps, illustrating where the models fits into the broader company framework, as well as how components within the model are interconnected
- Model caveats, biases and limitations, including previous validation findings and whether these were resolved

Documentation should be clearly maintained and updated, ensuring older versions are archived and complex models have greater detail than simpler models. Documentation needs to be tailored to the audience - for example, more detailed technical documentation will be needed for the model user guide to minimize uncertainty for the users of the model. Summarized documentation should be produced for items that are required by senior executives.

Conclusion

A well-executed model validation framework ensures that the output of actuarial models can be relied upon for decision-making purposes, improves key stakeholders' understanding of their business, and increases the overall integrity of the modeling infrastructure deployed at financial institutions. Having validated models also improves auditability of organizations' models and reduces the risk of material misstatements. This, in turn, will foster greater trust in financial institutions and reduce insolvency risks, fines, and reputational damage. Hence, model validation is an essential component of financial institutions' enterprise risk programs and the actuary's toolkit.

Disclaimer

The information provided in this whitepaper is for informational purposes only and does not constitute legal, financial, or professional advice. While we strive to ensure the accuracy and completeness of the information contained herein, we make no warranties or representations, express or implied, about the accuracy, reliability, suitability, or availability with respect to the whitepaper or the information, products, services, or related graphics contained in the whitepaper for any purpose. Any reliance you place on such information is therefore strictly at your own risk.