

FROM MANUAL OVERLAYS TO EVIDENCE-BASED RATING JUDGEMENT

How ICRA Analytics assessed whether rating overlays improved, weakened or preserved model performance



The Challenge

A financial-services institution needed an independent assessment of how effectively and accurately manual overlays were being used within its credit rating process. The central question was whether Risk Committee judgement enhanced the model's predictive and discriminatory performance, weakened it, or produced a neutral outcome when tested against actual performance.

Model Output VS Actual Performance

Backtest model-generated ratings against the identified default indicator to assess quantitative performance before overlays.

Overlaid Output VS Actual Performance

Backtest final risk committee-assigned ratings, including overlays, against the same identified default indicator.

Overlay Efficiency

Compare original and final backtesting results to determine whether overlays improved, deteriorated or preserved quantitative performance.

Granular Diagnostics

Evaluate effectiveness at the overall model-framework level, across models or segments, and at individual rating-grade level.

Our Solution

ICRA Analytics designed a comparative analytical framework that separately tested model ratings and final assigned ratings, and then drilled down into cases where an overlay had changed the model decision.



01 Data Structuring and Alignment

Prepared two performance-testing views: model-generated ratings with the identified default indicator, and final risk committee-assigned ratings with the same default indicator.



02 Model Rating Backtesting

Assessed original model ratings against actual performance to establish the model's quantitative performance before manual judgement.



03 Final Rating Backtesting

Assessed final ratings, including overlays, against actual performance to measure performance after risk committee intervention.



04 Overlay Population Identification

Isolated records where final ratings differed from model ratings to create a focused population of overridden cases.



05 Overlay Pattern Assessment

Analyzed overlay frequency, direction, concentration and magnitude, including upgrade versus downgrade rates and mean, median and maximum notch changes.



06 Model-Framework-Level Efficiency

Compared model-only and final-rating performance at the overall framework level and across relevant models or segments.



07 Rating-Grade-Level Efficiency

Reviewed where overlays were concentrated across rating grades and whether post-overlay outcomes demonstrated better, worse or neutral performance.



08 Overlay-Only Performance Comparison

For overridden cases, compared performance under the original model rating with performance under the final overlaid rating for the same records.



09 Validation Conclusion

Synthesized evidence on whether overlays enhanced or detracted from discriminatory and predictive power, while highlighting concentrations and patterns requiring review.

The Impact

The assessment provided a structured evidence base for evaluating the role of expert judgement in the rating process. The assessment strengthened the transparency and defensibility of expert judgement within the rating process by providing a structured, evidence-based framework for its evaluation. The analysis enabled informed decision-making, enhanced governance oversight, and increased confidence in the consistency and appropriateness of expert judgement applications across rating outcomes.

Objective View of Overlay Value

A like-for-like comparison of model-only and final ratings against actual performance enabled overlays to be classified as performance-enhancing, performance-deteriorating, or neutral.

Visibility Into Judgement Patterns

Frequency, direction, grade concentration and notch magnitude analysis made the nature and intensity of manual intervention transparent.

Multi-Level Effectiveness Testing

Results could be examined at overall framework, model or segment, rating-grade and overlay-case levels, avoiding reliance on a single aggregate conclusion.

Focused Management Review

Concentrated or recurring overlay patterns, and cases where final ratings did not improve performance, could be identified for further review.

Conclusion

ICRA Analytics' approach converted manual rating overlays from an opaque judgement layer into a testable component of the model framework. By backtesting model-generated ratings and final assigned ratings against the same actual-performance indicator, and by separately examining overridden cases, the analysis was designed to determine whether overlays added measurable value and where their use warranted closer management attention.

Core Analytical Lens

Model ratings vs actual performance | Final ratings including overlays vs actual performance
| Overlay cases before vs after intervention | Framework-level and rating-grade-level assessment