Automated Valuation
Model Back Testing

PropTrack

Automated Valuation Model Back Testing:

Why is back testing important

Back testing your AVM service is an essential practice for lenders, here's why:

Risk: It provides a quantitative understanding of AVMs accuracy and reliability, assisting lenders to **understand and quantify the risk** they are exposed to by using an AVMs in their valuation strategy.

Benchmarking: It enables lenders to **consistently compare the accuracy and reliability of different suppliers**, helping them identify those that perform the best and introduce the lowest amount of risk into your lending.

Governance and oversight: to meet the **compliance and critical regulatory requirements** clearly set out in APS 220 and APG 223.

Given AVMs do not carry the counter party risk mitigants of full valuations, ie no Professional Indemnity Insurance, a bank is taking all the risk in their utilisation if an AVM estimate is materially unreliable.

AVM back testing is a must do so that lenders can confidently use AVMs within their defined risk appetite and adhere to their prudential obligations.

Regulatory requirements for banks with respect to back testing

The Australian Prudential Regulation Authority (APRA) standards provide clear guidance and requirements for an Authorised Deposit Taking Institution (ADI) with respect to AVM back testing.

The most critical piece that bank must adhere to are the requirements set out in APS 220, Credit Risk Management.

The standard provides the following requirements for lenders with respect to the testing and understanding of alternate valuation methods such as AVMs.

- o The prudential standard states that when an ADI leverages alternative valuation methods such as...automated valuation methods (i.e., AVMs or data) to produce security valuation **it must have appropriate processes which address the monitoring, validation and reporting of valuation data**.
- o It also details that ADIs must have a capability that analyses the strength and weaknesses of relevant alternate methods (i.e., AVM or data) and have details of any back testing of a statistical random sample of these alternate methods.

The importance of these statements is reaffirmed by their repetition within the Prudential guide APG 223, Residential Mortgage Lending.

APRA place significant importance on back testing to ensure ADI's are holding AVM providers to the highest standard, ensuring their **models are sufficiently reliable and performing as statistically intended**.

This is due to the importance of the valuation in assisting lender's manage credit risk and in ensuring adequate capital is held to protect the bank and financial system in the event of downturn scenarios.

Additionally, for ADI's, alternate valuation methods generally do not carry any form of counter party risk mitigation features such as those attached to physical on-site inspection. Thus, if an AVM is inaccurate the risk of relying on this outcome is fully born by the lender.

The back test process

Each party plays a crucial role in effectively back testing an AVM service.

ADI's provide a statistically random and representative sample of loan origination activity over a defined period along with their valuation dates. Usually, this is a subset of their loan book portfolio and comprised of contract of sale or full valuations. The valuation amounts are not disclosed to the AVM service provider.

AVM service providers provide AVM's for the sample properties for the requested valuation dates using the information available at the date of valuation. The results returned should accurately mirror the results that would have been received at that time.

ADI's review the AVM's against the actual valuations for the sample set of properties using a set of performance measures and benchmarks. This provides a view of how the AVM's perform relative to their risk appetite.

Back testing is performed periodically (for example, quarterly or half-yearly) by the ADI for continuous monitoring of the AVM service provider's AVM performance in changing market conditions and to adhere to regulatory requirements.

For the most accurate back testing analysis, an ADI should compare AVM's to real observed market transactions. However, full valuations are often used as benchmarks in the absence of sales data, as they offer a detailed and professional assessment of the property's value at that time.

Contract of Sale vs Full Valuation: A contract of sale is the best evidence of a property's true market value. It is real evidence of what a buyer is willing to pay for a property. A full valuation is a valuer's opinion of value and may contain bias or error. APRA references full valuations as a benchmark for accuracy in refinance scenarios and so it may be used in back testing scenarios in lieu of a contract of sale.

Key things to be assess as part of a back test

Return rate

Also known as coverage or hit rate, this is the percentage of properties for which a valuation has been returned. From a ADI's perspective, return rate can be further qualified as a usable or acceptable return rate by adding lender's acceptance criteria on accuracy and confidence.

Accuracy of the estimate

Accuracy is a measure of how close the AVM is to a known benchmark value, which is either a subsequent sale price or value provided by a professional valuer. PropTrack recommends using sale price as a benchmark value as it's a price that a buyer has agreed to pay for the property. Accuracy can be represented as Within5 (within 5% of the benchmark), Within10 (within 10% of the benchmark), Within15 (within 15% of the benchmark) and Within20 (within 20% of the benchmark).

Levels of over and under valuation

An ADI should check the distribution of AVMs over and under the benchmarks. This provides insights into vendor AVM's bias towards over and underestimating property value. An AVM with over-estimation bias will increase lender's risk when providing loan for the property as the lender may approve higher loan amount than property's real value. Underestimation of AVMs may cause an ADI to miss lending opportunities. ADI's should pay particular attention to AVMs that are over 20% of the benchmark as part of their prudential risk management.
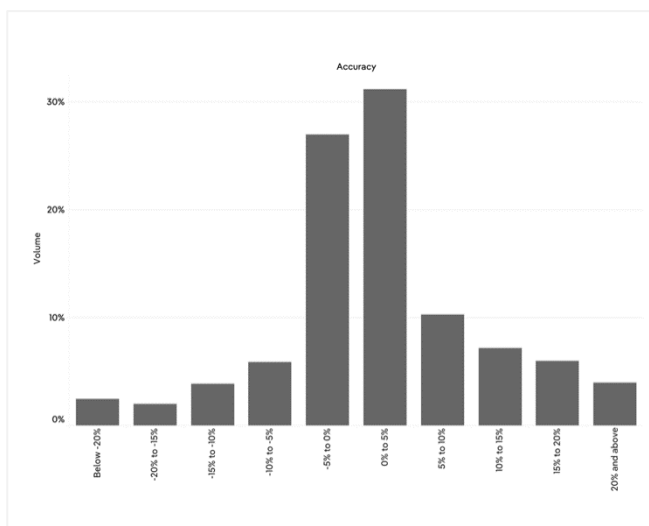

Confidence (FSD) and Statistical Integrity of this measure

Forecasted Standard Deviation (FSD) is a statistical measure of the model's precision, which is used to describe confidence in an AVM value. A low FSD implies high confidence in an AVM and vice versa. For lenders to have confidence to adopt AVMs, a best-in-class AVM should have FSDs that behave in line with the properties of a normal distribution.

FSD provides the highly probable value range around the valuation that the observed sale price will fall within. Approximately 68% of prices are expected to fall within 1 FSD of the valuation. Approximately 95% of prices are expected to fall within 2 FSDs of the valuation (as illustrated in the left diagram below).

Quoted FSDs need to be assessed to ensure they are performing as intended and that benchmark values fall within quoted FSDs 68% of the time (for 1 FSD) and 95% of the time (for 2 FSDs). The results of this test are illustrated below.
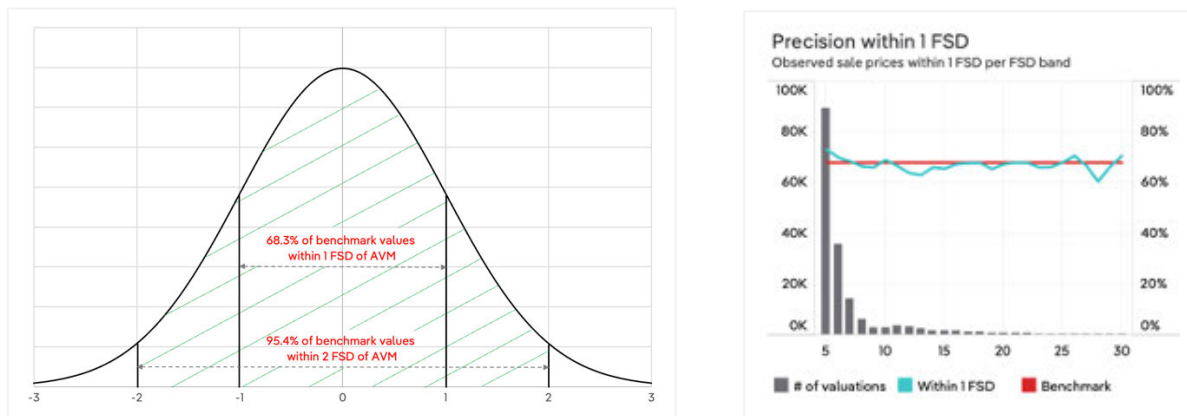
All FSDs quoted by an AVM provider should be as close to the 68% benchmark when assessed for performance. For example, for an FSD of 10 the final sale price is expected to fall within 10% of the AVM value 68% of the time.

If this occurs less than 68% of the time, then the FSD is too low and could open the lender to unnecessary overvaluation risk.

If this occurs more than 68% of the time, then the FSD is too high and represents a lost opportunity to digitize valuation flow.

Once lenders have confidence in the statistical accuracy of an FSD model, the FSD can help define the maximum acceptable FSDs for a given LVR given the current risk tolerance.

Performance by segments

AVM performance can be grouped by region, property type and value to provide visibility on where the AVM is performing well and identify areas that may need further analysis on the reliability of the AVM.

What should a lender do with their back testing results?

Utilizing back testing results is pivotal to the ongoing success of your lending operation. Here is how you can make the most of these insights:

- **Compare Performance Across Providers**: Analyze the results to identify how different providers perform in various market segments. Understanding these nuances allows you to strategically deploy each provider to maximize coverage and accuracy within your digital strategy. It's not just about finding the best provider, but the right provider for each specific task or region.
- **Update Rules/Policy Based on Changes in Performance:** Utilize insights from back testing to make necessary updates to rules or policies, including the ability to tighten or loosen restrictions on AVM's. By being responsive to the performance of

the AVM back testing can dynamically adjust your approach to extract maximum value, ensuring flexibility without sacrificing accuracy or compliance.

- **Collaborate with Service Provider:** Use the insights gained from back testing to foster collaboration with your chosen service provider. Together, you can improve the rigor in the back test and collaboratively improve the performance of the AVM service.

Monitor Key Performance Metrics: Back testing results are an ongoing tool for tracking key metrics, setting benchmarks, and evaluating the AVM's performance over time, ensuring you are on track to deliver your expected business outcomes.

By leveraging back testing results in these ways, you can turn raw data into actionable insights and strategy, enhancing both the agility and robustness of your lending practices.

How does PropTrack monitor performance?

PropTrack, as part of its quality assurance and model governance procedures, produces a monthly AVM Performance Report which shows the return rate, accuracy and FSD performance of the AVM. Results for each month are based on a 6-month rolling window of observed sales data. The observed sale prices are compared to the model estimates calculated immediately prior to the sale event and aggregated across a number of key segments. All AVM valuations used in the report are generated based on information known just prior to the sale date. This aligns to the regulatory standards on AVM performance reporting.

What lenders should look for with their AVM back test results

Results to use data available at the time of the valuation only

In relying on an AVM, lenders are taking risk based on the result of the model at the point in time an estimate would be provided within the live environment. Back testing data used to test models should replicate this process, with AVM service providers always providing results that are a true representative of how the service will perform in a live environment. This is typically known as as-at AVM results.

For example, if an AVM service provider is experiencing system data lag in receiving sales information, that data must not be used in the AVM estimation. There are cases where the sales may have occurred before the valuation date, but received after the valuation date and they must be excluded from the model. If they are included, then the AVM accuracy will be artificially higher than it really is.

One example of how AVM providers may limit the impact of data lag is by leveraging agent advised sales. This is a way for providers with a strong connection to listings data to reduce lag and incorporate this data ahead of the lag created by waiting for VG data.

Don't always rely on the AVM with the lowest FSD

FSD is critical for lender to understand the risk they are taking when relying on an AVM. Lenders must independently test that the FSD provided are performing as statistically intended (i.e. 68% of the time the true value falls within +/-1 FSD) at each FSD.

A lower FSD does not mean a better AVM and relying on an AVM that produces the lowest FSD can expose you to significant valuation risk depending on your risk appetite.

AVM values that are equal to the benchmark

There should be almost no AVM values that are equal to the benchmark. This could indicate the sales data received after the valuation date might be used in the AVM estimate. Check if the future sale price is being used.

Performance against Contract of Sale valuations is materially higher than against full valuations

This may be an indicator that sales information after the valuation date is being used in the AVM estimate. Check for any bias towards owners estimates in your valuations data.

Time taken for AVM service provider to return results

AVM back test results should be system driven and fully automated.

1. This involves an integrated address matching service along with an automated append of AVMs to each address and valuation date.
2. An agreed SLA should allow for any operational process to complete to ensure that the results are error free.
3. If the time to return the result is long there is an increased risk that the file is being manually reviewed by the AVM service provider and post processed manually.

About PropTrack

PropTrack is Australia's most reliable choice for property data, market insights and valuations solutions – utilising the latest artificial intelligence, machine learning techniques and software capabilities to power REA Group apps, like realestate.com.au, and our customers. As part of our commitment to changing the way the world experiences property, PropTrack is helping organisations use transformative property data, develop new property experiences and automate valuations, enabling new solutions and services for the modern consumer. Join the millions of Australians who already use REA Group apps every month and transform your property experience. **proptrack.com.au**