2012 Appraisal Institute Annual Meeting
Aug. 1–3
Loews Coronado Bay Hotel
San Diego, CA
appraisalinstitute.org/aiconnect
AVM Testing and Model Validation

Mike Coyne
AVMetrics
Collateral Process Manager
Statistical Analysis and Metrics

• What statistical analytics are being used?
  • How do these statistical tests relate to model performance?
• What standards of measurement are being used?
  • Are they adequate?
• Confidence Scoring Methods
  • Why so many?
  • Is there a correlation to predicted value?
  • Standardizing the confidence score
AVM Advantages and Disadvantages

**Advantage**
- Speed, fast
- Objective, non-biased
- Inexpensive, cost-effective
- Consistent

**Disadvantage**
- Cannot confirm or deny a property exists
- Limited ability to address property condition (equal)
- Limited ability to account for external influences
- Limited coverage in area with data limitations
The following graphs present examples of presentation styles for each metric.

AVMetrics has successfully used these presentation formats for prior model validation projects from the National level down to the county and sub-county levels.
**Histograms**

*Histograms* are one of the most flexible tools for evaluating model's performance though more so for its ability to isolate on one model rather than for comparative purposes.

The bins can be created to key on specific variance groupings or segmented to present a traditional “bell curve” view (as seen here). Introduction of complementary metrics provides a clear synopsis of a model’s performance.

---

**Overall Deviation Distribution (%) - M12**

- ME: 0.06%
- MAE: 11.5%
- SD: 17.8%
- HR: 75.6%

**Median Error:** -0.02%

**Mean Error:** 0.06%

(AVM Value - Benchmark) / Benchmark
Standard Deviation

Vendor Standard Deviation - All Models
Mean Absolute Error

Vendor Mean Absolute Error - All Models

Absolute Average (LAVM Value - Benchmark) / Benchmark
Mean Error

Error rates, commonly referred to as deviation or variance can be presented in two ways; the arithmetic and absolute averages.
State Level Metrics

As the data is segmented for regional level results, model strengths and weaknesses begin to appear.
County Metrics

County level results are backbone of model validation results. Scoring is performed at the county level and Model Preference Tables are generally applied at this level as well.
Mixed metric charts provide unique views of related metric results. The chart below is a standard presentation of the relationship between effective hit rate and mean absolute error.
State Mix Metric

Refinement in mix metric charts show a similar process of increased definition to the results.
County Mix Metric

At the county level the refinement results in three equal groups containing four models each.
<table>
<thead>
<tr>
<th>County</th>
<th>Score</th>
<th>Blind Name</th>
<th>Records</th>
<th>ME</th>
<th>5%AR</th>
<th>10%AR</th>
<th>20%AR</th>
<th>Over 20%AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Los Angeles</td>
<td>9.71</td>
<td>Model 21</td>
<td>2,080</td>
<td>-1.2%</td>
<td>53.1%</td>
<td>81.3%</td>
<td>93.7%</td>
<td>6.3%</td>
</tr>
<tr>
<td></td>
<td>9.11</td>
<td>Model 6</td>
<td>5,429</td>
<td>2.0%</td>
<td>53.4%</td>
<td>72.0%</td>
<td>86.7%</td>
<td>13.3%</td>
</tr>
<tr>
<td></td>
<td>7.68</td>
<td>Model 15</td>
<td>3,669</td>
<td>3.5%</td>
<td>23.8%</td>
<td>47.2%</td>
<td>75.6%</td>
<td>24.4%</td>
</tr>
<tr>
<td>Price Tier</td>
<td>Score</td>
<td>Blind Name</td>
<td>Records</td>
<td>HR %</td>
<td>ME</td>
<td>5%AR</td>
<td>10%AR</td>
<td>20%AR</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>------------</td>
<td>---------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>&lt;100K</td>
<td>9.71</td>
<td>Model 21</td>
<td>50</td>
<td>62.0%</td>
<td>13.6%</td>
<td>32.3%</td>
<td>58.1%</td>
<td>74.2%</td>
</tr>
<tr>
<td></td>
<td>9.11</td>
<td>Model 6</td>
<td>123</td>
<td>73.2%</td>
<td>10.2%</td>
<td>35.6%</td>
<td>52.2%</td>
<td>72.2%</td>
</tr>
<tr>
<td></td>
<td>7.68</td>
<td>Model 15</td>
<td>76</td>
<td>59.2%</td>
<td>45.4%</td>
<td>2.2%</td>
<td>2.2%</td>
<td>4.4%</td>
</tr>
<tr>
<td>100K-300K</td>
<td></td>
<td>Model 15</td>
<td>1417</td>
<td>76.2%</td>
<td>11.3%</td>
<td>20.2%</td>
<td>40.1%</td>
<td>67.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 21</td>
<td>768</td>
<td>70.4%</td>
<td>-0.1%</td>
<td>43.6%</td>
<td>73.8%</td>
<td>90.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 6</td>
<td>2075</td>
<td>85.7%</td>
<td>3.7%</td>
<td>48.4%</td>
<td>66.6%</td>
<td>83.7%</td>
</tr>
<tr>
<td>300K-500K</td>
<td></td>
<td>Model 15</td>
<td>1230</td>
<td>80.4%</td>
<td>-1.9%</td>
<td>30.1%</td>
<td>58.2%</td>
<td>86.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 21</td>
<td>716</td>
<td>76.8%</td>
<td>-2.4%</td>
<td>56.0%</td>
<td>85.8%</td>
<td>96.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 6</td>
<td>1835</td>
<td>85.4%</td>
<td>0.7%</td>
<td>60.7%</td>
<td>79.3%</td>
<td>90.1%</td>
</tr>
<tr>
<td>500K-700K</td>
<td></td>
<td>Model 15</td>
<td>466</td>
<td>80.5%</td>
<td>-3.4%</td>
<td>23.5%</td>
<td>49.9%</td>
<td>84.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 21</td>
<td>256</td>
<td>76.2%</td>
<td>-2.4%</td>
<td>65.1%</td>
<td>88.2%</td>
<td>97.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 6</td>
<td>702</td>
<td>83.0%</td>
<td>1.0%</td>
<td>58.7%</td>
<td>77.5%</td>
<td>90.6%</td>
</tr>
<tr>
<td>700K-900K</td>
<td></td>
<td>Model 15</td>
<td>250</td>
<td>79.6%</td>
<td>-1.4%</td>
<td>23.1%</td>
<td>46.2%</td>
<td>74.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 21</td>
<td>145</td>
<td>85.5%</td>
<td>-2.2%</td>
<td>65.3%</td>
<td>89.5%</td>
<td>96.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 6</td>
<td>345</td>
<td>84.1%</td>
<td>-0.5%</td>
<td>49.7%</td>
<td>67.9%</td>
<td>84.8%</td>
</tr>
<tr>
<td>900K-1.1MIL</td>
<td></td>
<td>Model 15</td>
<td>106</td>
<td>82.1%</td>
<td>0.5%</td>
<td>25.3%</td>
<td>41.4%</td>
<td>72.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 21</td>
<td>73</td>
<td>83.6%</td>
<td>-1.6%</td>
<td>55.7%</td>
<td>82.0%</td>
<td>96.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 6</td>
<td>159</td>
<td>81.1%</td>
<td>3.4%</td>
<td>45.0%</td>
<td>69.0%</td>
<td>86.0%</td>
</tr>
<tr>
<td>&gt;1.1MIL</td>
<td></td>
<td>Model 15</td>
<td>124</td>
<td>80.6%</td>
<td>-7.9%</td>
<td>12.0%</td>
<td>31.0%</td>
<td>60.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 21</td>
<td>82</td>
<td>85.4%</td>
<td>-1.0%</td>
<td>54.3%</td>
<td>80.0%</td>
<td>91.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Model 6</td>
<td>190</td>
<td>77.9%</td>
<td>-1.4%</td>
<td>41.2%</td>
<td>60.1%</td>
<td>82.4%</td>
</tr>
</tbody>
</table>
**Property Type Metrics**

*Property Type* charts can be generated at a model level or as a county sub-group allowing further insight and refinement of model selection and best practice rules.
Confidence Score Correlation

Understanding how confidence scores correlate to a model’s value provides two critical elements in any model validation process.

Los Angeles FSD Score Mean Error Model 21

State: CA
County: Los Angeles
Model: PASS

ME

CS Group
Confidence Score Correlation

Understanding how confidence scores correlate to a model's value provides two critical elements in any model validation process.

Los Angeles FSD Score Mean Error Model 21

SD

40.0%
35.0%
30.0%
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%

4 - 7 8 - 11 12 - 15 16 - 19 20 - 24

State CA County Los Angeles Model PASS

SD

CS Group
Generic AVM Origination Cascade Process

Developed At County Level

AVM 1

Rule based decisions met:
- CLTV
- Loan Type
- FICO

AVM 2

AVM 3

Fail · Value Returned · Meets Confidence Score Threshold

Fund Loan

Increase Service Level and Proceed through normal Bump Logic

Loan Request Only
- 2nd Equity

Property Eligibility NO
- Manufactured Housing
- 2-4 Units
- Nonconforming
Reporting the Results of Testing

• What reporting format best meets the users needs?

• What minimum elements should the reporting structure contain?

• Purpose of specific reporting formats

• Reporting commonalities

• Internal vs. external reports
Translating Test Results into Business Decisions

• How can the testing results fuel risk based decision engines?

• What are the key metrics that decision makers need to key on?

• What are some bad decisions that can be made with good data?

• Key elements

• Risk decisions
Best Practices

• Policy and Procedures
  – Define the Key elements of P&P
  – P&P audit trails

• Independence of Testing
  – In-house verses outsourced
  – Transparency issues

• Cascading or Waterfall Logic Systems
  – Valuing shopping potential
  – Deferring to Product Aggregators
Appendix B-(Interagency Guidelines)

• An institution should establish policies and procedures that provide a sound process for using various methods or tools. Such policies and procedures should:
  – “Ensure staff has the *requisite expertise* and training to manage the selection, use, and validation of an analytical method or technological tool. If an institution does not have the in-house expertise relative to a particular method or tool, then an institution should employ additional personnel or engage a third party”.
  – Address the selection, use, and validation of the valuation method or tool.
Appendix B - (Interagency Guidelines)

With the latest guidance the Federal Regulators recognize that AVM’s can be a valuable tool, HOWEVER....

• **Professional driver/closed course, do not attempt on your own**

  – You Need to have some knowledge and expertise in the following areas
  – AVM testing and Model Validation
  – Evaluating the testing to produce good decisions
  – Selecting the appropriate model for your risk decisions
  – Setting policies around how you will deploy (or not) AVM’s and when

If you do not have this “requisite expertise”, Then you need to seek it by hiring, Training or from outside third parties.
AVM Use  (When and Where)

- You can still use AVMs as a (1) supplement to an appraisal or (2) in situations in which the Guidelines recognize an appraisal to be impractical or (3) where an appraisal is not required (i.e. portfolio monitoring).
  - AVMs are given substantial discussion in the Interagency Guidelines - VALIDATING their use
  - Basic AVM Guidelines did not change; AVMs can be used for some originations when used in conjunction with additional tools
  - An AVM by itself is not solely sufficient if it does not address property condition
  - AVMs must comply with “safe & sound” banking practices
  - AVMs are often recommended as a best practice (e.g. Freddie Mac Bulletin 2009-18)
Let’s talk for a minute about AVM’s in the origination space and why they should not present any additional collateral risk over the traditional appraisal process…. I call this prudent use of AVM’s the “Best Practices” BOX.... This box is designed by the guidelines (and best practices). This box defines the area in which AVM’s are allowed to operate and it is fairly small in comparison to the traditional appraisal products. Let’s go though the basic qualification criteria that makes up the walls of our AVM box...

- The main qualifier is the deminimus level (Currently set at $250,000 or below)
- Higher qualifying credit score (Than would be required for “valuations)
- Loan programs restrictions (lower LTV, portfolio, GSE PIW Programs)
- Market condition restrictions (No adverse conditions, flood, hurricane, fraud, etc....)
- Property type restrictions (No manufactured housing, 2-4 units, high rise Condo’s, etc..)
- AVM acceptance thresholds
  - Confidence Score (or FSD)
  - Valuation Floor (below which and AVM will not be accurate enough)
  - Valuation Ceiling (above which and AVM will not be accurate enough)
  - Acceptable Value Variance Range (from a stated value)

With a combination of these parameters your custom AVM BOX will contain properties that are very conforming and with negative property physical / market condition factors that are substantially minimized.
New rules to use AVMs?

• When do I need to verify actual (current) physical condition?
  – No set circumstances

• How do I verify condition?
  – Appraisal
  – Home Inspection
  – BPO
  – Condition Report

• What are acceptable proxies (i.e. “research”) for condition inspections?
  – Recent appraisal or BPO report?
  – How far back in time are on-site valuations considered “current”?
  – “Recent” record of property on MLS with subject property photos?
Lender Responsibilities for AVM Use

- Lenders are ultimately responsible for all third-party arrangements and should maintain the expertise and resources for proper oversight.

- Valuation providers who are stepping up compliance and/or maintaining high standards of QC should serve to enhance the valuation process for a lender partner – not replace it.

- Interagency Guidelines are very clear that lenders are responsible for any valuation outsourcing – this includes AMCs, BPOs, AVM, consultants, third-party reviewers and others.

- The lender still holds ultimate responsibility with the regulator to make sure proper controls and independence are in place.
Guidance on Model Risk Management

- Section V Model Validation:
  - "rigor and sophistication of validation should be commensurate with the bank’s overall use of models”
  - "generally, validation should be done by people who are not responsible for development or use and do not have a stake in whether a model is determined to be valid”
  - Individuals doing validation should have a significant degree of familiarity with the line of business using the model and the model’s intended use.

- 3 Key Elements of Comprehensive Validation
  1. Evaluation of conceptual soundness, including developmental evidence
  2. Ongoing monitoring, including process verification and benchmarking
  3. Outcomes Analysis, including back-testing