C6: Data Quality and the Impact on Catastrophe Modeling

David Lalonde FCAS, FCIA, MAAA

CAS Spring Meeting

June 20, 2007

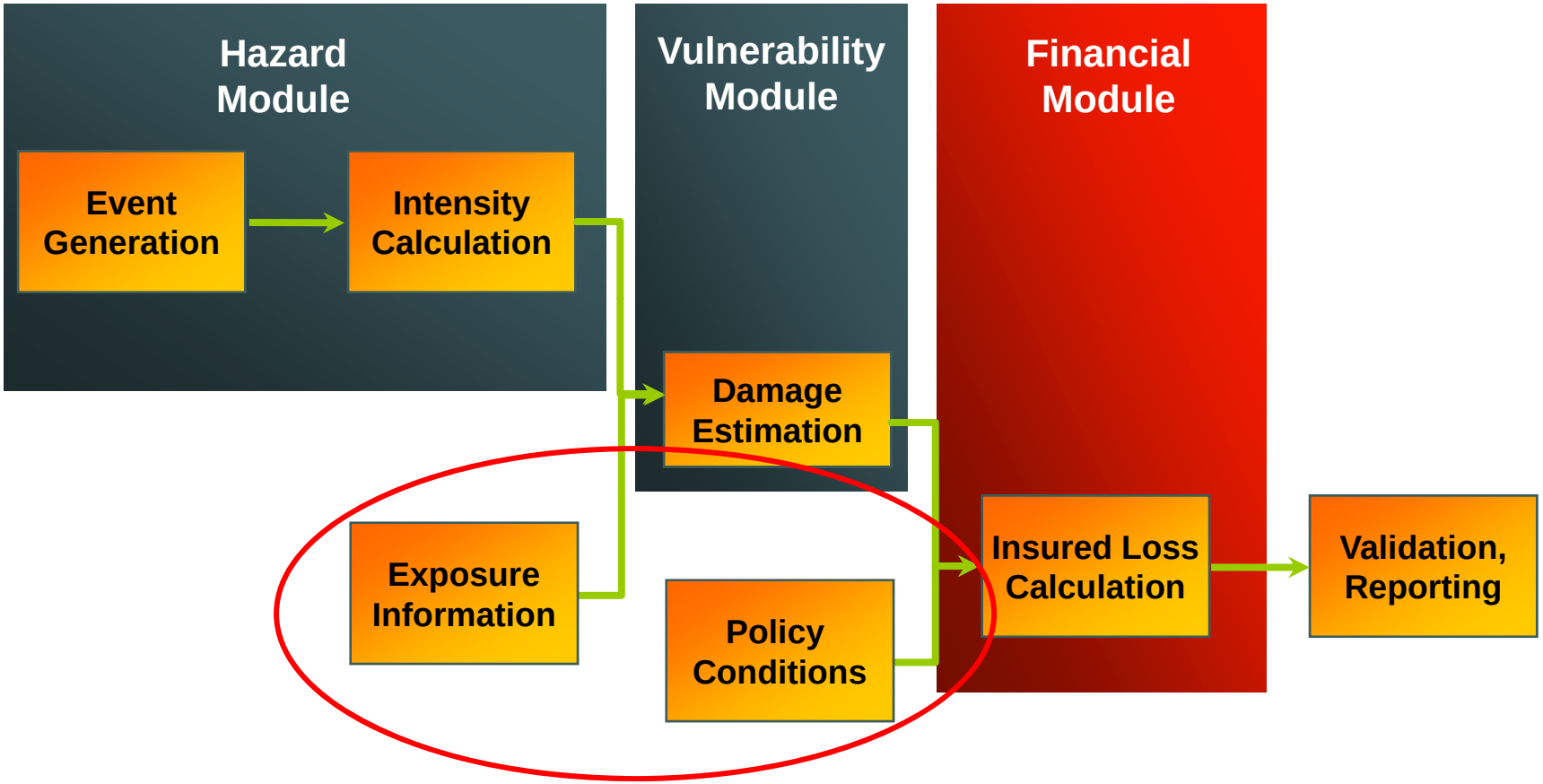
Lake Buena Vista, FL

BETTER TECHNOLOGY
BETTER DATA
BETTER DECISIONS

AIR

www.air-worldwide.com

Primary Model Components



Important Factors Contributing to Accuracy in Modeling

- Location information
  - Level of resolution
  - Accuracy of data
- Replacement values
  - Damage is calculated based on replacement values
  - Underestimating replacement values leads to underestimated losses
- Physical characteristics
  - Vulnerability is dependent on construction and occupancy characteristics
  - Results can be refined using secondary risk characteristics

Significant Exposure Data Quality Issues Discovered in Post-Katrina Analysis

- AIR reviewed exposure data from companies representing more than 50 percent of the total U.S. property market
- Nine out of ten commercial properties analyzed had replacement values less than the amount estimated using a standard engineering-based cost estimation process.
- Over 50 percent of companies analyzed lacked construction and/or occupancy information for more than a third of policies.
- Accurate analysis of multiple-location policies requires an address for each location.
- The coverage limit should not be used as a proxy for the replacement value.

Data Required for Modeling - Hurricane

Location	Replacement Value	Building Characteristics	Secondary Characteristics	
Street Address	Building	Construction	Window Protection	Large Missile Source
Zip Code	Appurtenant Structures	Occupancy	Glass Type	Small Debris Source
	Contents	Year Built	Glass Percent	Roof Pitch
	Time Element	Height	Building Condition	Building Foundation Connection
			Roof Anchorage	Year Roof Built
			Roof Deck	Exterior Doors
			Roof Geometry	Roof Covering
			Roof Deck Attachment	Roof Attached Structures
			Roof Cover Attachment	Wall Attached Structures
			Wall Type	Wall Siding

If I Can Only Collect One Piece of Information for My Hurricane Book, Which One Should I Collect?

- Analyzed commercial book of business, selecting best and worst level of each characteristic

Characteristic	Range
Roof Anchorage	22%
Roof Geometry	21%
Window Protection (Shutters)	19%
Roof Deck Attachment	14%
Roof Covering	12%

It is important to understand how characteristics work together. For a given construction, simultaneously selecting the best vs worst level for each characteristic yields losses from 50% lower to 65% higher than the base loss for a total loss range of 115%.

Data Required for Modeling - Earthquake

Location	Replacement Value	Building Characteristics	Secondary Characteristics	
Street Address	Building	Construction	Foundation	Torsion
Zip Code	Appurtenant Structures	Occupancy	Proximity Exposure	Shape
	Contents	Year Built	Soft Story	Structural Irregularities
	Time Element	Height	Special EQ Resistant Systems	Building Foundation Connection
			Retrofit Measures	

If I Can Only Collect One Piece of Information for My Earthquake Book, Which One Should I Collect?

- Analyzed commercial book of business, selecting best and worst level of each characteristic

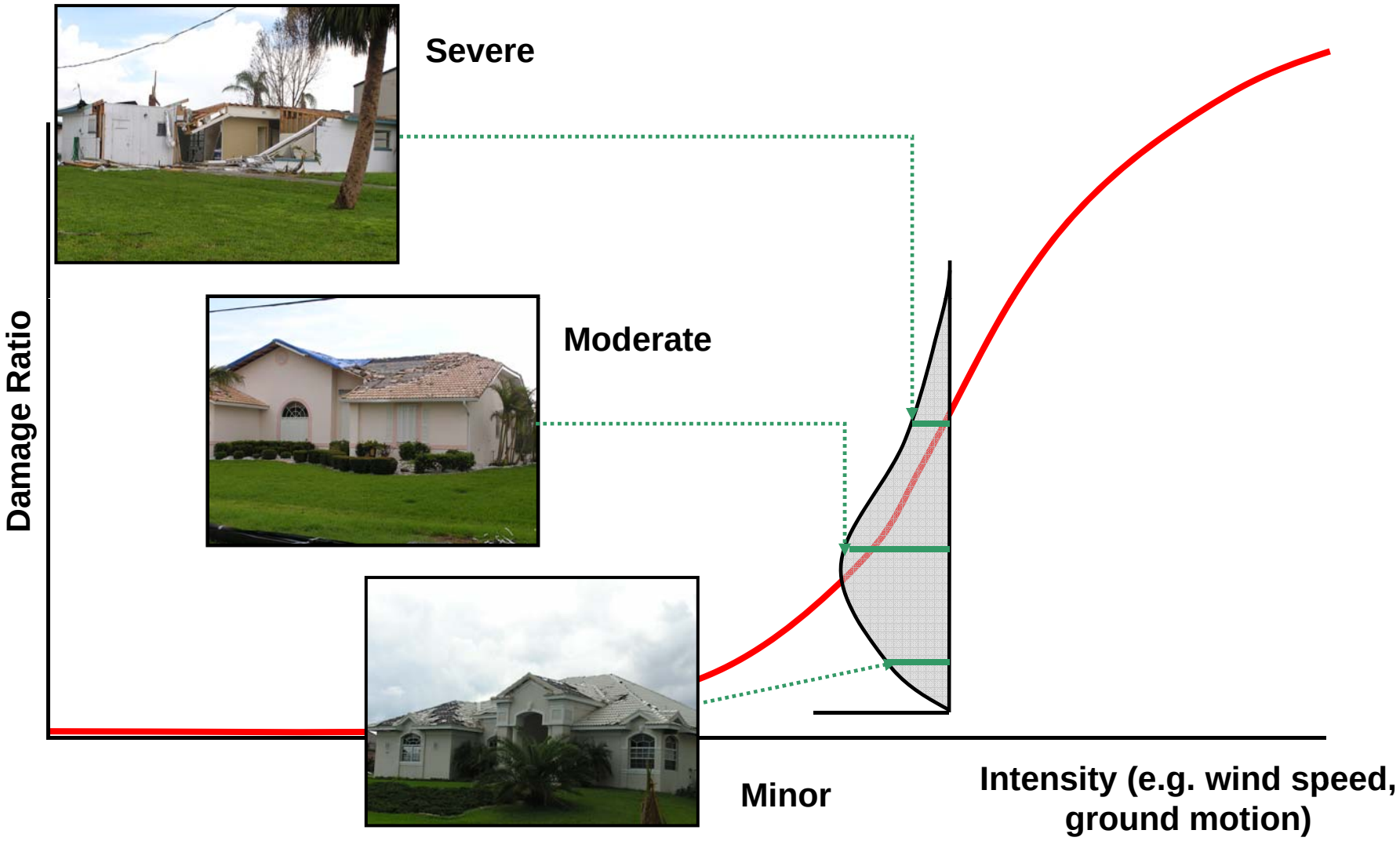
Characteristic	Range
Building Foundation	28%
Building Shape	25%
Torsion	9%
Building Foundation Connection	8%
Building Condition	7%

Insurance and Reinsurance Terms

Limits		Deductibles		Reinsurance*	
Location	*Policy*	*Location*	*Policy*	*Location*	*Policy*
Site Limits	Blanket Policy Limits	Combined ($ or %)	Attachment Point	Proportional Facultative	Proportional Facultative
Coverage Specific Limits	Excess Policy Limits	Combined Excluding Time ($ or %)	Blanket	Non-Proportional Facultative	Non-Proportional Facultative
Building	Blanket Policy Sub-Limits	Coverage Specific ($ or %)	Franchise	Surplus-Share	Surplus-Share
Appurtenant Structures	Excess Policy Sub-Limits	Building	Minimum / Maximum		
Contents		Appurtenant Structures	Percent of Loss		
Time Element		Contents			
		Time Element			
		Annual			
		CEA Mini-Policy			

*** Catastrophe Excess of Loss, Aggregate Excess of Loss, Quota Share, Surplus Share, and Per Risk Treaties are also available for application to the entire portfolio**

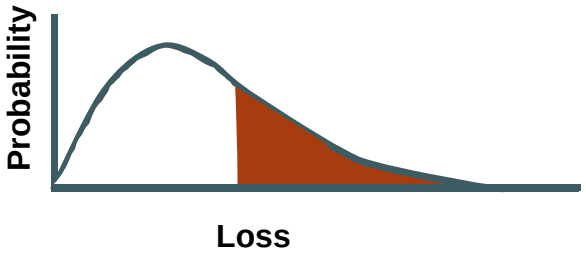
Variability in Damage for a Given Intensity



All Policy Conditions Must be Applied Probabilistically

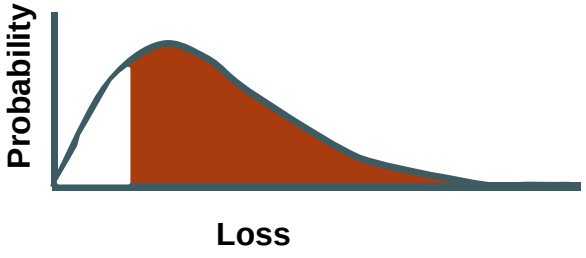
$250,000 Building Value

5% mean damage = $12,500



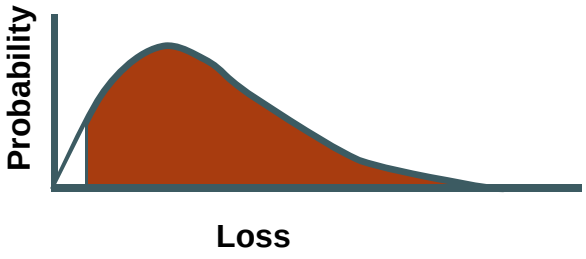
5% Deductible

Loss = $4,215



2% Deductible

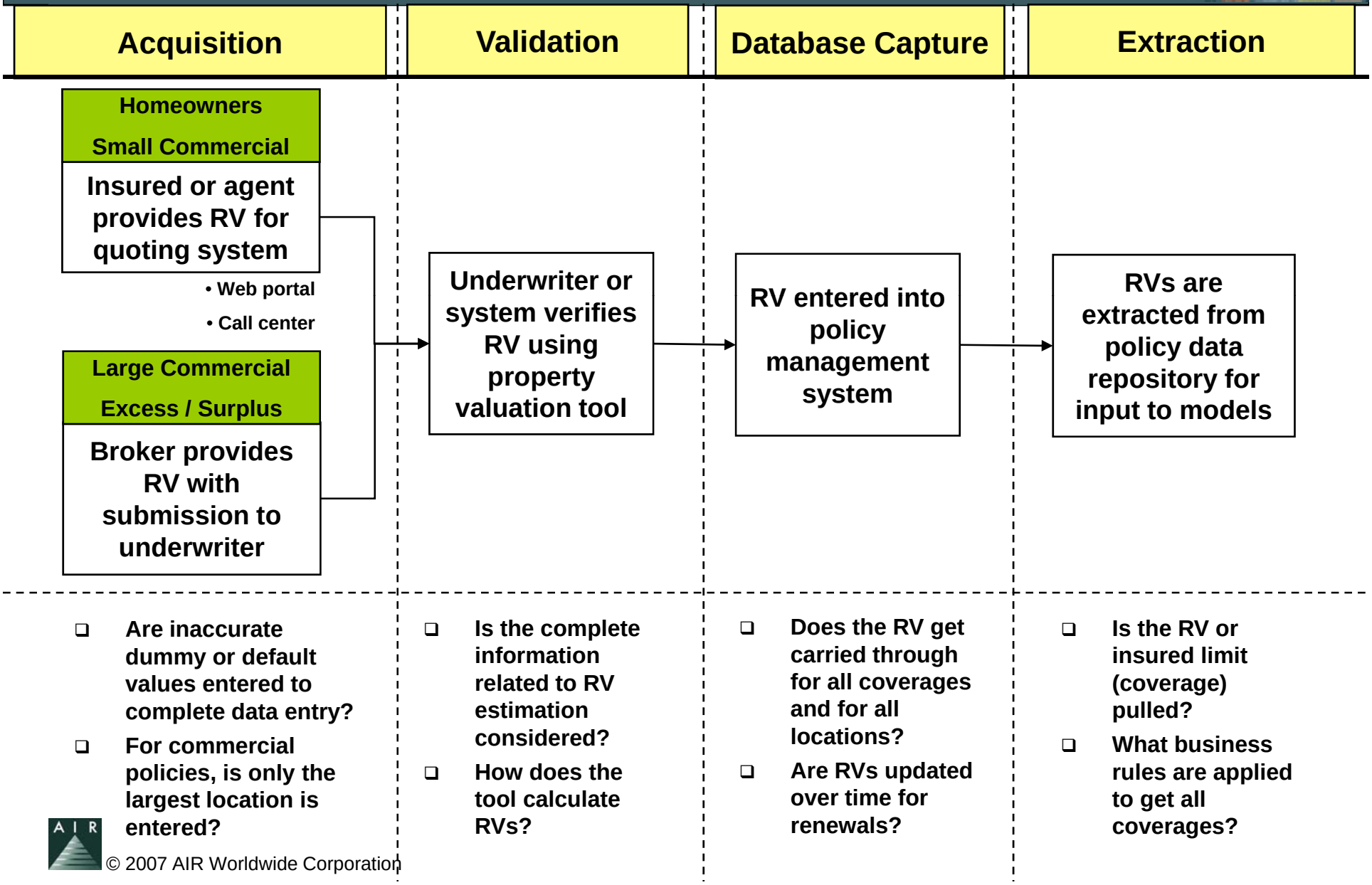
Loss = $7,750



1% Deductible

Loss = $9,750

Potential Sources of Data Quality Loss in Workflow

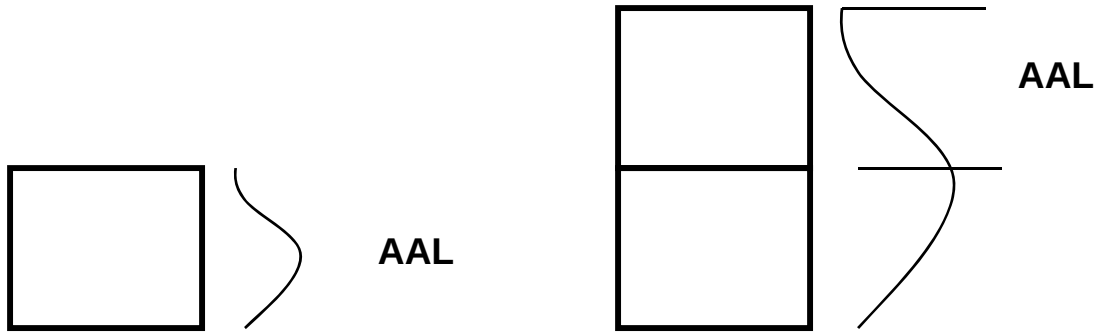


How Do I Determine if My Input Data is Reasonable?

- Residential exposure data tends to be homogeneous allowing checks by broad category
- Check coverage relationships
  - Ratios of coverage limits to replacement values
  - Are maximum and minimum limits within my underwriting guidelines?
  - What is ratio of appurtenant structure, contents and ALE to building limits
- Check Deductibles
  - What is my average deductible by state?
  - Any zero deductibles? Is this possible?
  - Any deductibles greater than limits?
  - Range of deductible percentages by total insured value
- Check for completeness
  - Percentage of the total exposure with no construction information
  - Percentage of total exposure with unmodelable location information (no street address, invalid zip code)
- General filtering capabilities and a variety of reports for reviewing data

Data validation is more difficult in commercial business due to the heterogeneity of exposures

- Building reported a replacement value of $5 million
- If building characteristics are unknown there may be some straight-forward ways to determine
  - Company website may have photo of building yielding construction and roof info
  - Tax assessor databases often contain square footage
- Cost per square foot is a common reasonability check
  - Using industry expertise and proprietary exposure sets, AIR experts estimate a unit cost per square foot for building
  - Multiplying that value by the recorded square footage for the building returned a replacement cost over $10 million
- Communication with the client revealed that AIR unit cost estimations were consistent with internal benchmarks and further investigation revealed this policy was a $5M excess of $5M layer; the insurer had entered the layer for which they were responsible as replacement value



Why is Location Information Important?

- Location information is key to the accuracy of model results
  - Models calculate damage at the latitude and longitude at which a risk is located
  - Difference between geocoding level can affect how a location is affected by an event, and particularly by the impact of storm surge on coastal locations
- Resolution of provided data is vital to quality of model output analysis
  - Street address – geocoded at exact location
  - Relaxed address – geocoded at a city block or street extension
  - ZIP code – geocoded at a population-weighted ZIP code centroid
  - City centroid – geocoded at a geographic city centroid
  - County centroid – geocoded at a county centroid

Impact of Location Information on the Quality of Loss Estimates

Full Portfolio

	Exact Address	ZIP Code Centroid	% Difference
AAL	$1,698	$ 1,711	0.77%
STD	$4,918	$ 4,972	1.09%
1.0% EP	$25,325	$25,388	0.25%
0.5% EP	$33,119	$33,009	-0.33%
1.0% TVAR	$40,425	$40,852	1.06%
0.5% TVAR	$50,317	$50,914	1.19%

Collier County, FL

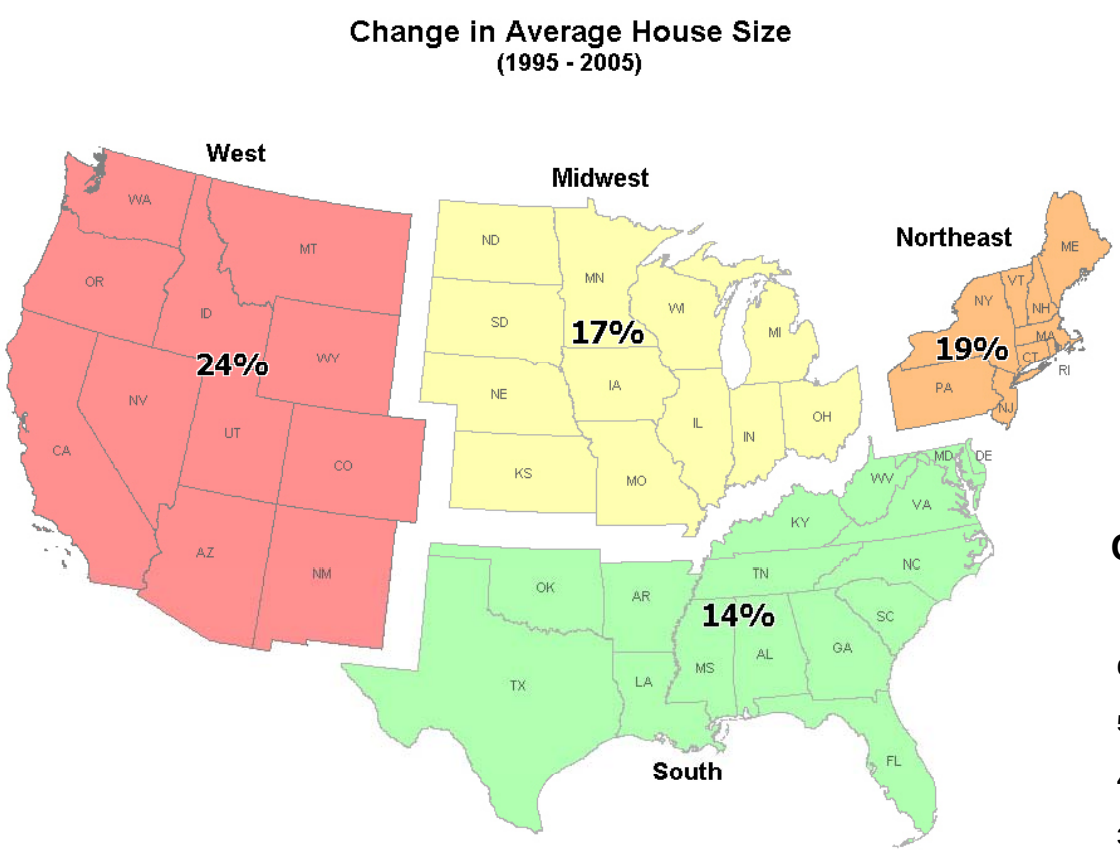
	Exact Address	ZIP Code Centroid	% Difference
AAL	$140	$145	3.37%
STD	$903	$933	3.32%
1.0% EP	$2,614	$2,914	11.48%
0.5% EP	$5,045	$5,331	5.67%
1.0% TVAR	$7,334	$7,632	4.07%
0.5% TVAR	$11,053	$11,445	3.55%

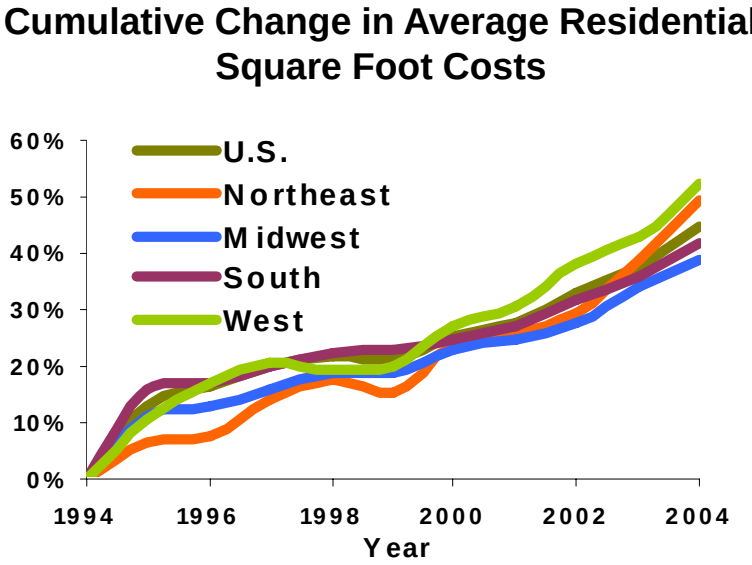
- On the portfolio level, the difference between exact address information and ZIP Code aggregate data is in the range of approxmately 1%.
- At a higher geographic resolution, however, the difference can be greater

Why Are Accurate Replacement Values Important?

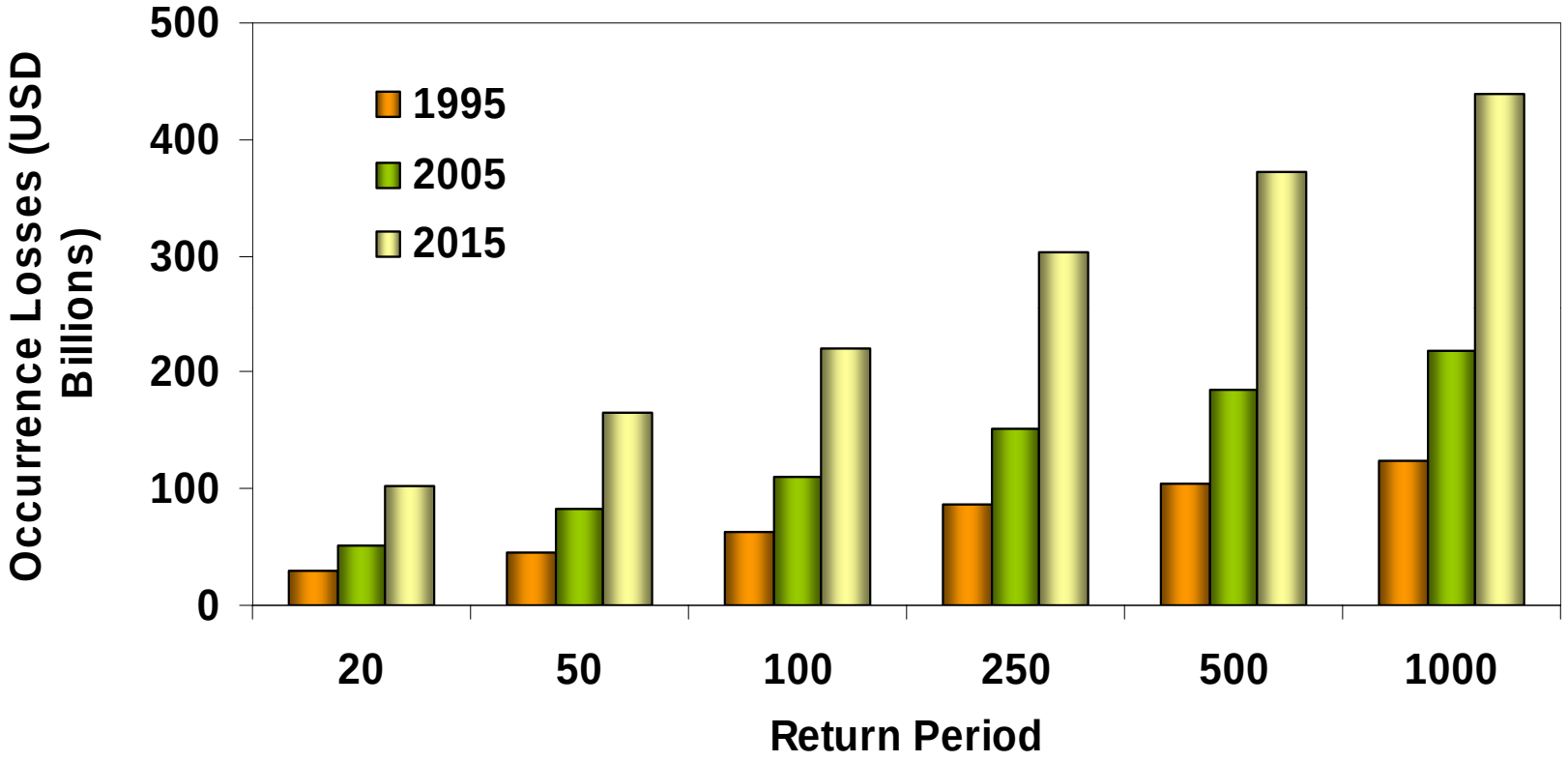
- Replacement value = full cost to replace the building in the event of a total loss
- Underestimation of the replacement value results in an underestimation of the estimated loss
- Coverage limits substituted for replacement values will not provide accurate loss estimates

New Homes Are Getting Larger, Architecturally More Complex and More Expensive





Exposure Growth Alone Results in a Doubling of AIR Industry Loss Estimates Over Past Decade



Losses include hurricane, earthquake, fire following, winterstorm, severe thunderstorm

Losses include aggregate demand surge

Impact of Accurate Replacement Value Information on the Quality of Loss Estimates

Full Portfolio

	Accurate Replacement Value	Replacement Value = Limits	% Difference
AAL	$1,698	$1,594	-6.15%
STD	$4,918	$4,640	-5.65%
1.0% EP	$25,325	$23,534	-7.07%
0.5% EP	$33,119	$31,006	-6.38%
1.0% TVAR	$40,425	$38,167	-5.58%
0.5% TVAR	$50,317	$47,614	-5.37%

Why Is Capturing Accurate Occupancy and Construction Important?

- Damage to buildings and their contents is a function of construction type and occupancy class
- At a given wind speed, light metal structures, for example, are more than three times as vulnerable as those built using reinforced concrete
- For a given construction type, occupancy provides insight into some of the key characteristics of a building, e.g. shape, window area and type of contents likely to be present

Impact of Accurate Construction Information on the Quality of Loss Estimates

Full Portfolio

	Specified Construction	Unknown Construction	% Difference
AAL	1,698	1,567	-7.72%
STD	4,918	4,406	-10.41%
1.0% EP	25,325	22,053	-12.92%
0.5% EP	33,119	29,121	-12.07%
1.0% TVAR	40,425	35,625	-11.87%
0.5% TVAR	50,317	44,931	-10.70%

Collier County, FL

	Specified Construction	Unknown Construction	% Difference
AAL	140	109	-22.05%
STD	903	710	-21.35%
1.0% EP	2,614	2.225	-14.90%
0.5% EP	5,045	4.461	-11.56%
1.0% TVAR	7,334	5.858	-20.12%
0.5% TVAR	11,053	8.783	-20.54%

- On the portfolio level, the difference between known and unknown construction information is in the range of 7 – 13%
- Looking at Collier County, however, where much of our exposure is of a multi-family occupancy class, changing the construction class to unknown causes vulnerability to increase to a greater degree than we experience for the portfolio as a whole

Is Knowing a Building is Made of Steel Construction Enough Security?

Construction Class - Steel Frame Occupancy Class - General Commercial Building Replacement Value - $50M Location - Zip Code 36602 (Mobile, Alabama)		
	AAL	**1 in 250-Year Loss**
No Secondary Risk Characteristics	81K	4.3M
Most Vulnerable Characteristics	141K	7.7M
Least Vulnerable Characteristics	29K	1.5M

Impact of Accurate Occupancy Information on the Quality of Loss Estimates

Full Portfolio

	Specified Occupany	Unknown Occupancy	% Difference
AAL	$1,698	$2,163	27.4%
STD	$4,918	$6,350	29.1%
1.0% EP	$25,325	$33,153	30.9%
0.5% EP	$33,119	$43,127	30.2%
1.0% TVAR	$40,425	$52,019	28.7%
0.5% TVAR	$50,317	$63,853	26.9%

Collier County, FL

	Specified Occupancy	Unknown Occupancy	% Difference
AAL	$140	$213	52.6%
STD	$903	$1,263	39.9%
1.0% EP	$2,614	$4,083	56.2%
0.5% EP	$5,046	$7,697	52.6%
1.0% TVAR	$7,334	$10,567	44.1%
0.5% TVAR	$11,052	$15,516	40.4%

- Occupancy yields important information about the types of contents subject to damage

Impact of Capturing Year Built Information on the Quality of Loss Estimates

Full Portfolio

	Specified Year Built	Unknown Year Built	% Difference
AAL	$1,698	$1,761	3.7%
STD	$4,918	$5,124	4.2%
1.0% EP	$25,325	$26,282	3.8%
0.5% EP	$33,119	$34,932	5.5%
1.0% TVAR	$40,424	$42,224	4.5%
0.5% TVAR	$50,317	$52,515	4.4%

Collier County, FL

	Specified Year Built	Unknown Year Built	% Difference
AAL	$140	$161	14.6%
STD	$903	$1,021	13.0%
1.0% EP	$2,614	$3,023	15.6%
0.5% EP	$5,046	$5,767	14.3%
1.0% TVAR	$7,334	$8,337	13.7%
0.5% TVAR	$11,053	$12,528	13.3%

- Year built gives us information about the building code and practices under which the exposure was constructed

Individual Risk Measures

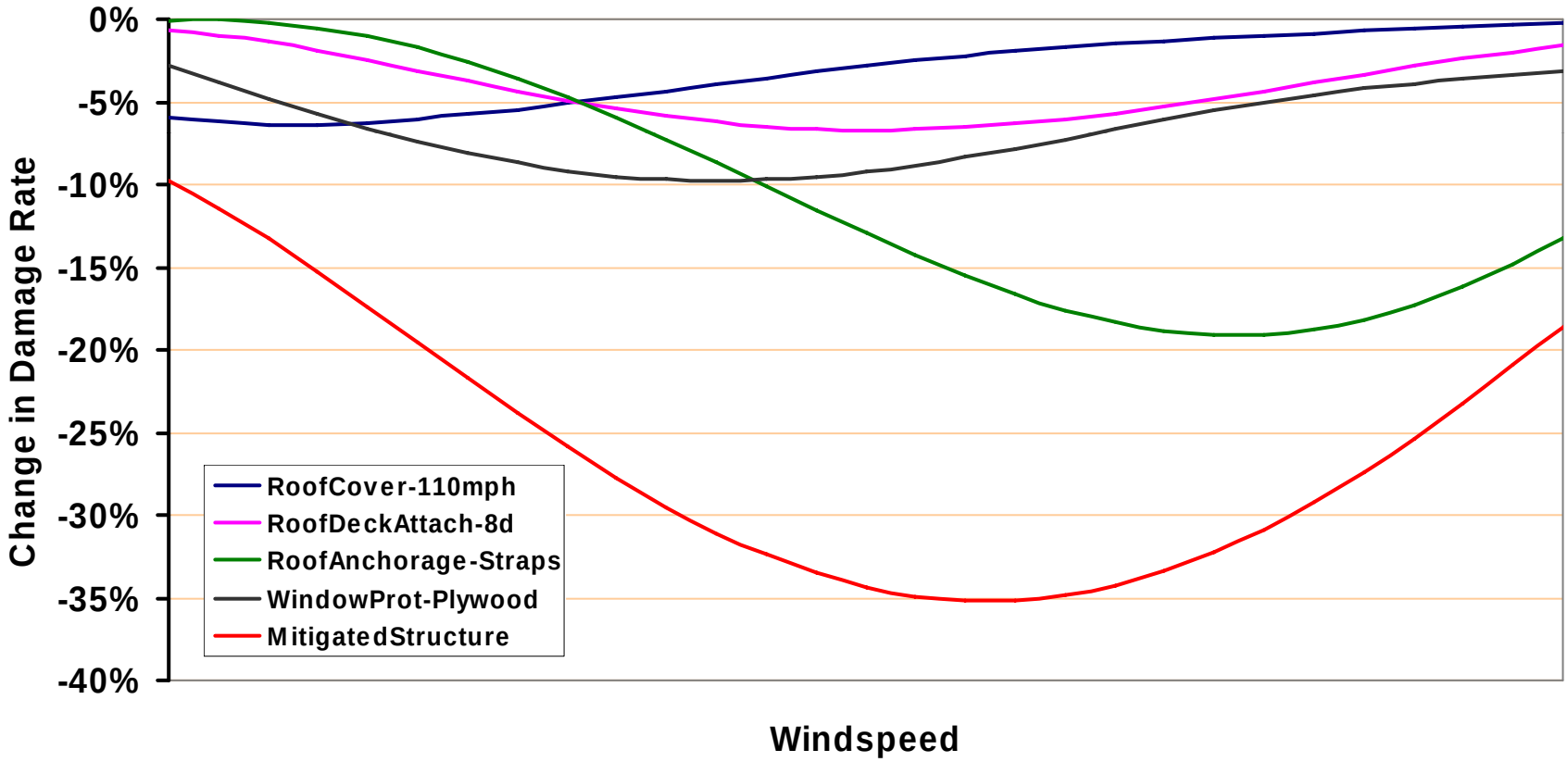
- Individual risk methodology follows a logic-based approach in combining multiple parameters to reduce to a modified damage function
- Considers both structural and non-structural damage
- Significance of parameters varies by construction class and intensity level
- Supports all possible combinations based on engineering principles and knowledge
- Allows study of mitigation effects including retrofitting

Options of Individual Risk Characteristics

Characteristic	Valid Options and Indices	Comments
Window glass type	0 = unknown/default 1 = Annealed 2 = Tempered 3 = Heat Strengthened 4 = Laminated 5 = Insulating Glass Units	Different glass types have varying degrees of resistance to wind loads and debris impact.
Window Protection	0 = unknown/default 1 = No protection 2 = Non-engineered Shutters 3 = Engineered Shutters	Protecting the windows can reduce the potential damage to a building.
Roof pitch	0 = unknown/default 1 = Flat (0°) 2 = Low (<10°) 3 = Medium (10° to 30°) 4 = High (> 30°)	Low pitch roofs have greater uplift forces acting on the roof as compared to high pitch roofs.
Roof covering	0 = unknown/default 1 = Asphalt Shingles 2 = Wooden Shingles 3 = Clay/Concrete Tiles 4 = Light Metal Panels 5 = Slate 6 = Built-up Roof with Gravel 7 = Single Ply Membrane 8 = Standing Seam Metal Roofs 9 = Built-up Roof without Gravel 10= Single Ply Membrane Ballasted	Damage to the roof covering can result in significant water damage to the interior of the building and contents.
Roof deck	0 = unknown/default 1 = Plywood 2 = Wood Planks 3 = Particle Board/OSB 4 = Metal Deck with Insulation Board 5 = Metal Deck with Concrete 6 = Precast Concrete Slabs 7 = Reinforced Concrete Slabs 8 = Light metal	Roof decks transfer the roof loads to the underlying joists and purlins. Damage to the roof deck results in the breach of the building envelope resulting in significant building and interior damage.

Impact of Individual Risk Characteristics Varies Based on Perils and Location of Exposures

- The impact of individual risk characteristics varies over the ranges of potential intensities



Sample Impact of Individual Risk Characteristics

- Primary risk characteristics
  - Construction type - Reinforced concrete
  - Occupancy - General commercial
  - Height - 3 stories
  - Age - Unknown
- Effect of individual risk characteristics

	Expected	250 Year	500 Year	Expected	250 Year	500 Year
				Relative Change from Base Case		
Base Case	270	16,362	20,886			
Base Case + Soft Story	395	25,096	33,897	1.46	1.53	1.62
Base Case + L Shape	331	20,484	27,297	1.23	1.25	1.31
Base Case + Soft Story + L Shape	423	27,197	36,898	1.57	1.66	1.77

Impact of Individual Risk Characteristics on the Quality of Loss Estimates

Full Portfolio

	Window Protection Provided	Unknown Window Protection	% Difference
AAL	$1,692	$1,761	4.1%
STD	$4,932	$5,124	3.9%
1.0% EP	$25,265	$26,282	4.0%
0.5% EP	$33,462	$34,932	4.4%
1.0% TVAR	$40,693	$42,224	3.8%
0.5% TVAR	$50,694	$52,515	3.6%

Collier County, FL

	Window Protection Provided	Unknown Windown Protection	% Difference
AAL	$153	$160	4.8%
STD	$981	$1,021	3.9%
1.0% EP	$2,872	$3,023	5.2%
0.5% EP	$5,432	$5,767	6.2%
1.0% TVAR	$7,984	$8,337	4.4%
0.5% TVAR	$12,046	$12,528	4.0%

Individual Risk Characteristics

- For a commercial writer in North Carolina, inclusion of storm shutters can mitigate hurricane losses by over 17%
- The inclusion of storm shutters on the same mix of commercial exposure shifted to Massachusetts mitigates losses by only 8%
- In North Carolina, the inclusion of storm shutters had a greater impact than roof geometry. In Massachusetts, roof geometry has a greater impact than storm shutters (12% mitigation versus 8%)

- Collection of individual risk characteristics should be based both on the perils you cover and the location of exposures

Summary

- Model results rely on the quality of data inputs
- Models are a valuable tool to evaluate the impact of various levels of data resolution
- The value of more data varies by peril, location and nature of your underlying book of business
- Models can be used in determining the cost/benefit of collecting more data

- ***Run the model !***