Institute and Faculty of Actuaries

Nat Cat Validation Framework Working Party

Caroline McMullan (AIR Worldwide)
Junaid Seria (SCOR SE)
Jessica Turner (Lloyds Banking Group)

12 April 2016

Institute and Faculty of Actuaries

Agenda

1. Working Party mandate
2. Observed pitfalls
3. What is validation?
4. Validation Framework
5. Case Studies
6. Next Steps

We need your feedback!

Institute and Faculty of Actuaries

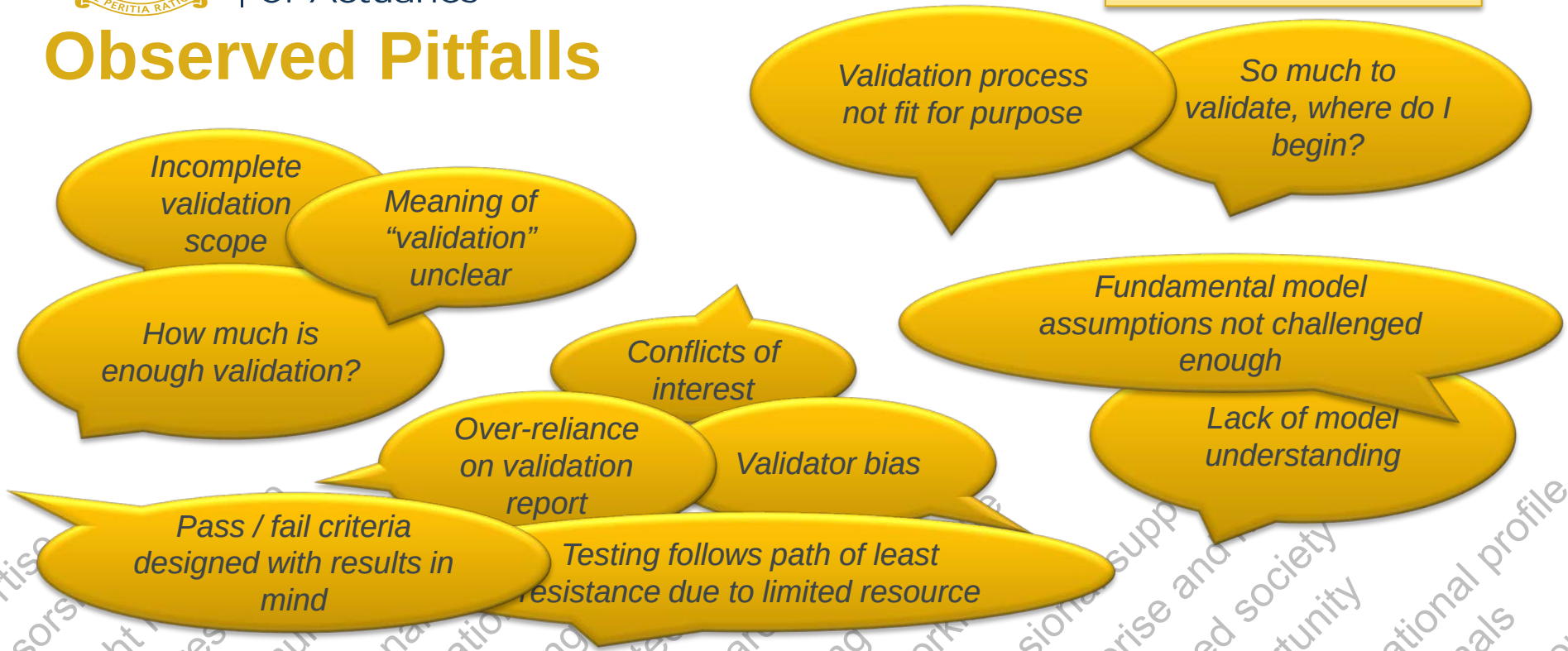
Working Party Mandate

1. **Assemble** a group of cat modelling industry practitioners with relevant skill sets
2. **Consider** the complexity and constraints of cat models, available cat vendor validation documentation, Solvency II requirements and current issues.
3. **Provide guidance** to cat modellers, internal and external validators and business users of cat model outputs on what constitutes “good enough” validation of catastrophe risk.
4. **Provide** practical validation **examples** that illustrate the principle of proportionality

Institute and Faculty of Actuaries

We need your feedback!

Observed Pitfalls



Institute and Faculty of Actuaries

Incomplete validation scope

Meaning of "validation" unclear

What is Validation?

1. A means of gaining **confidence** over all material aspect of the modelling process
   - Inputs
   - Cat model components
   - Results
   - Purpose / Use: primarily in the internal model, for pricing, reinsurance purchase, exposure management and business planning
   - Control environment

Institute and Faculty of Actuaries

What is Validation?

Validation process not fit for purpose

Meaning of "validation" unclear

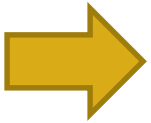
1. Evaluation of the suitability of the science underlying the model?
2. Validation of each model component's "performance" against scientific and statistical expectations and empirical evidence?
3. Validation of modelled losses against historical losses?
4. Assess whether checks and controls were performed through the modelling process and that these are documented?
5. Assess the extent to which the modelling process complies with relevant internal guidelines?
6. Evaluate the firm's understanding of cat models?
7. Validate adherence to Solvency II requirements?

Institute and Faculty of Actuaries

Testing follows path of least resistance due to limited resource

What is Validation?

1. Tick-the-box exercise.
2. Evidence selected to support the current approach and calibration
3. Heavily biased by previous choices
4. Little / no use by C-suite



1. Genuine attempt to find an alternative hypothesis to the adopted view
2. “Model inquisitiveness”
3. Unbiased / awareness of bias
4. A communication tool

Institute and Faculty of Actuaries

How much is enough validation?

So much to validate, where do I begin?

Validation Framework

Institute and Faculty of Actuaries

Validation Framework

Test Topics

1. Data
2. Model Design
   - hazard,
   - event set
   - vulnerability
   - financial
3. Results
4. Key drivers:
   - Expert judgements
   - Key assumptions
   - Key switches/options
   - Key distribution choices

5. Governance / Control environment (of items1- 4 above)

Test Structure

1. Test description
   - Context / Issue
   - Materiality
   - Scope
   - Objective
   - Limitations
2. Quantitative or Qualitative
3. Pass / fail criteria
   - what is the hypothesis / expectation?
4. Test result and rationale
5. Conclusion / Recommendation

Test Tools

What
- Analysis of change test
- **Back-test**
- **Benchmarking test**
- **Stress test**
- Scenario test
- **Sensitivity test**
- Functional test
- **Reverse Stress Test**
- Risk attribution test

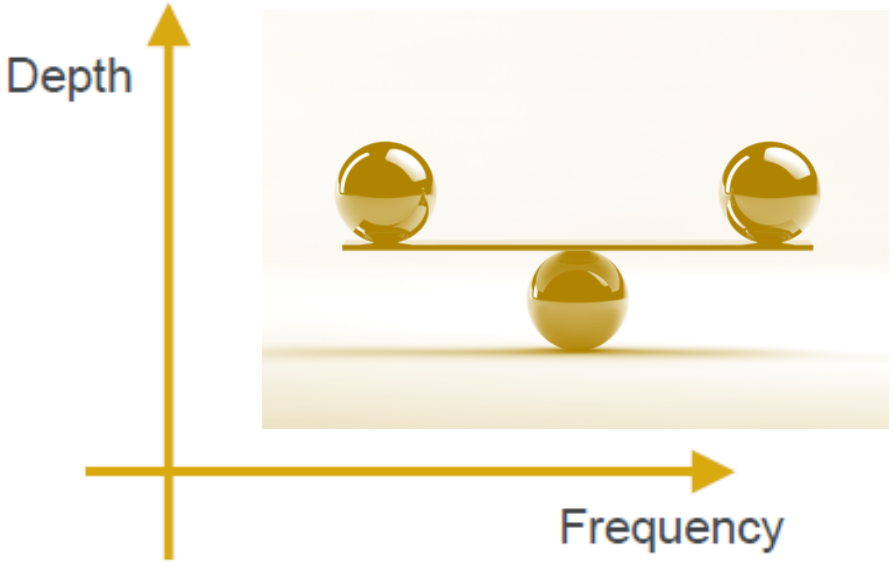
Why
- Completeness
- Consistency
- Reconciliation
- Replication
- Transparency
- Stability
- Comparison of observed against modelled

What do you think?

Institute and Faculty of Actuaries

Validation Process

1. "Deep dive" of cat model component(s)
   - Leverage vendor validation
   - Infrequent (every 3 – 5 years)
   - Requires expert knowledge
2. *Basic* frequent:
   - Renewing tests (with each model run)
   - Does not require expert knowledge
   - Could be automated
3. *Intuitive* frequent
   - Rules of thumb
   - Easy to understand, hence more likely to be referenced in C-suite discussion



Validation process not fit for purpose

Institute and Faculty of Actuaries

Region-Peril Case Studies

Case Study 1: UK Wind back-test
Case Study 2: Taiwan EQ sensitivity test
Case Study 3: US Wind Stress test
Case Study 4: JP Wind Stress test
Case Study 5: Global Reverse Stress test
Case Study 6: EU Wind Benchmarking test

Test Design

Results & Conclusions

All data and results are hypothetical

Case Study 1

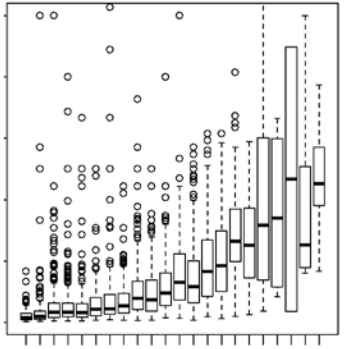
UK Wind back-test
Test Design

Design	Test Topic > Model design > UK Wind Vulnerability		Test Tool > back-test
Context & Objective	• A **fundamental component** of cat model is the assumed damage to insured property resulting from the operation of an insured peril. • Vendors use **market claims data** to calibrate damage (as % of exposure) across hazard intensities, supplemented by risk engineer expert judgement where data is scarce. • This test aims to assess the **appropriateness** of this calibration and **suitability** of expert judgements made for a sub-set of vulnerability curves	**Test inputs**	1. Wind intensity data from your vendor model 2. Location-level claims from the event(s) 3. In-force exposures at time of event (s) 4. Loss adjuster expertise 5. Published Engineering studies
Materiality, Scope & Limitations	• **High** (>5% impact on net cat risk 1:200 AEP VaR / 1:100 AEP TVaR, for example – depends on purpose) • **Property** class (PD + consequential loss (where applicable)) • **UK Residential Windstorm** peril (most suited to residential property insurers) for specific events • **Scarce data** at high hazard intensities • **Data quality** / coding errors in wind damage claims • **Regional vulnerability** differences may be difficult to test	**Test Steps**	• Assign wind intensity to internal claims data • Calculate mean conditional damage ratio by wind intensity and vulnerability region for affected policies • Compute claim frequency at above granularity • Compute observed damage ratio and compare to modelled mean damage ratio (MDR) where data is credible. • Validate assumed modelled damage at high hazard intensities with internal claims adjustors.
Nature	• Quantitative (observed vs. modelled) • Qualitative (evaluation of the suitability / reasonableness of engineering judgements at high hazard intensities)	**Pass / Fail criteria**	• Overall **shape** of vulnerability curve ("…vendor residential curves reflect changes in the rate of damage across hazard intensities..") • Calibrated **damage ratios** ("…observed damage within x% of modelled…") and observed exceptions don't reveal any systemic bias

Case Study 1

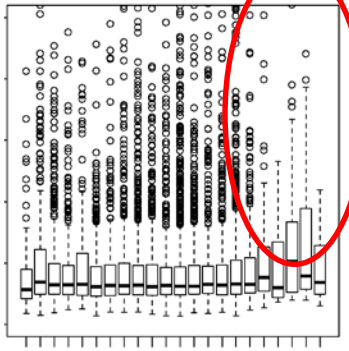
UK Wind back-test
Results & Conclusion

Claim Frequency



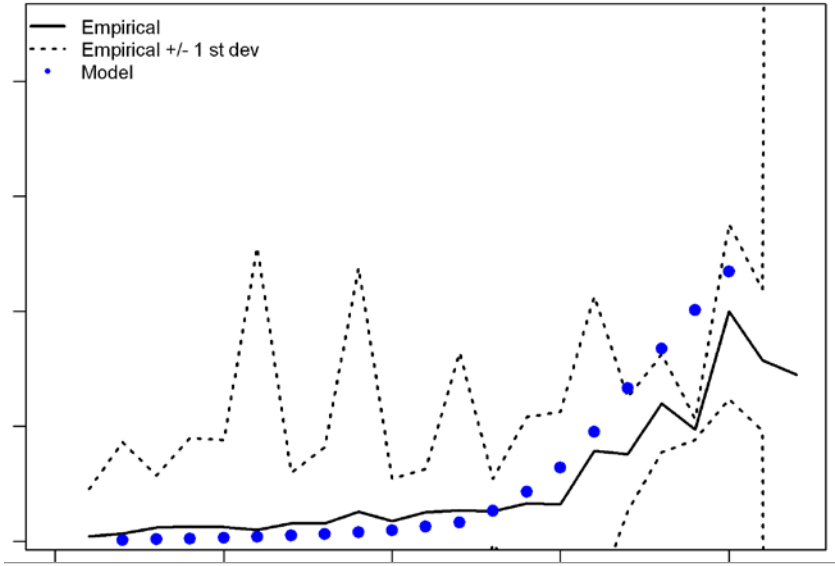
*

CDR



=

MDR



$$MDR = \overline{claim} / \overline{SI} = \overline{cf}_{v,dc} * \overline{CDR}_{v,dc}$$

cf = claim frequency

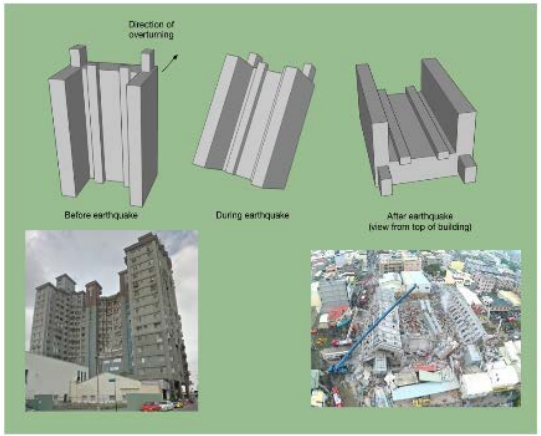
v = peak gust / wind intensity measure

dc = damage curve (vulnerability) region

CDR = conditional damage ratio

For illustration: execution will depend on cat model used and data available

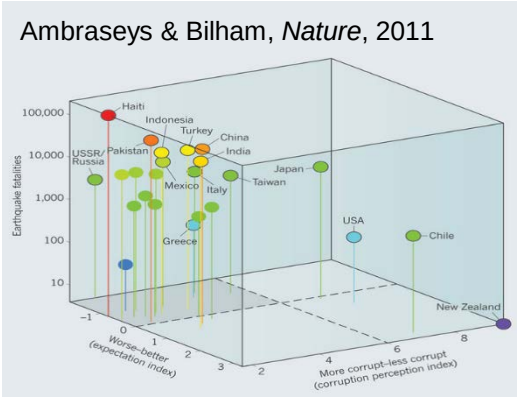
Case Study 2



M6.3 Tainan, Taiwan, 2016

M7.6 Chi-Chi, Taiwan, 1999

Ambraseys & Bilham, *Nature*, 2011



Taiwan EQ (TWEQ)
Type: sensitivity test
Context

Modelled vulnerability of insured property assumes seismic resistance based on specified building characteristics.

However, corrupt business practices may lead to certain buildings displaying no seismic resistance despite classification as “reinforced concrete” thus potentially under-stating losses.

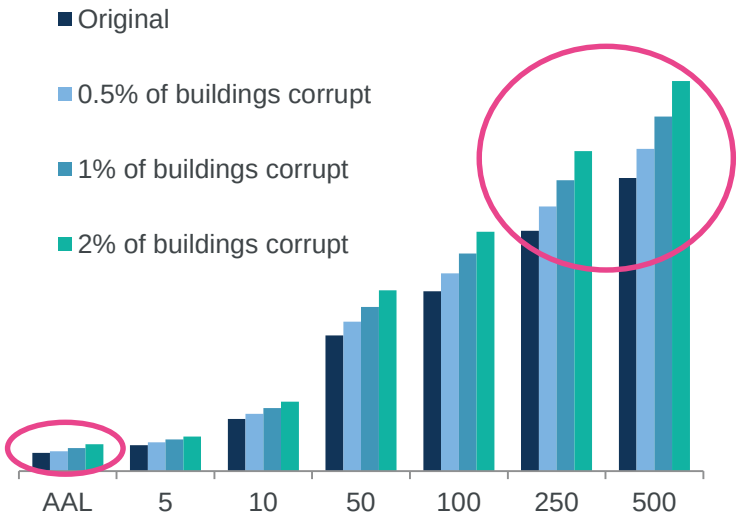
While systemic non-adherence due to corrupt practices will likely reflect in historic claims, this sensitivity test assesses impact on TWEQ results due to sporadic non-adherence where construction practices are generally adequate.

Institute and Faculty of Actuaries

Case Study 2

Test Topic > Model design > Construction Data	
Key question	Is the sensitivity of modelled losses to different levels of assumed corrupt building practices sufficiently high to apply an uncertainty loading to modelled results?
Scope	Taiwan Quake property portfolio, however as per *Ambraseys & Bilham* (2011) the test could be applied to other countries where there are concerns about corrupt construction practices
Nature	Quantitative
Pass/Fail Criteria	If ratio of original loss to loss under sensitivity test > some pre-defined X then apply a loading factor of Y (consider non-linear adjustment if necessary
Test Steps	Recode X amount of locations or replacement value within a country portfolio with a construction code that provides little/no seismic resistance

Taiwan EQ (TWEQ)
Type: sensitivity test
Test Design and Results

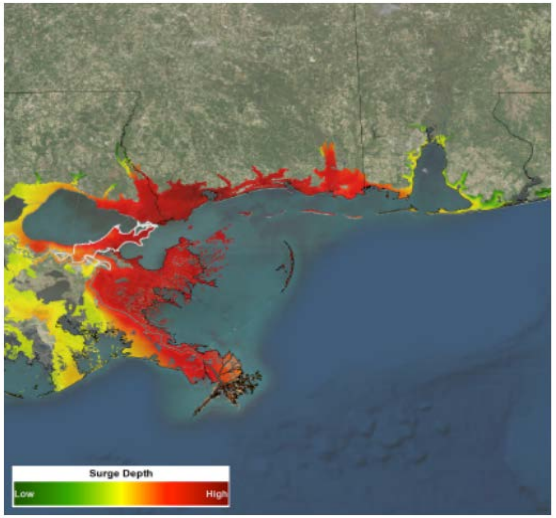


Institute and Faculty of Actuaries

Case Study 3

Test Topic > Results	
Context	2005 Hurricane Katrina was the worst natural disaster the insurance industry faced (Insured loss = $41bn). While property exposures and as-if claims costs are significantly higher, levee enhancements, tighter policy wordings and improved building code adoption reduce severity significantly. The new levee system provides protection for water levels with a return period of 100 years
Key question	Does the Katrina levee failure scenario produce losses within the range of modelled outcomes / within 1:250 OEP VaR?
Materiality	High (US Wind is a peak peril)
Nature	Quantitative
Scope	US Wind + Surge. Property class
Pass/Fail Criteria	Test passes if the estimated loss scenario is within the range of modelled loss outcomes
Test Steps	Use vendor published losses and apply market share or run historical event ID where available.

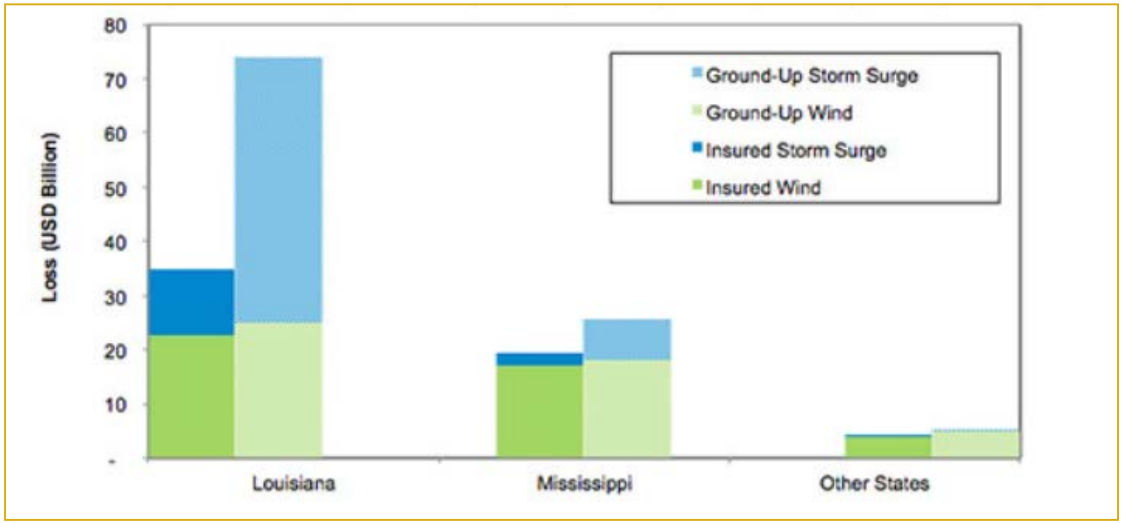
Scope: US Wind region-peril
Type: Stress test
Test Design



Storm surge footprint for Katrina scenario with levee failure

Institute and Faculty of Actuaries

Case Study 3



Insured and ground-up losses by wind and storm surge for a storm with the characteristics of the 2005 Hurricane Katrina, with the New Orleans flood defenses failing. (Source: AIR)

Scope: US Wind
Type: Stress test
Test Result

Apply market share to $60bn loss estimate considering possible portfolio biases

Pass: Resulting loss corresponds to 1: 150 on modelled OEP VaR

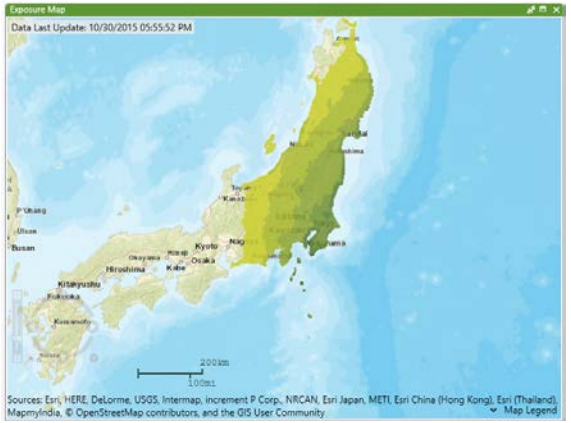
Institute and Faculty of Actuaries

Case Study 4

Scope: JP Wind
Type: Stress test
Test Design

Test Topic > Results	
Context	Typhoon Vera (1959) is the costliest weather-related event in Japanese recorded history.
Key question	Does the vendor derived stress scenario (Vera track with increased intensity and more proximal Tokyo path) produce losses within the range of modelled outcomes?
Materiality	Medium / High
Nature	Quantitative
Pass/Fail Criteria:	Stress test passes if estimated losses are within the range of modelled outcomes?
Test Steps	1. Source AIR extreme-disaster scenario (EDS) for the modified Vera path and wind speeds and apply an appropriate market share 2. Alternatively, find closest stochastic match from event set of chosen cat model vendor.

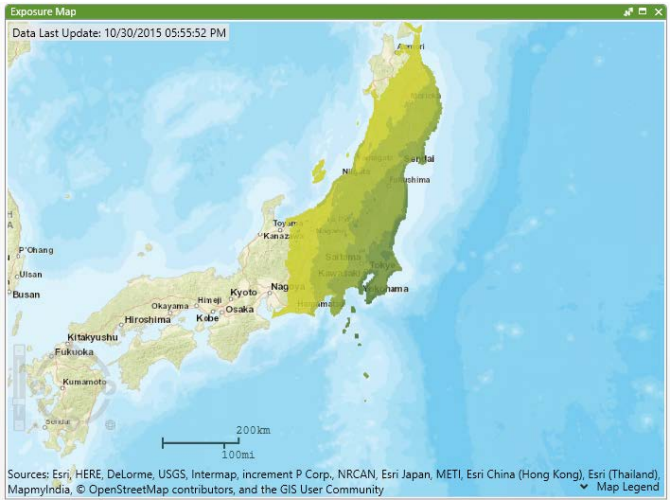
- Track shifted to hit between Tokyo & Kagoshima
- Vera's central pressure reduced by 10mb
- Astronomical tide phased to cause max surge height



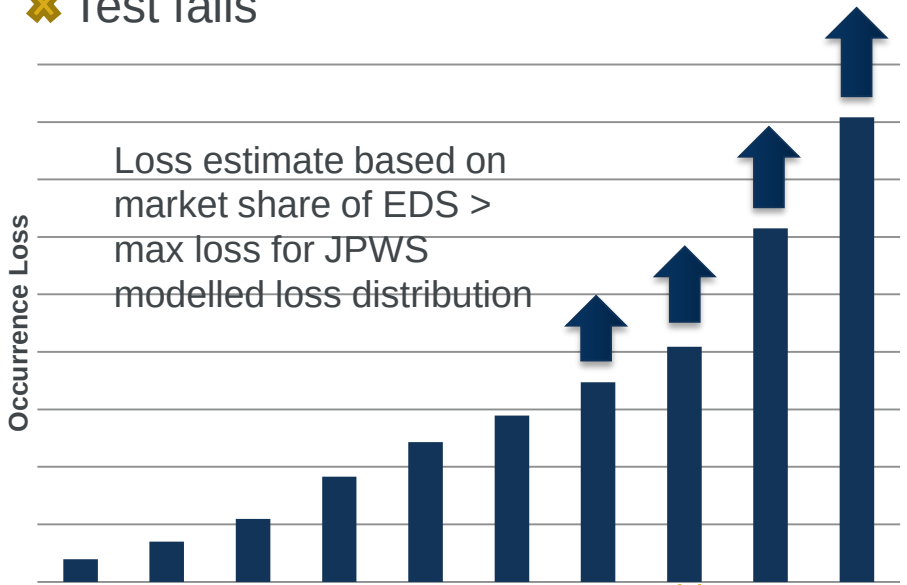
Case Study 4

Scope: JP Wind
Type: Stress test
Test Results

Extreme Disaster Scenario:
Typhoon Vera variant
wind speed footprint



✖ Test fails



Institute and Faculty of Actuaries

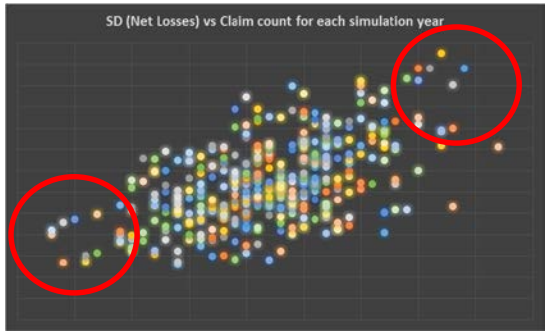
Case Study 5

Scope: Global perils
Type: Reverse Stress test
Test Design

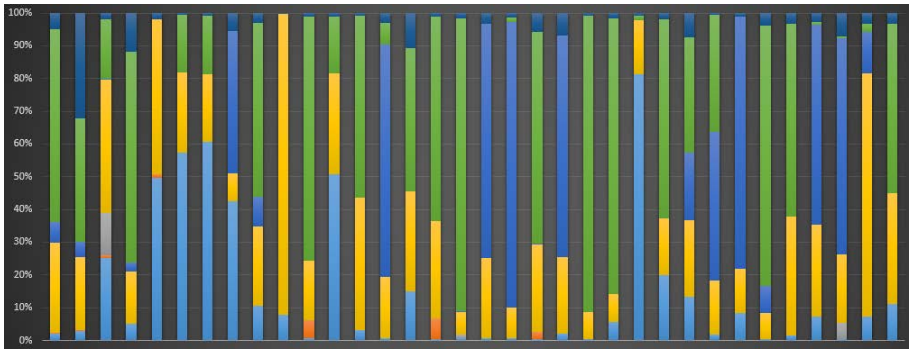
Test Topic > Results	
Context	As part of the demonstration that the capital requirements resulting from the internal model are appropriate, …undertakings shall **compare the coverage and scope** of the internal model… Identify the **most probable stresses** that would threaten the **viability** of the insurance or reinsurance undertaking." (Art 242 (6))
Materiality	High
Nature	Quantitative, while pass / fail criteria assessed qualitatively
Pass/Fail Criteria	Type 1: Test fails if available capital eroded by an as-if historical event / series of events not considered extreme – further management action required (e.g., buy additional Cat XoL protection) Type 2: Test fails if Internal model simulation years around the 1:200 VaR are not reflective of the firm's cat risk profile (e.g., results not driven by key risk drivers)
Test Steps:	1. Compute independent of the cat model the largest as-if historical losses and compare against available capital (including allowance for expected non-cat losses) 2. Extract net losses for simulation years exceeding a capital threshold / in a defined range and assess the suitability of how losses in the each simulation year are composed

Case Study 5

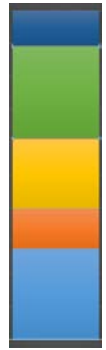
Scope: Global perils
Type: Reverse Stress test
Test Results



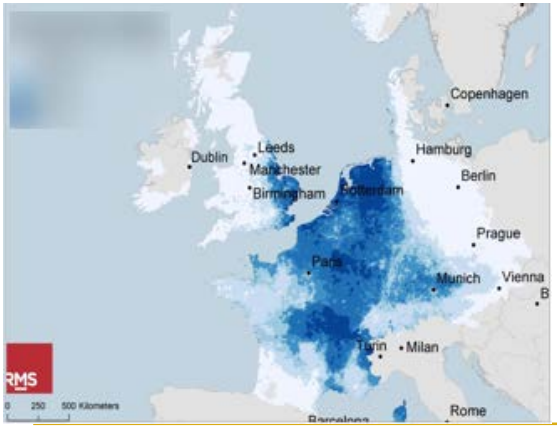
Evaluation of modelled simulation years



Selected



Assess Management Action



Plausible, but too extreme (~5x Windstorm Daria) therefore not selected

Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

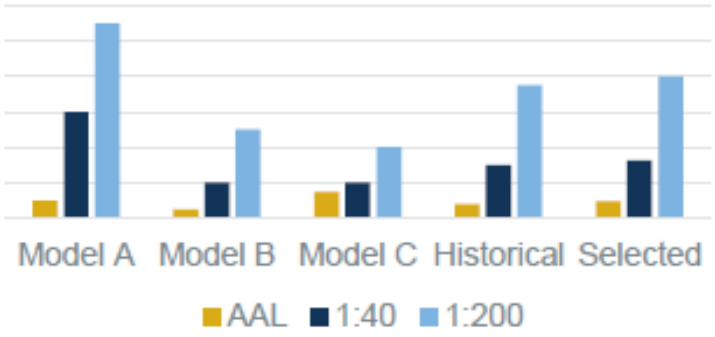
Case Study 6

Scope: EU Wind
Type: Benchmark test

Benchmark test: Clustering

1. Benchmark selected choice against vendor models and history
2. Various perspectives available:

- AEP / OEP
- Var (Claim count/yr) / mean claim count per yr
- Your AEP / Poisson AEP
- AEP clustered / unclustered (where vendor provides these views)
- Likelihood of having 2 events as large as [Daria] in the same year (benchmark models)
- Likelihood of having 3 events as large as [Herta, Vivian, Kyrill] – EU wide / country level



Despite the uncertainties, is our selection in a reasonable range?

Institute and Faculty of Actuaries

Next steps

- Volunteers to apply framework and conduct tests for region-perils globally that illustrate principle of proportionality
- Consider tests applied in the validation of ESGs, internal model dependencies and banking models that could be applied in the validation of nat cat risk
- Conduct a literature review of relevant material to help cat risk validation practitioners / C-suite
- Provide examples of what a high-level / executive summary to a validation report could look like to aid engagement by C-suite members

Questions

Comments

Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenter.

Junaid Seria jseria@scor.com

Institute and Faculty of Actuaries