Big Data Impact on Optimal (Re)Insurance Premium Pricing and Limits Accumulations

Ivelin Zvezdov, AIR Worldwide

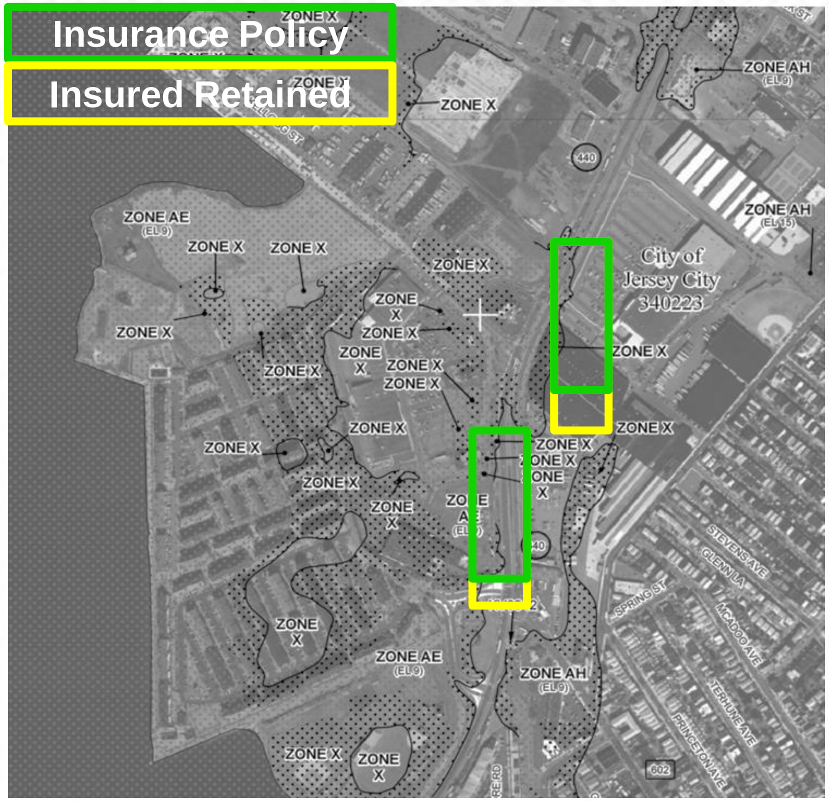
AIR WORLDWIDE®

Agenda

- Set up the objective and the case study
- Premium components and “thin data” pricing
  - Catastrophe platforms & big data tools for premium pricing
- Dependencies in premium accumulation: umbrella policies
- Dependencies in limits accumulation: likelihood of full payout
- Constructing fair premium bounding intervals

Setting Up the Objective and Defining the Case Study

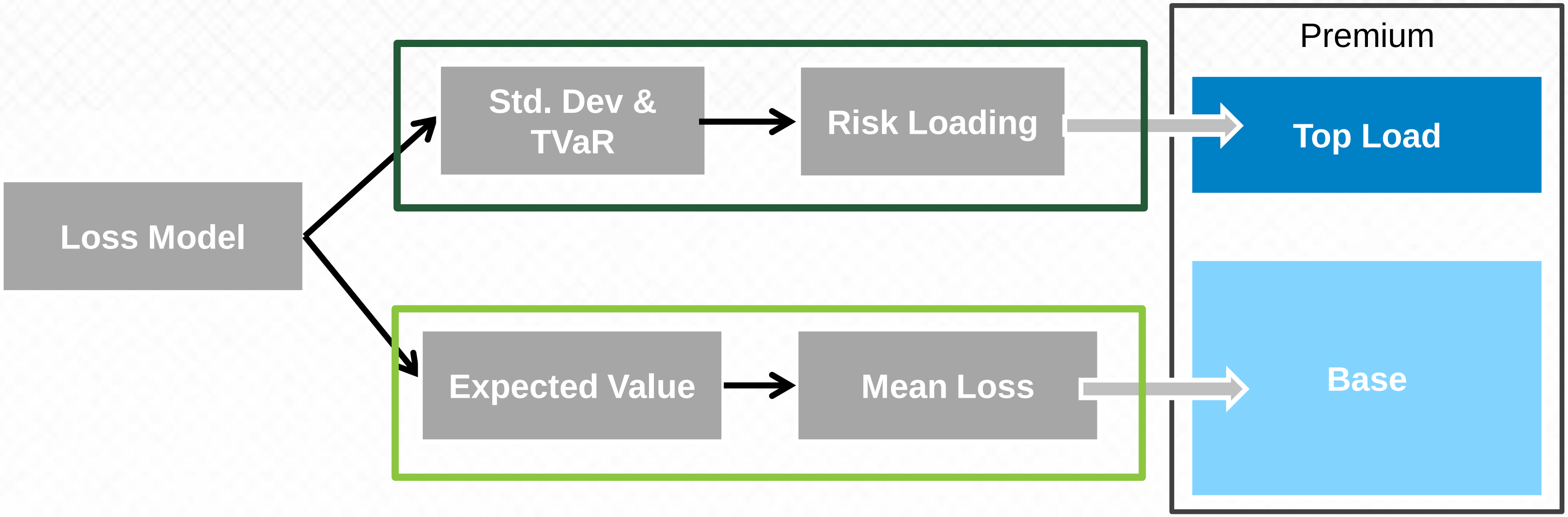
- The Objective
  - Derive fair technical flood CAT premiums for commercial risks



- The Case Study
  - Two risks in a geo-spatial area with minimal or none historical claims
  - A Historical flood map is available
- The Problem
  - Define the shortcomings of ‘thin data’ in premium pricing
  - Discuss the role of catastrophe models and big data platform in ‘sustainable’ pricing and limits accumulation

Premium Components and ‘Thin Data’ Pricing

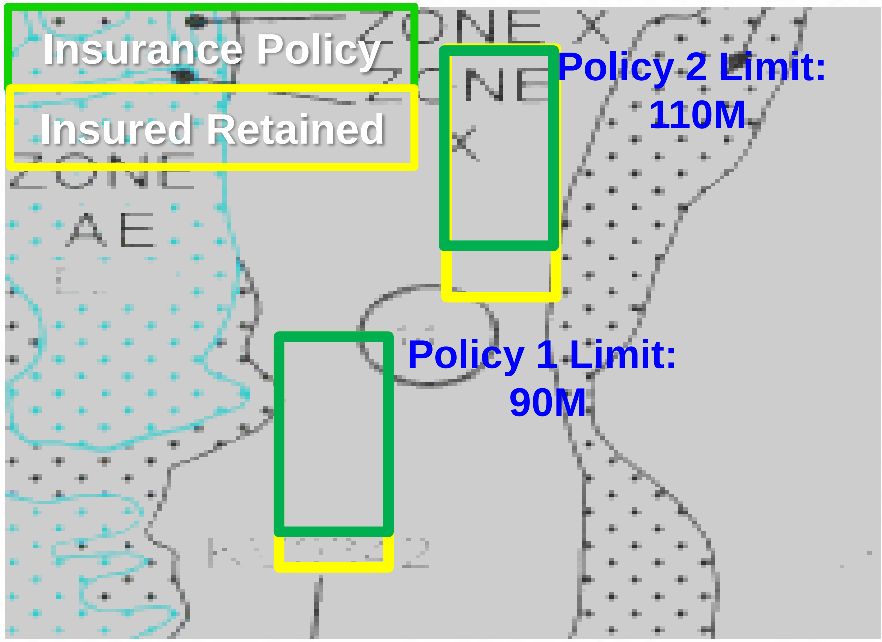
Technical Components of Catastrophe Premium Pricing



- Base Premium depends on modeled means, or averages of historical data
  - Average historical loss
- Top Premium Load(s) depend on:
  - Variability of loss and standard deviation
  - TVAR & full loss distribution

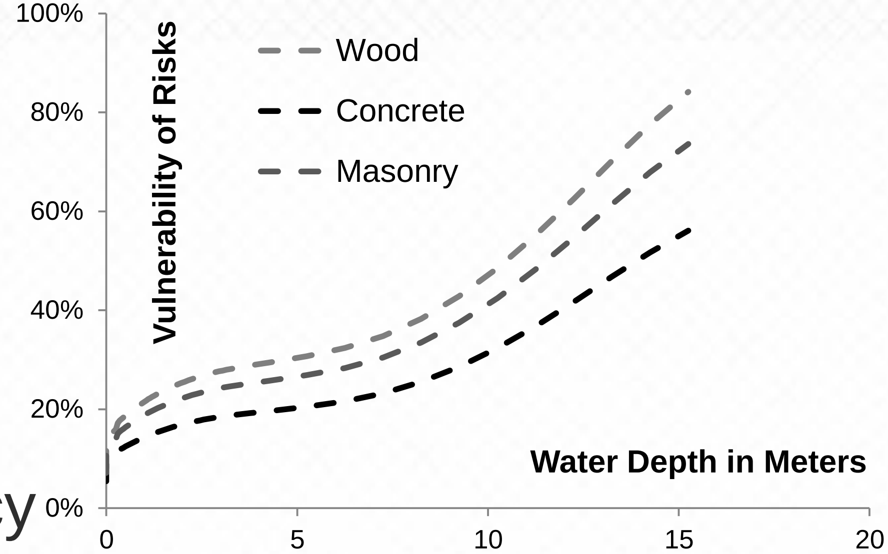
Base Premium and Components of 'Thin Data' Pricing: Historical Flood Map

- Average historical flood intensity at the insured risk location
  - Insured Risks are not in a Historical flood zone
  - Geo / Proximity to:
    - 100 year Zone with 1 feet of Flood depth
    - 500 year Zone with 3 feet of Flood depth



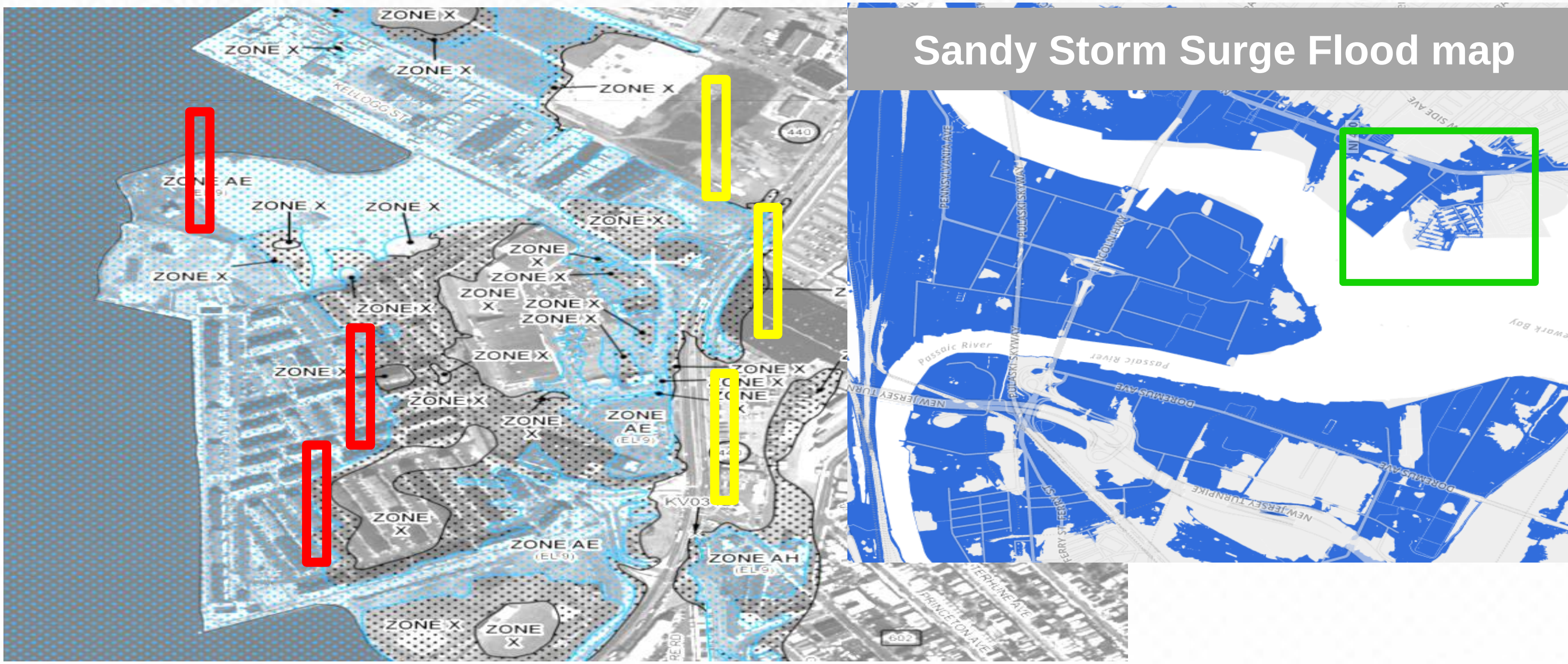
Base Premium Definition:
Damage and Vulnerability Factor Table

- Relationship between physical peril intensity and Insured risk damageability
- Derived from historical Engineering experience
- Using Water depth in nearest Flood Zone – i.e., low accuracy



Nearest Flood Zone Water Depth	Approximate Vulnerability	Actual Water Depth & Vulnerability
100 Year, 1 foot	9%	None
500 Year, 3 feet	12%	None

Some Shortcomings with the Base Premium Definition with Minimal 'Thin Data'

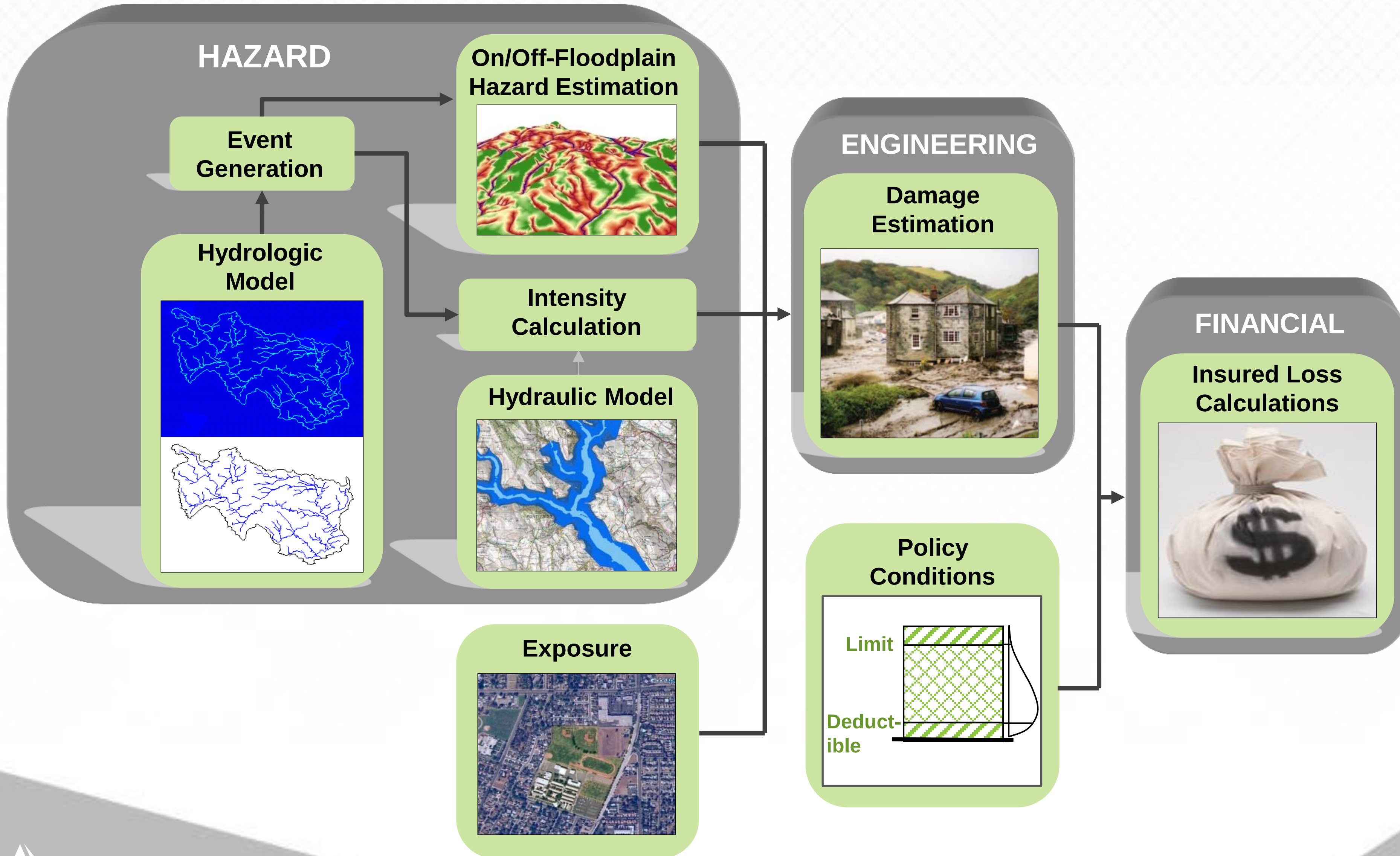


- Risks outside of 0.2% or 1% annual chance flood planes - no intensity data
- Uniformity of Flood Intensity for Large Areas – flat riskiness and vulnerability
- Blend for Coastal and River-Rain Flood maps is not available in near future

Catastrophe Modeling Platforms and Big Data Tools for Premium Pricing

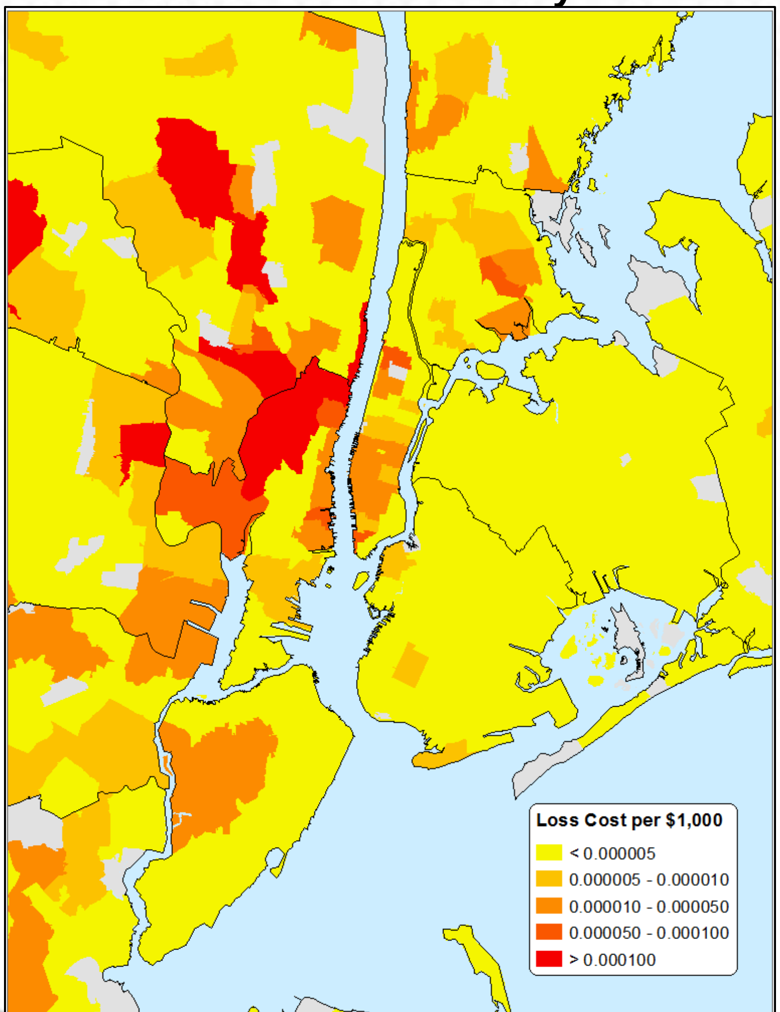
AIR

Financial & Hazard Big Data Components Are Contained in a Catastrophe Modeling Platform

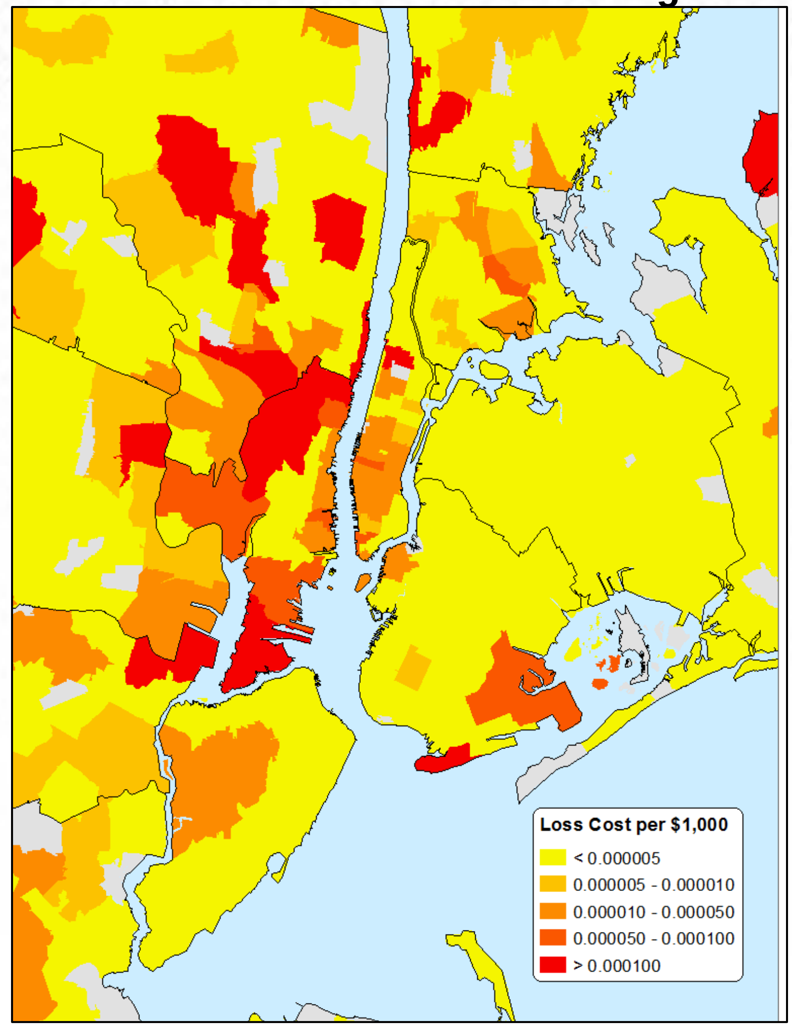


Catastrophe Modeling Platforms Provide Aggregate River Flood and Storm Surge Flood View of Risk at High Resolution

Inland Flood Only

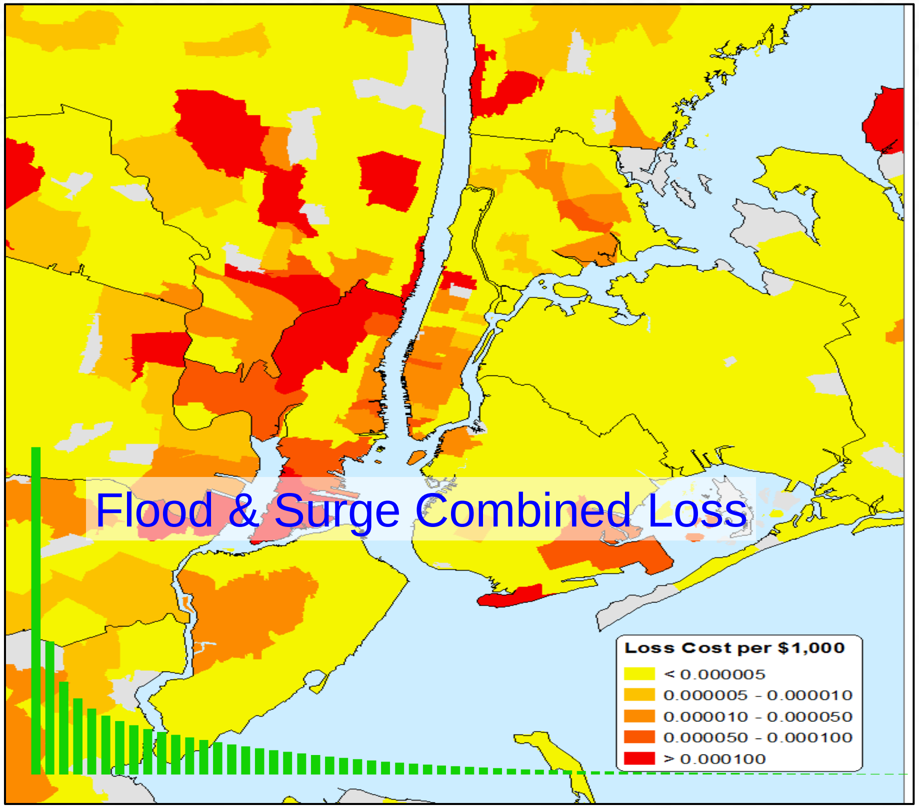


Inland Flood and Storm Surge



Financial Data Layer: Premium Pricing Components from Modeled Stochastic Loss

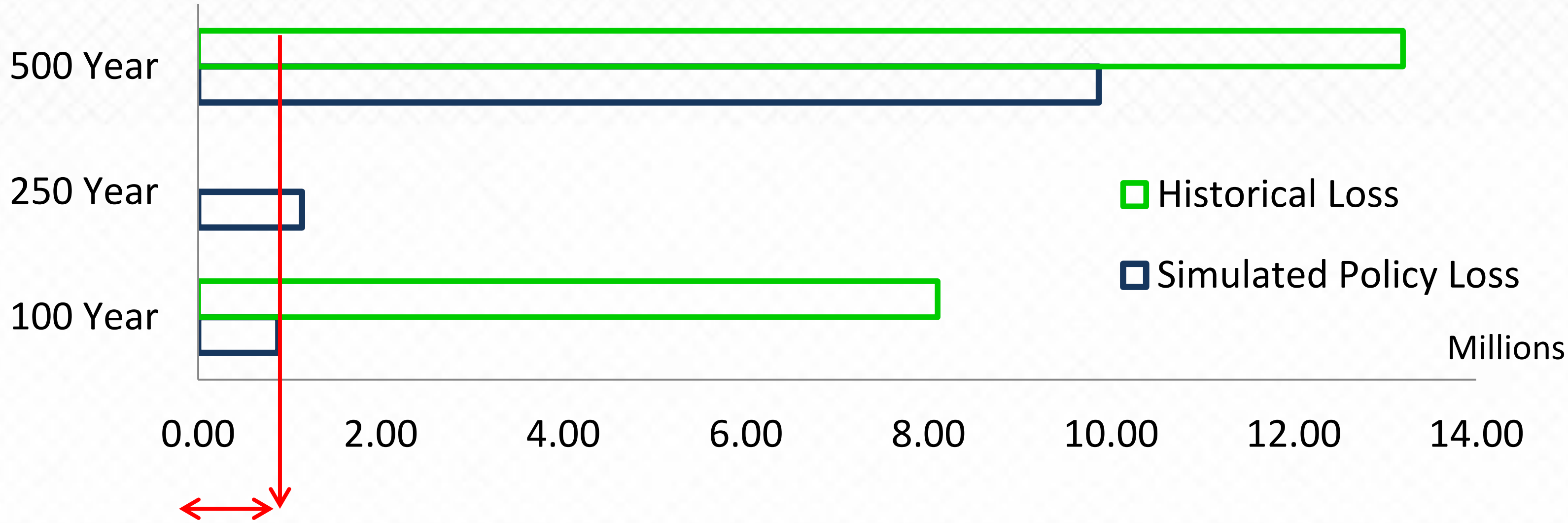
- Stochastic full distributions for combined inland flood and coastal storm surge loss



- Summary per-risk stats
  - Mean and Standard Deviation
  - PML and TVaR
- Basic Premium Pricing Principle

$$Premium(risk) = Expected\ Loss + k * Standard\ Deviation$$

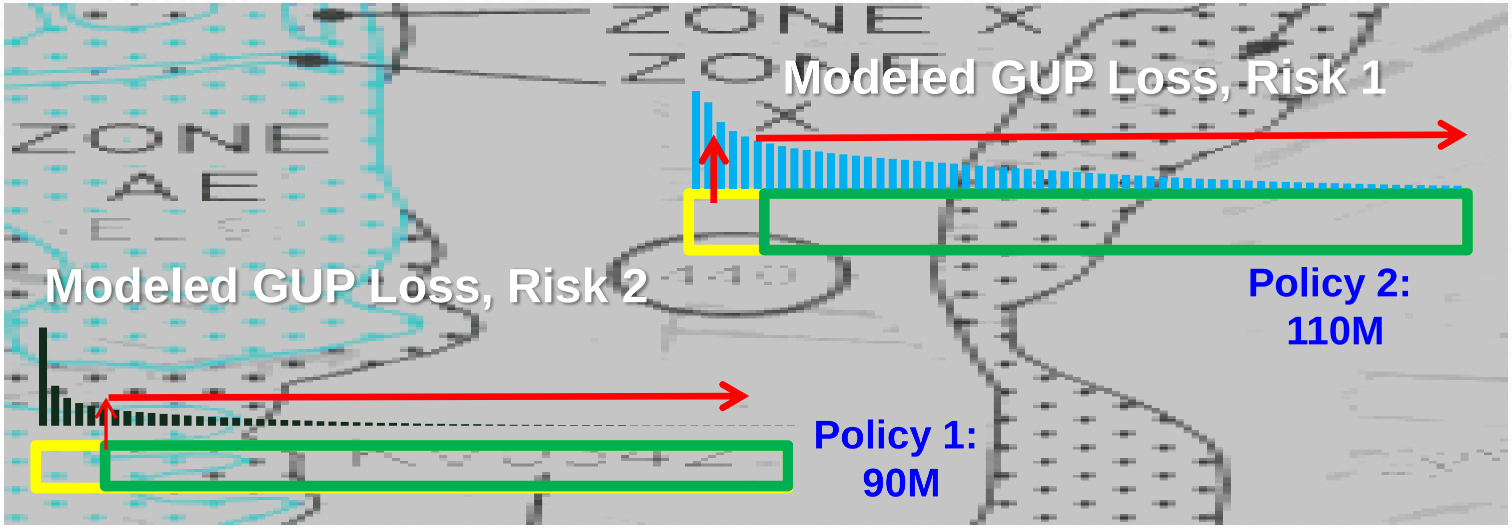
Fair Technical Premium Pricing: Basic Methods



$$\textit{Pure Technical CAT Premium} = AAL + k * \textit{Standard Deviation}$$

- Historical loss is fully arbitrary, interpolated from nearest flood zone intensity
- Risk loading measured by standard deviation is not available with the historical data approach

Complex Underwriting Metrics: Probability of breaching Deductibles



Deductible % of TIV	Probability of Breaching Deductible	
	Policy 1	Policy 2
1%	0.046	0.041
2%	0.036	0.033
3%	0.358	0.327

- Complex risk and underwriting metrics require ‘rich’ data sets
- Modeling and big data platforms preserve full single risk distributions

Dependencies in Premium Accumulation: Umbrella Policies

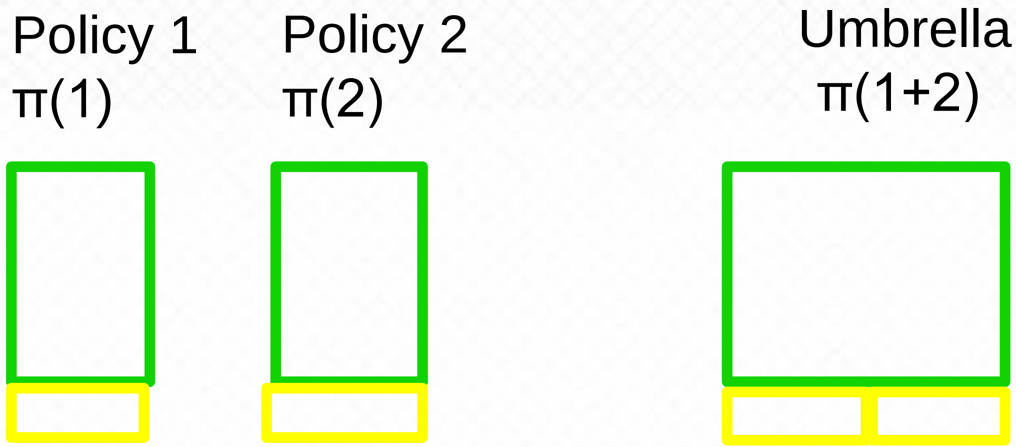
AIR

A Business Case for Accounting for Dependencies in Premium Pricing

Policy Set Up	Coverage
Policy 1	90M
Policy 2	110M
Umbrella π(1 + 2)	200M

- What dependencies impact pricing of a combined umbrella policy?
- Does accumulation of premiums lead to diversification, and does it allow for premium discount?
- What data components are needed to model dependencies?

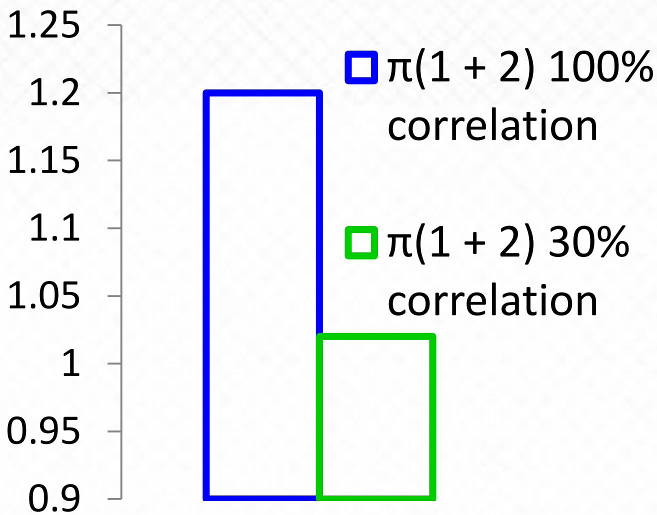
Distance Dependencies and Correlations Have an Impact on Pricing of Accumulated Combined Premium



$$Pure\ Technical\ CAT\ Premium = AAL + k * Standard\ Deviation$$

Policy Set Up	Coverage	Premium
Policy 1: π(1)	90M	512K
Policy 2: π(2)	110M	725K

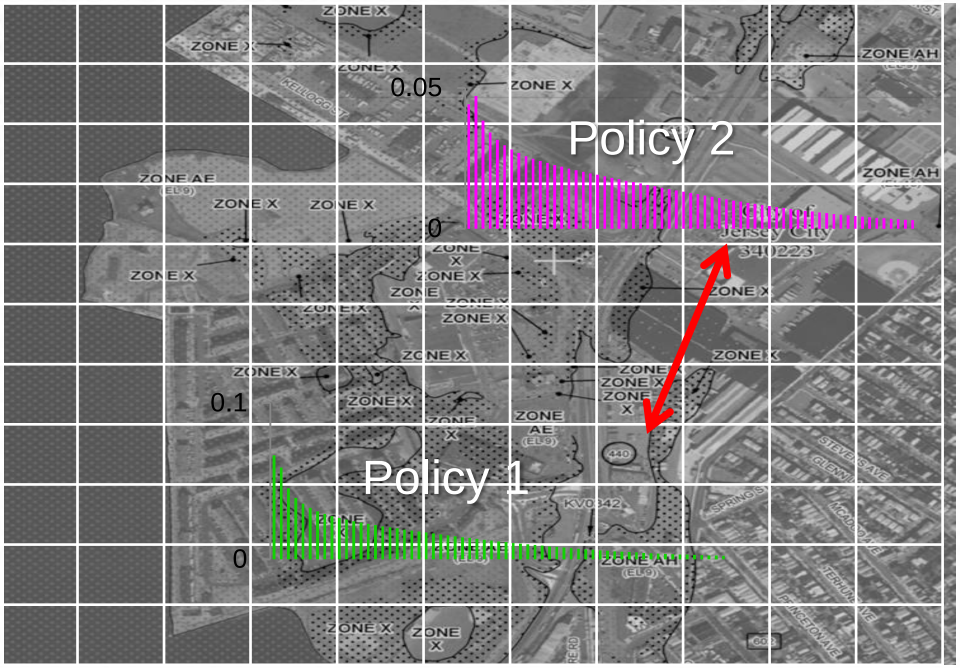
Diversification and Independence as Business Logic for Premium Economies of Scale



Policy Set Up	Premium
Policy 1: π(1)	512K
Policy 2: π(2)	725K
π(1+2): 100% correlation	1.24M
π(1+2): 30% correlation	1.02M

- For **Fully Dependent (100% correlation)** risks in close proximity, the sum of single risk premiums approaches the price of an umbrella product
- For **Less Dependent (30% correlation)** and **Independent** risks, the price of a combined product could be less than the sum of single risk premiums

Big Data Components for Dependencies and Correlations Modeling

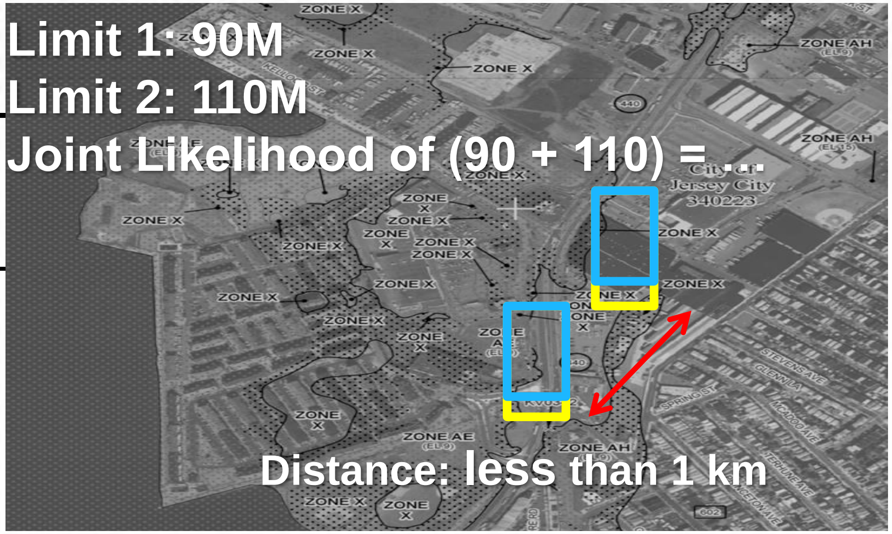


- Geospatial grid application
- Distances between single risks
- Full single risk simulations
- Covariance and correlations

- High resolution geospatial grid application for estimation of distances and dependencies between single risks
- Preserving full single risk 10K simulations for estimation of distance-based covariance matrices

Dependencies in Limits Accumulation: Likelihood of Full Payout

Business Case: What Is the Likelihood of Paying Out the Full Insured Limits



Likelihood of full payout

- Capital reserving impact
- Risk management and MI

Scenario outcomes:

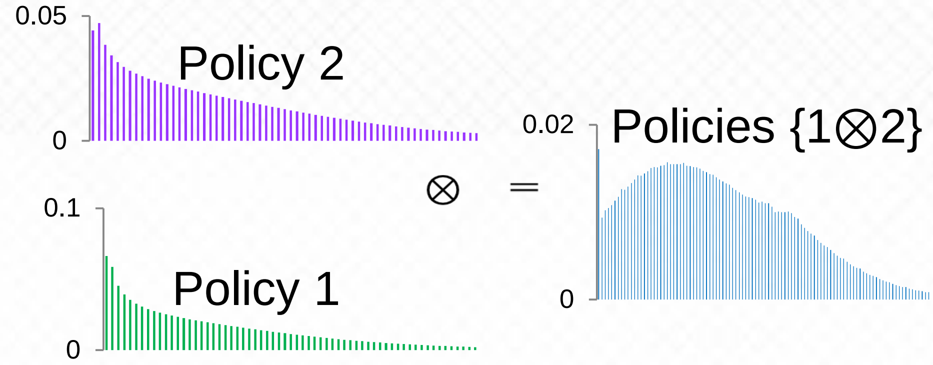
- Dependent policies
- Fully dependent policies

Limits Roll-Up

$$Limit(policy\ 1 + policy\ 2) = Limit(policy\ 1) + Limit(policy\ 2)$$

Data Components for Computing Joint Likelihood Weighted Outcome in Full Payout

Loss accumulation with dependencies

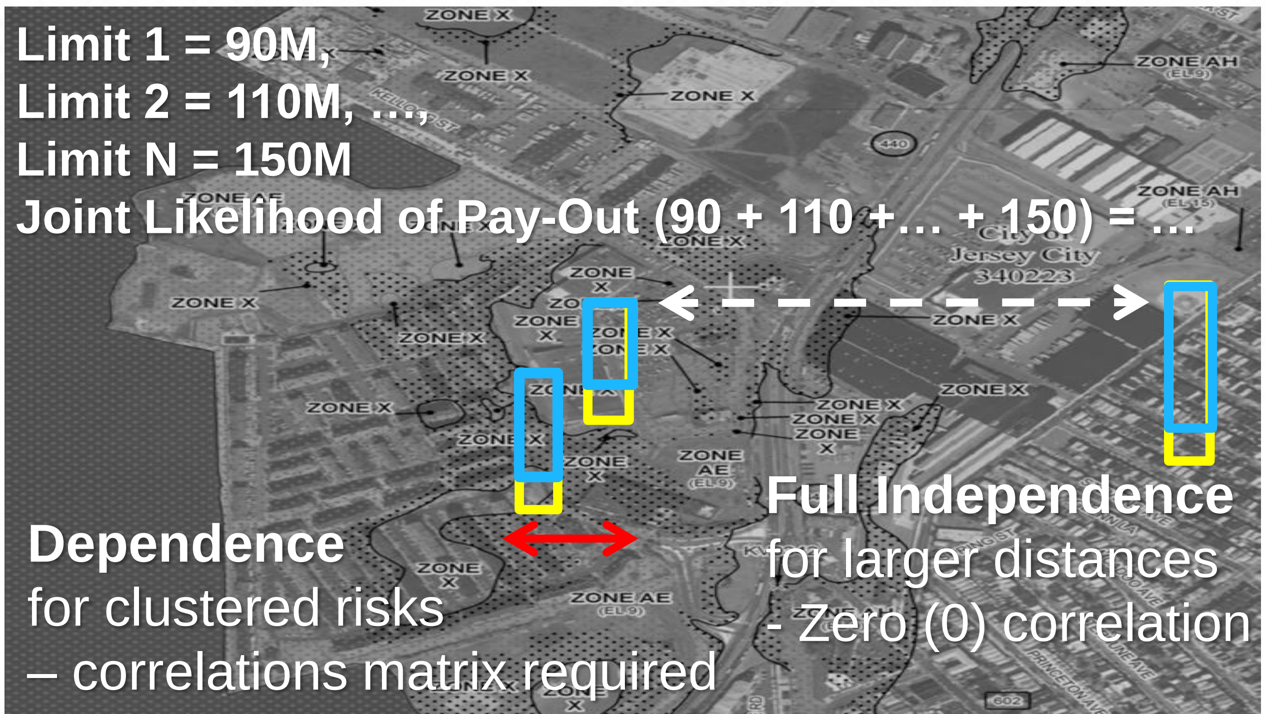


Single and Joint Likelihoods of Full Policy Payout

Limits' Set Up	Limit Value	Likelihood
Limit Policy 1	90M	0.04
Limit Policy 2	110M	0.09
Limit (P1+P2): 100% correlation	200M	0.09
Limit (P1+P2): 30% correlation	200M	0.07
Limit (P1+P2): 0% correlation	200M	0.004

Business Conclusions: Data Components and Tools for Modelers

- Detailed full probabilistic curves at highest risk granularity
- Modeling platform with Big Data capability and fast compute
- Geospatial grid and distance compute functions
- Distance based correlation matrices
- Accurate dependence and accumulation algorithms



Constructing Premium Bounding Intervals: Competitiveness and Cost Savings

The Business Case for Fair Technical Premium Bounds

- Definition and demand for lowest and upper fair and sustainable technical bounds for insurance premium

$$Average\ Historical\ Loss \leq Policy\ Premium \leq Modeled\ VaR_{0.004} \leq Modeled\ TVaR_{0.004}$$

- What data tools are needed for the task?

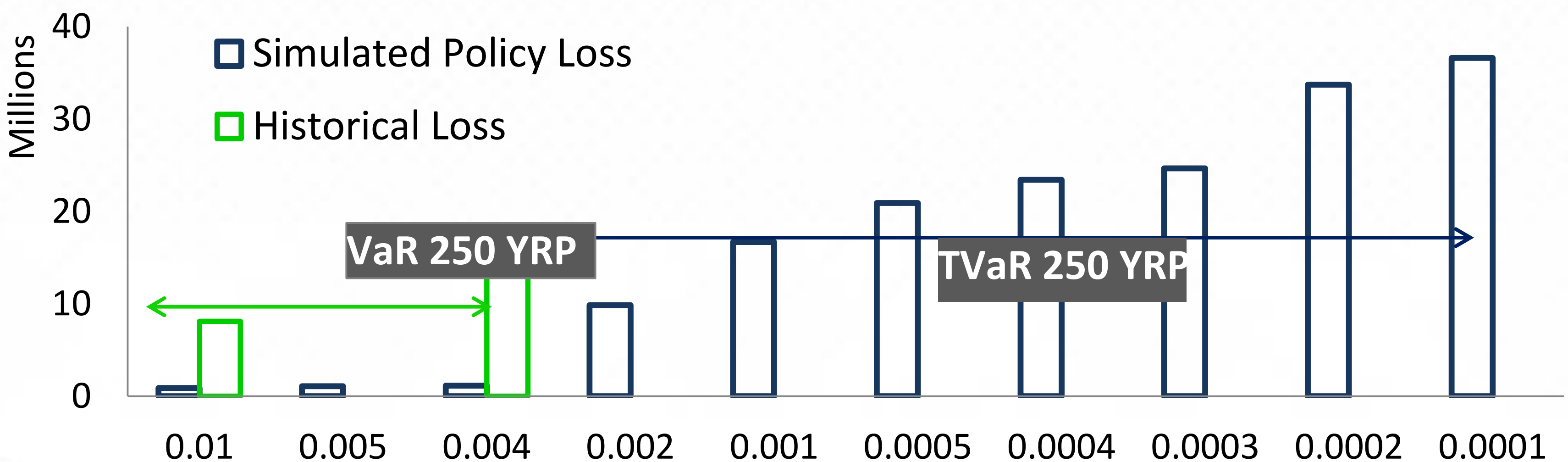
- Desired outcome: long term sustainable premium pricing guidelines

Historical and Simulated Loss: Data Components

CAT Premium 725K

Average Historical Loss	Base Modeled AAL 64K	Risk Load STD. DEV. 661K	VAR 250 YRP 910K	MAX Historical Loss 9M	TVAR 250 YRP 15.6M

$$Average\ Historical\ Loss \leq {\color{red}Policy\ Premium} \leq Modeled\ VaR_{0.004} \leq Modeled\ TVaR_{0.004}$$



Case Conclusions: Functions of Premium Bounds

	CAT Premium i 512K			
Policy 1 90M	Base Modeled AAL 43K	Risk Load STD. DEV. 469K	VAR 250 YRP 870K	TVAR 250 YRP 13.8M
	CAT Premium ii 725K			
Policy 2 110M	Base Modeled AAL 64K	Risk Load STD. DEV. 661K	VAR 250 YRP 910K	TVAR 250 YRP 15.6M

- **Lower Bound** guarantees that economies of scale and diversification do not degrade underwriting discipline
- **Upper Bound** provides a fair technical sustainability of premium at macro level – insureds, insurers and share-holders
- **Risk Ranking** of policies with comparative profiles is possible and recommended with TVaR - a consistent and sub-additive risk measure

Some Final Thoughts and Further Work

Advantages of Big Data & Compute Platforms

- Aggregate and blended inland flood and storm surge flood view of risk
- Tail risk modeling at high granularity and sensitivity to physical environment
- Probabilistic tail risk metrics
- Solving complex actuarial and pricing problems

Further work

- Continuing improvements in computational performance
- Expansion to larger simulations to enhance view of tail risk by high granularity of risk – single policy & location
- Visualization: Fast printing out of loss cost and physical intensity (flood depth) maps for historic and stochastic events
- Transparency in peril and financial modeling operations

Contributors

Many thanks for help and contribution from

Vinu Kuriakose, FCAS, CPCU, MAAA - Liberty Mutual
Cheng Khang Saw, ACAS - Nationwide Insurance

Appendix

AIR

Premium Accumulation and Dependence

Additivity of two premiums:

Notation: π−premium; Q_t, S_t− modeled loss for risk 1 and risk 2; E[.] − expected mean loss σ[.] − standard deviation of expected loss; $COV[S_t : Q_t]$ - covariance

$$\boldsymbol{\pi}(Q_t) + \boldsymbol{\pi}(S_t) = E[S_t] + \sigma[S_t] + E[Q_t] + \sigma[Q_t]$$

$$\boldsymbol{\pi}(Q_t) + \boldsymbol{\pi}(S_t) = E[S_t + Q_t] + \sigma[S_t] + \sigma[Q_t]$$

Dependencies in accumulation of Premium

$$\boldsymbol{\pi}(Q_t) + \boldsymbol{\pi}(S_t) = E[S_t + Q_t] + \sqrt{\sigma^2[S_t] + \sigma^2[Q_t] + 2 * COV[S_t : Q_t]}$$

Dependence, Sub-Additivity and Diversification

- 100% correlation models full dependence and leads to full additivity

$$\boldsymbol{\pi}(Q_t) + \boldsymbol{\pi}(S_t) = E[S_t + Q_t] + \sqrt{\sigma^2[S_t] + \sigma^2[Q_t] + 2 * 100\% * \sigma[S_t] * \sigma[Q_t]}$$

- Less than full dependence (i.e., less that 100% correlation) allows for sub-additivity

$$\boldsymbol{\pi}(Q_t) + \boldsymbol{\pi}(S_t \geqq E[S_t + Q_t] + \sqrt{\sigma^2[S_t] + \sigma^2[Q_t] + 2 * 30\% * \sigma[S_t] * \sigma[Q_t]}$$

- Sub-additive effects for less than full dependence support diversification

$$\boldsymbol{\pi}(Q_t + S_t) \leqq \boldsymbol{\pi}(Q_t) + \boldsymbol{\pi}(S_t)$$

Exposed Limits with Probability Weighted Outcomes

- Independent Joint Probability of Limits Roll-Up

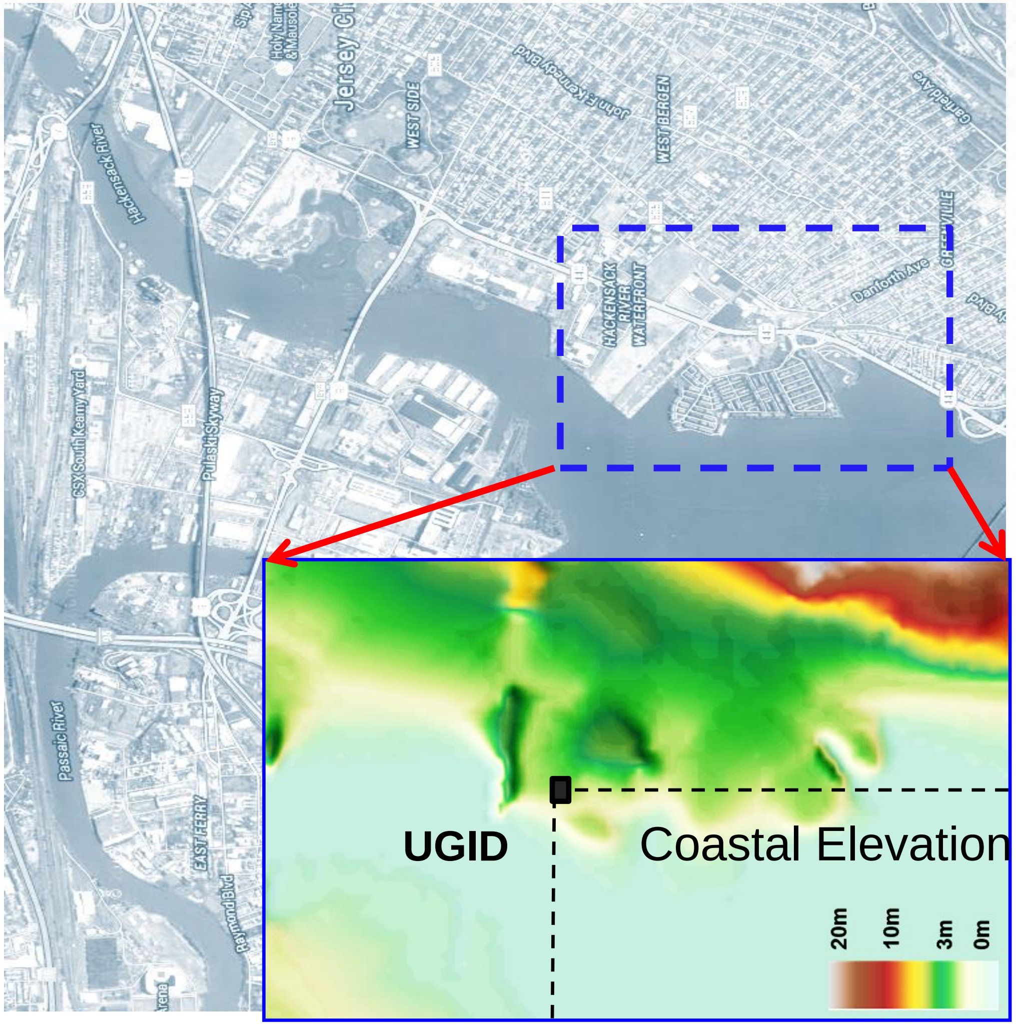
$$p_{1+2}[Lim(S_t + Q_t)] = p_1[Lim(S_t)] * p_2[Lim(Q_t)]$$

- Dependent Joint Probability of Limits Roll-Up

$$p_{1+2}[Lim(P1_t + P2_t)] = f_{P1_t,\, P2_t}\ \{Lim(P1_t) + Lim(P2_t)\}$$

- Critical Knowledge: $f_{P1_t,\, P2_t}$ - the joint distribution function with a dependence structure, described by a Covariance Matrix - $COV[P1{:}\,P2]$

Physical Geospatial Data for Definition of Insurance Policy Premium



- Coastal Elevation
- Distance to Water Body
- Soil Type
- Flood Defenses
- Land Use / Land Cover
- Flood Intensity Grid resolution