Predictive Modeling for Homeowners

David Cummings
Vice President – Research
ISO Innovative Analytics

Opportunities in Predictive Modeling

- **Lessons from Personal Auto**
  - Major innovations in historically static rate plan
  - Increased competition
  - Profitable growth for adopters of advanced analytics
  - Hunger for the next innovation
- **In comparison, much less modeling has been done in Homeowners**
  - Translates into greater opportunity
  - By peril modeling is an important tool

ISO's approach to predictive modeling

- **Highly qualified modeling team**
  - Technical staff has more than 25 advanced degrees in math/statistics/computer science
- **State of the art statistical/data mining approaches**
- **Enabling company customization**
  - Not a "one size fits all" solution
- **De-mystifying the "black box"**

ISO Risk Analyzer® - Homeowners Framework

Traditional Rating Plan

- Territory
- State
- Construction
- Protection
- Amount of Ins

- Prior Claims
- Demographics
- Credit

New By Peril Rating

Environmental Module

Risk Characteristics

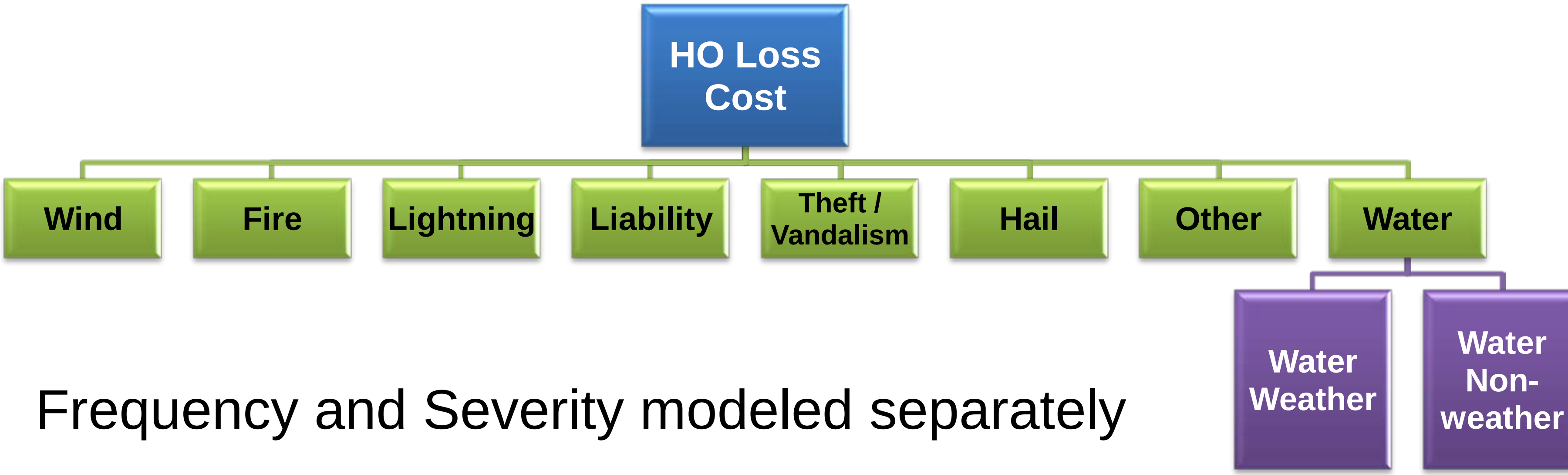
Human Factors

Total Policy Risk

Interactions of all indicators

Features of the Model

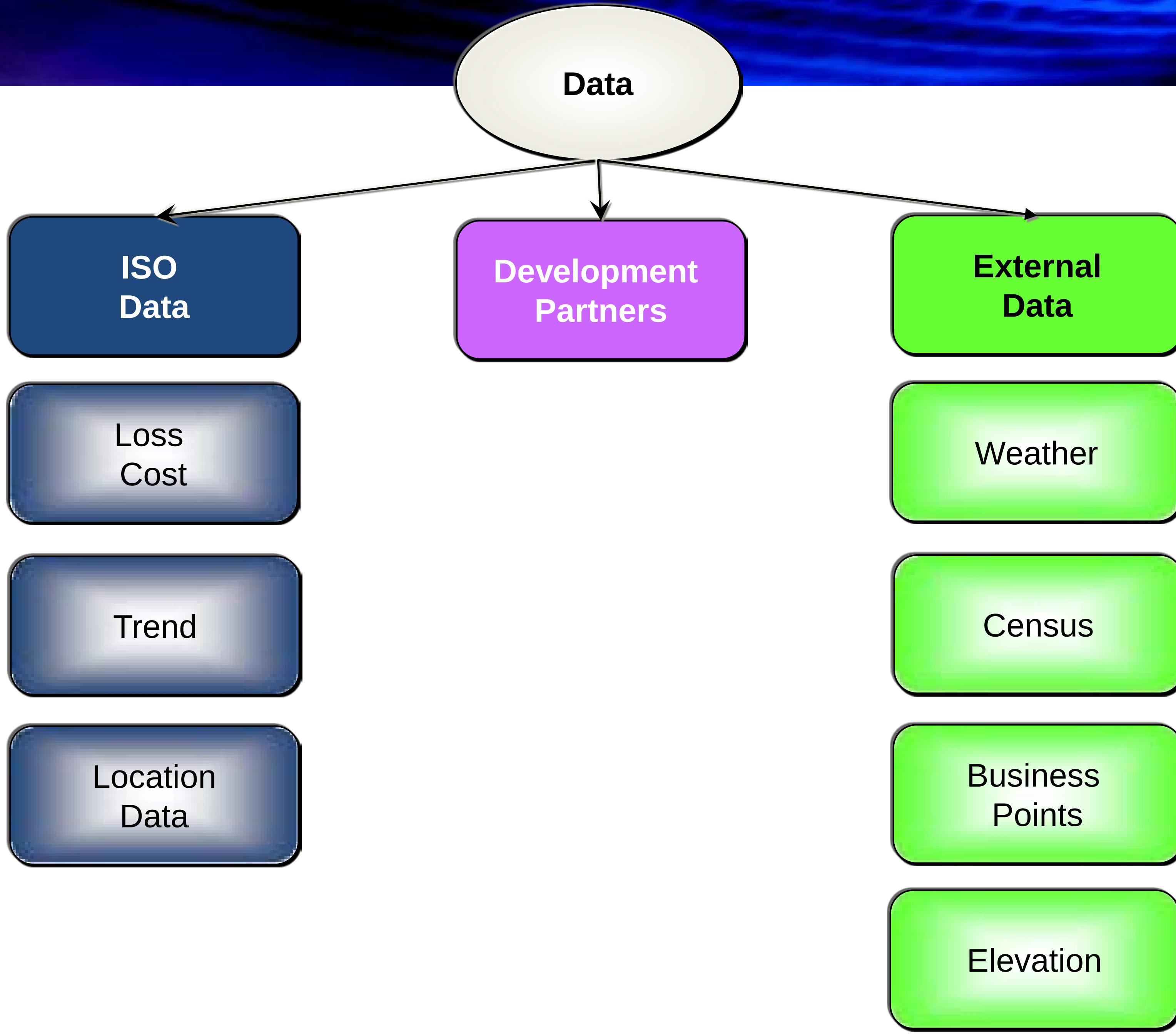
- Modeled by peril (excluding hurricane)



- Frequency and Severity modeled separately

- Combine to form ‘all peril loss cost’ – multiplied frequency and severity – added across perils

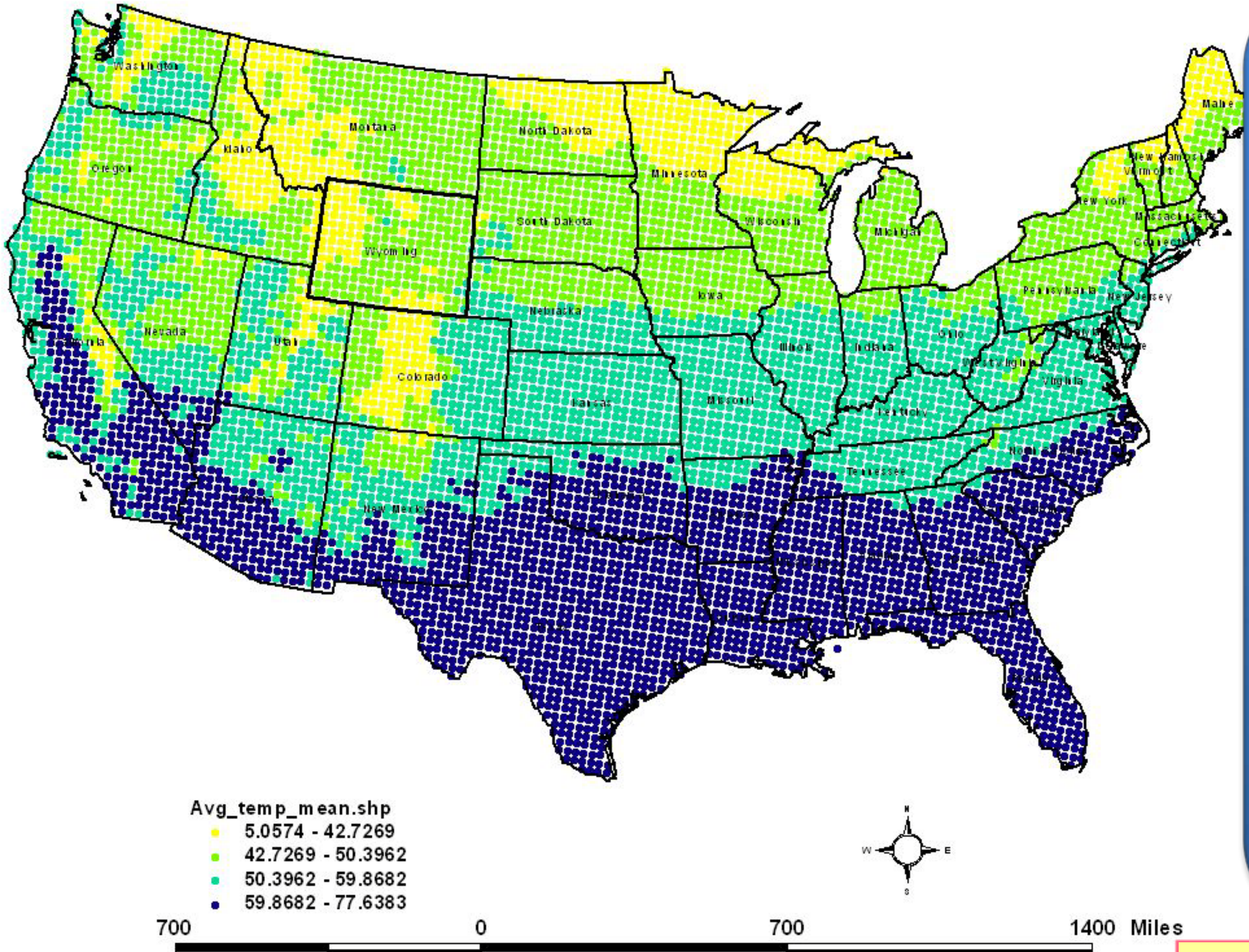
The Environment is the Exposure


Data
ISO Data
Development Partners
External Data
Loss Cost
Trend
Location Data
Weather
Census
Business Points
Elevation

Modeling Techniques Employed

- Variable Selection – univariate analysis, transformations, known relationship to loss
- Sampling
- Regression / general linear modeling
- Sub models/data reduction – splines, principal component analysis, variable clustering
- Spatial Smoothing

External Data – Weather



Mean of daily average temperature

Source:
North America
Regional Reanalysis

Length:
27 years of data
(1979 -2005)
8 daily readings

Resolution:
32 x 32 km
Interpolated using 4 nearest grid centroids (weights = inverse distance)

2 person-years work

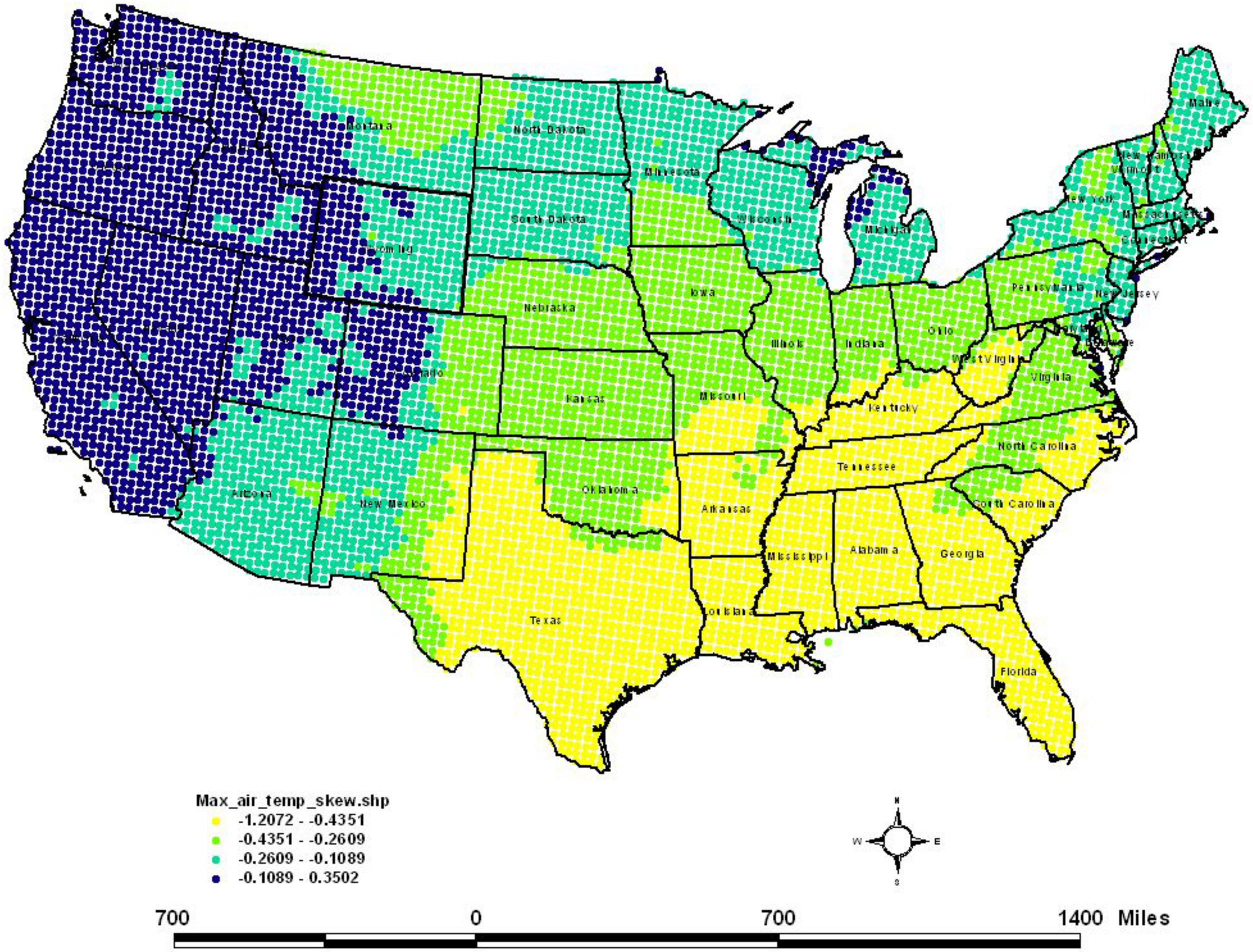
Mean of daily average temperature in the last 27 years

External Data – Weather
Derive Novel Data Features

(Indicators, daily, consecutive days, number of days)

- **Temperature**
  - Below freezing / High temperatures
  - Variations / Average / min / max / deviation
- **Precipitation, Wind and Snow**
  - With / Without
  - Average / min / max / deviation
- **Interactions**
  - Weight of snow (snow + temp)
  - Ice (rain + temp)
  - Fire (no rain, high temp + high wind)
  - Blizzards (snow + wind)

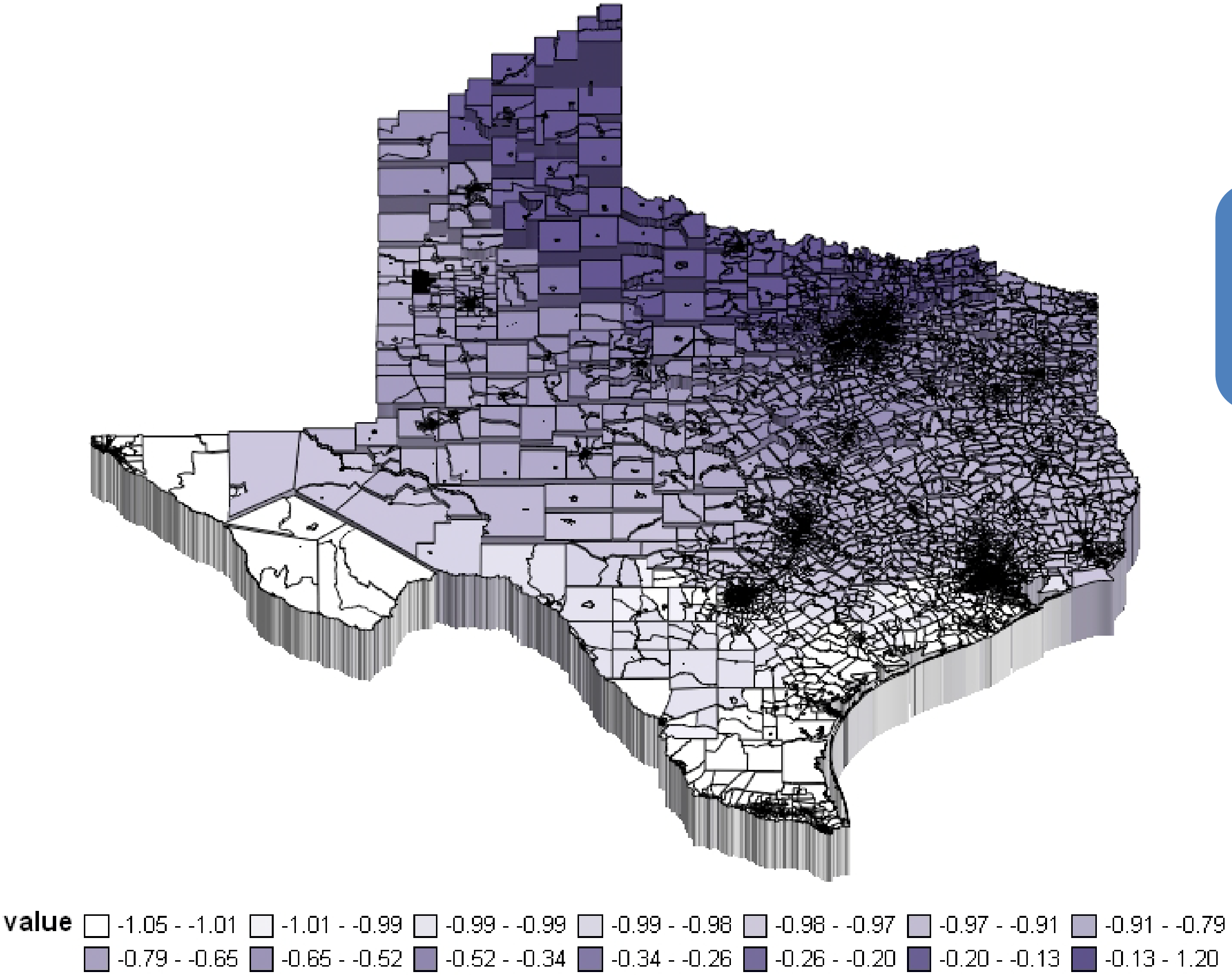
External Data – Weather



Skewness of high air temperature

Visualizing of Weather Interactions

% of days with High < 32 and % of days with Low > 72 (Texas)



Positive coefficient in Wind Frequency model

Using SAS/Graph

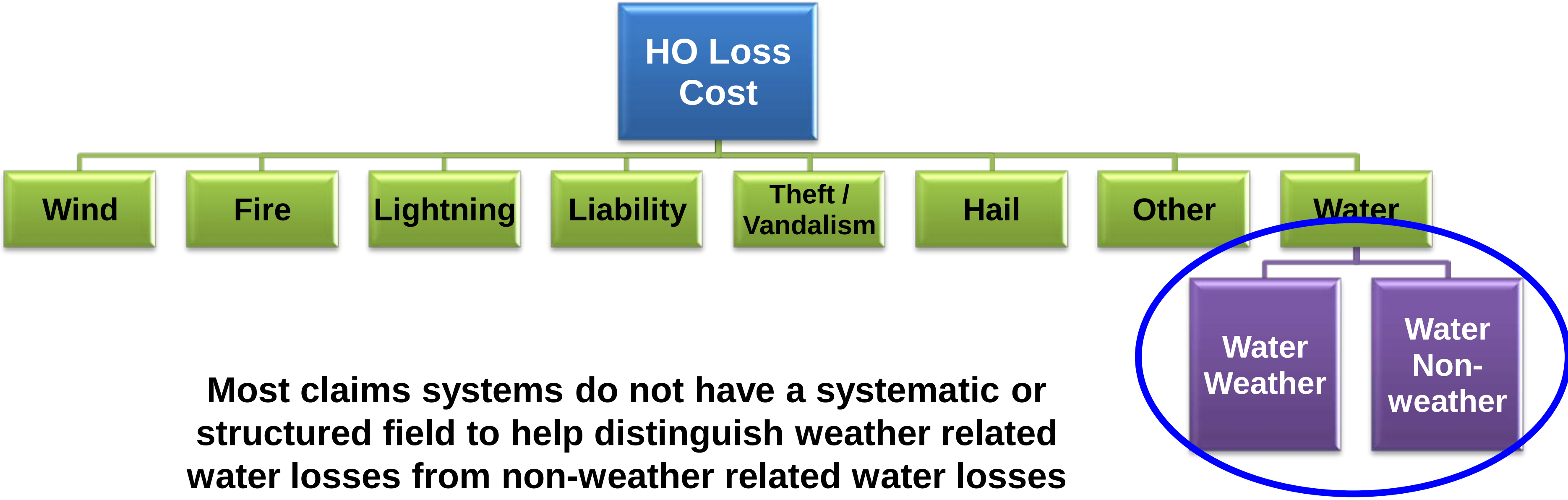
By-Peril Modeling – Serendipitous Discoveries

Weather & Elevation	FIRE	LIGHT	WIND	HAIL	WW	LIAB	THEFT
Elevation							
Temperature							
Precipitation							
Relative Humidity							
Snow							
Wind							
Ice Pellets							

External Validation:

Ellen Cohn. "Weather and Crime". *The British Journal of Criminology* 30:51-64 (1990)

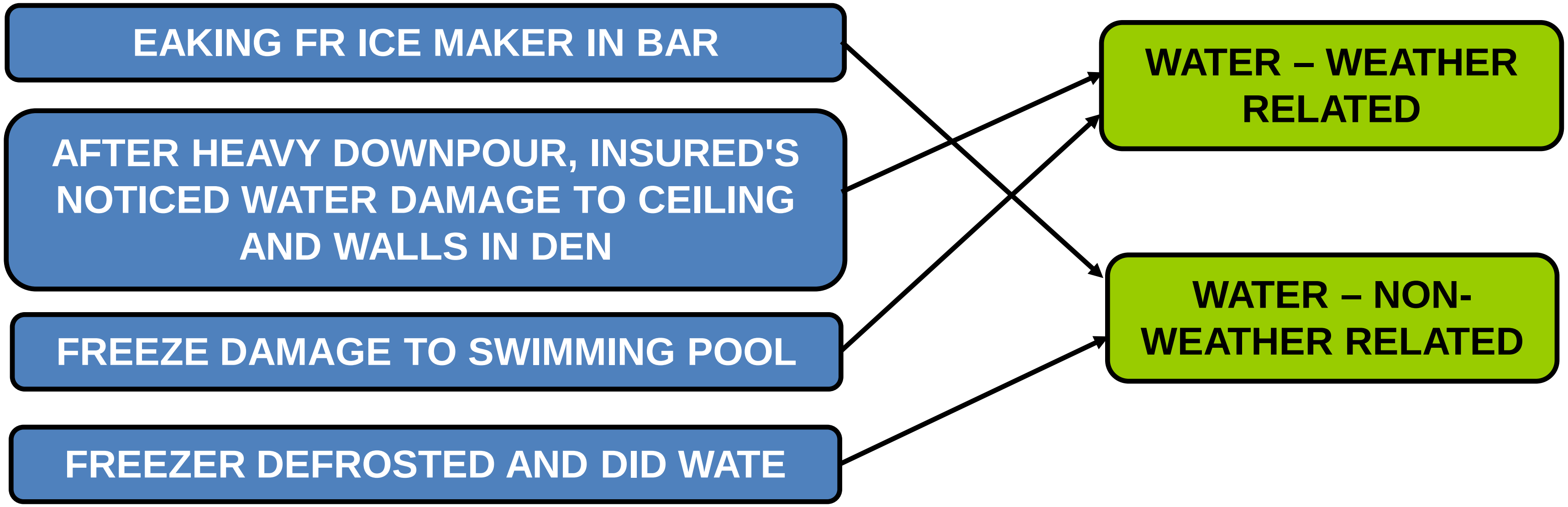
Decomposing Water Losses



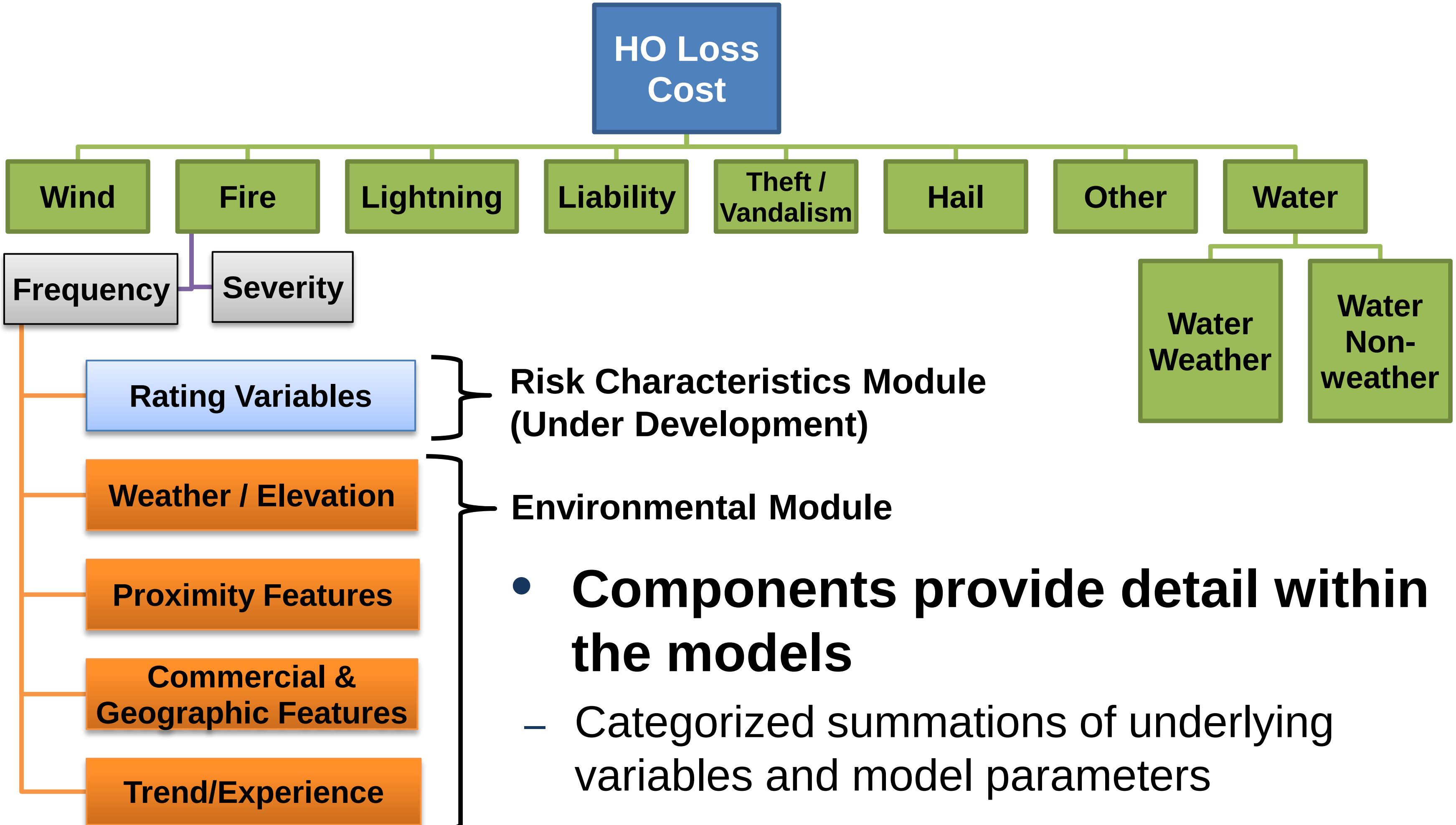
Most claims systems do not have a systematic or structured field to help distinguish weather related water losses from non-weather related water losses

Text Mining for Cause-Of-Loss

- Rich information buried in Unstructured data, such as Loss Descriptions or Adjuster Notes
- E.g., Extracting the "Type of Loss" from the Loss Description



Components



- **Components provide detail within the models**
  - Categorized summations of underlying variables and model parameters
- **Enables Customization**
  - Short circuiting the variable selection process

Example of Variables in Components

- **Unique for each peril model (freq/severity)**
  - **Weather / Elevation:**
    - Elevation
    - Measures of Precipitation
    - Measures of Humidity
    - Measures of Temperature
    - Measures of Wind
  - **Proximity:**
    - Commuting patterns
    - Population variables
    - Public Protection Class
  - **Commercial & Geographic Features:**
    - Distance to coast
    - Distance to major body of water
    - Local concentration of types of businesses (i.e. shopping centers)
  - **Trend / Experience**
    - Peril's proportion of ISO Loss Cost
    - Trend
    - Base Level parameters for:
      - HO Form
      - Construction type
      - Amount of insurance
      - Liability amount
      - Deductible amount
      - Wind and hail deductible
      - Construction age

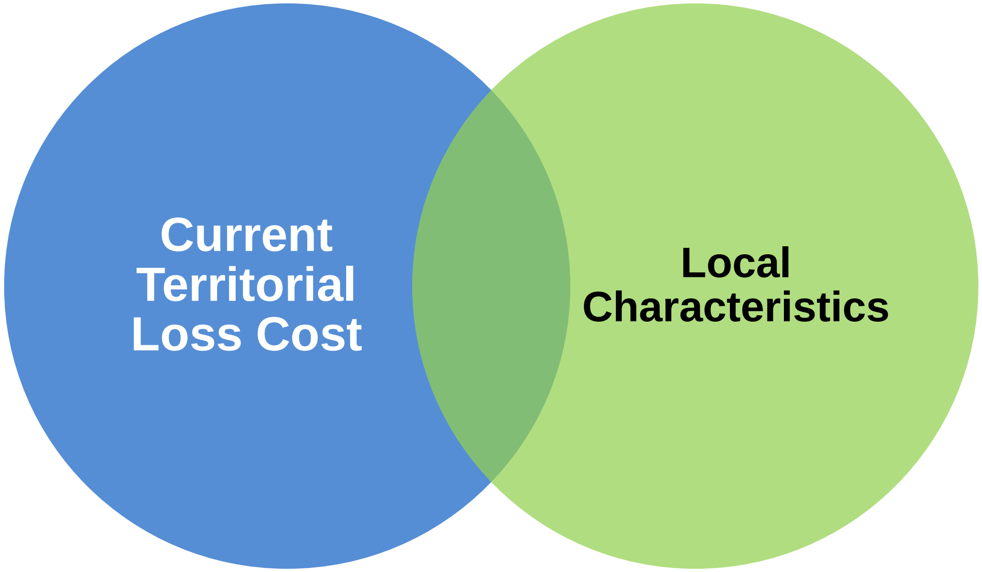
Risk Characteristics Module (Under Development)

Improving Accuracy by Combining Geographic Ratemaking Methods

- **Use traditional territorial loss cost as predictor variable in models**
  - Enables model to capture effects not identified by other predictor variables
  - Helps to “true up” model predictions with traditional estimates
- **Need to be aware that some effects of predictor variables may already be embedded in current territory loss costs**

Improving Accuracy by Combining Geographic Ratemaking Methods

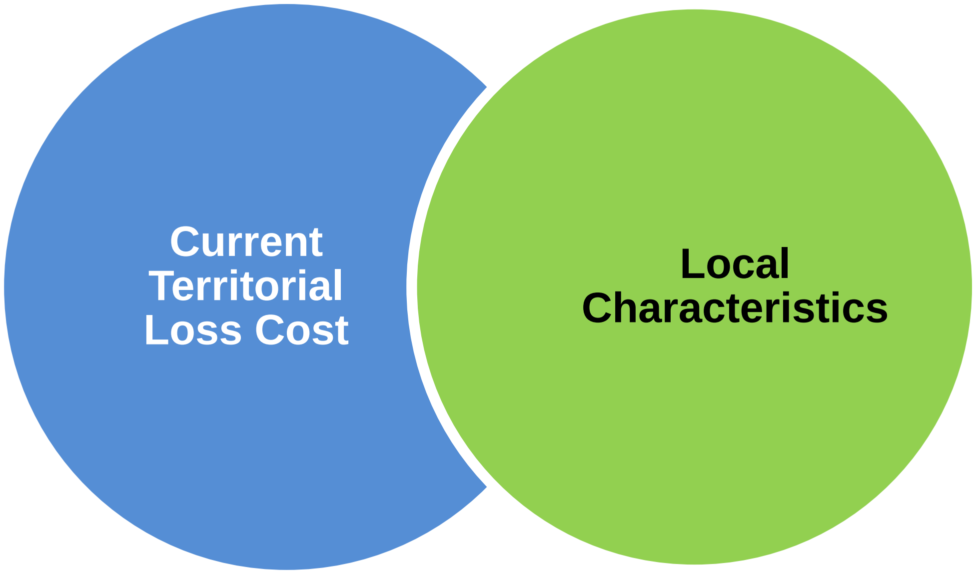
- **Shared Predictive Effects**



- **Multivariate methods can address the overlap without double counting**

Improving Accuracy by Combining Geographic Ratemaking Methods

- Separated Predictive Effects – Same Prediction

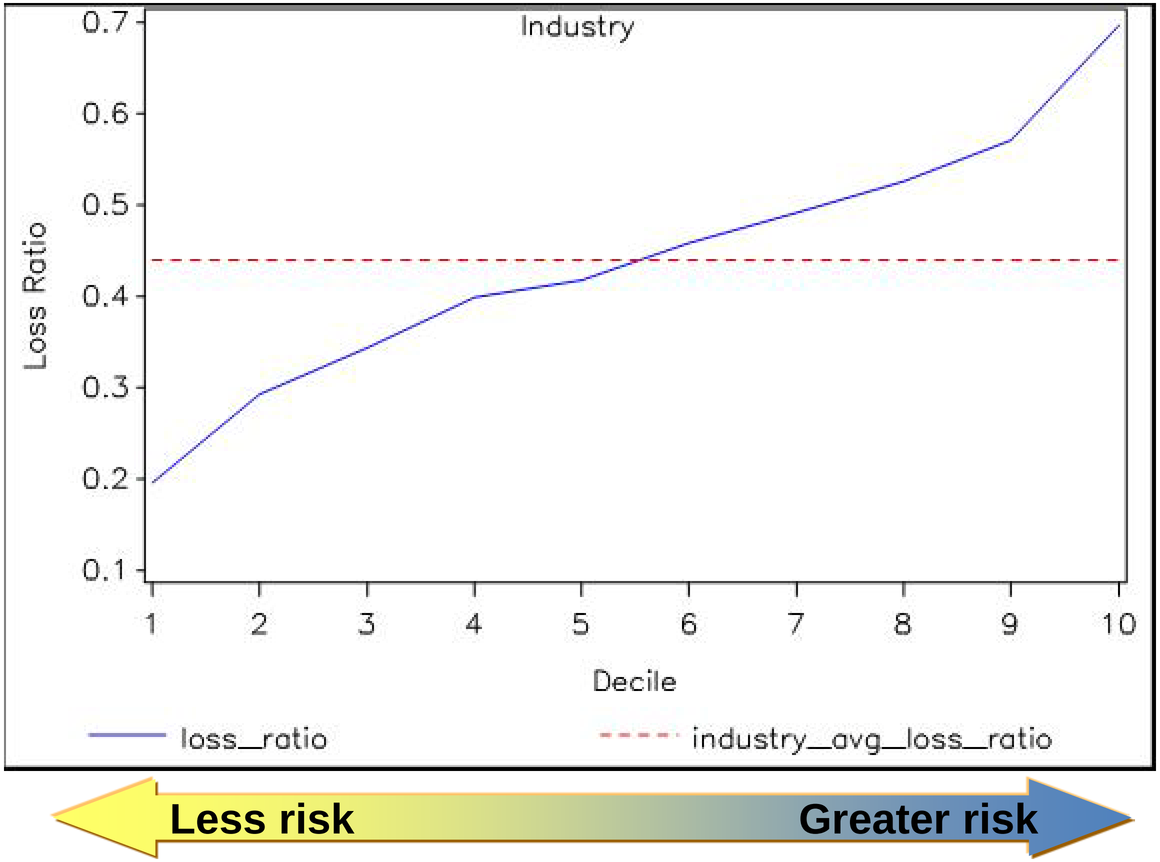


- Estimate the portion of current loss cost not explained by other predictors
- Use “Loss Cost Residual” as predictor

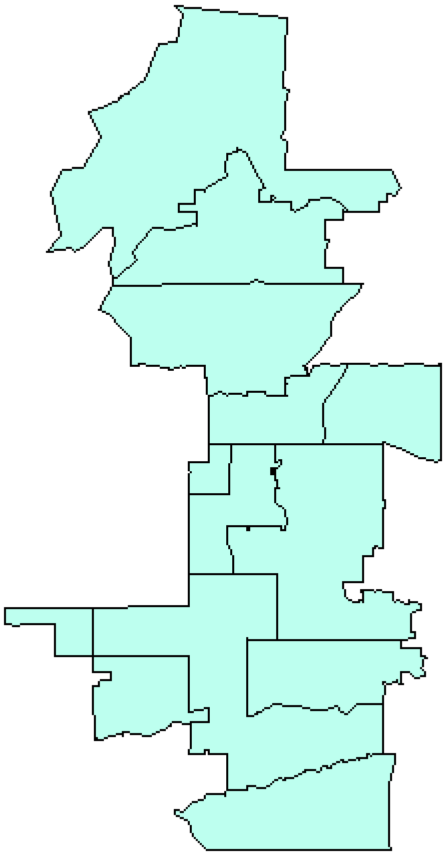
Model Testing

- **Validation of model performance on hold-out dataset**
- **Look at results on maps**
- **Statistical reports to quantify the effect of changes**
  - Examine adjacent loss cost differences
  - Compare to current territorial base rates
  - Examine largest changes from current loss costs
- **External review**

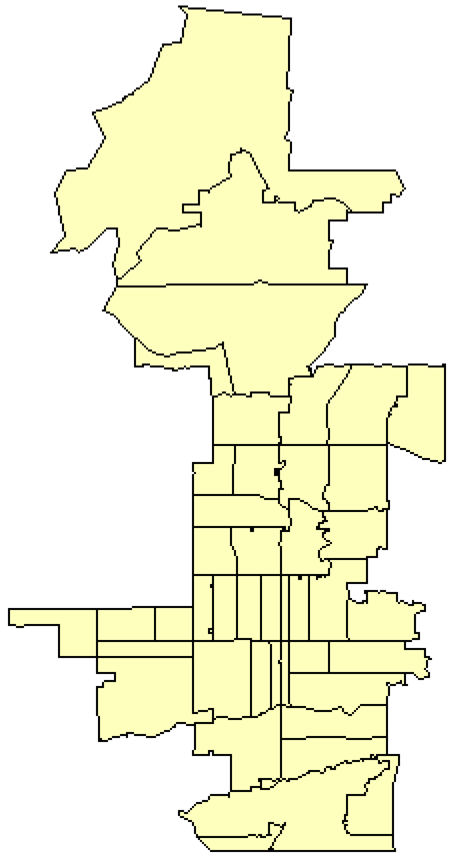
Industry Total Loss Cost
Loss Ratio by Premium Decile



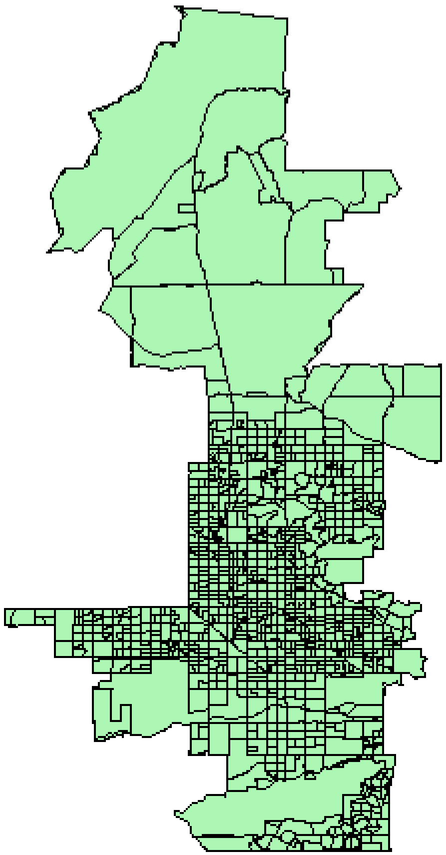
Phoenix, AZ Geographic Area



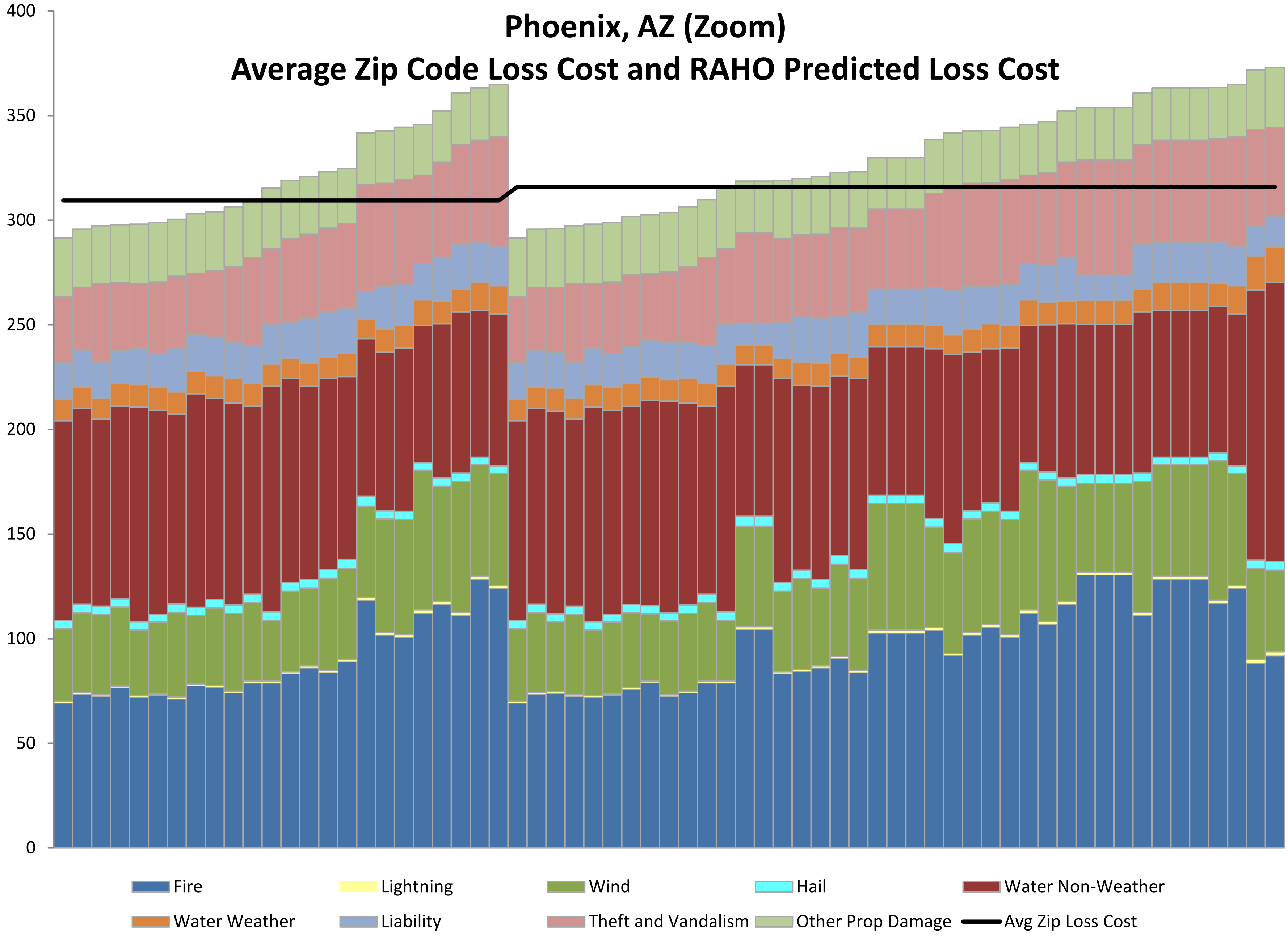
ISO Territories: 9



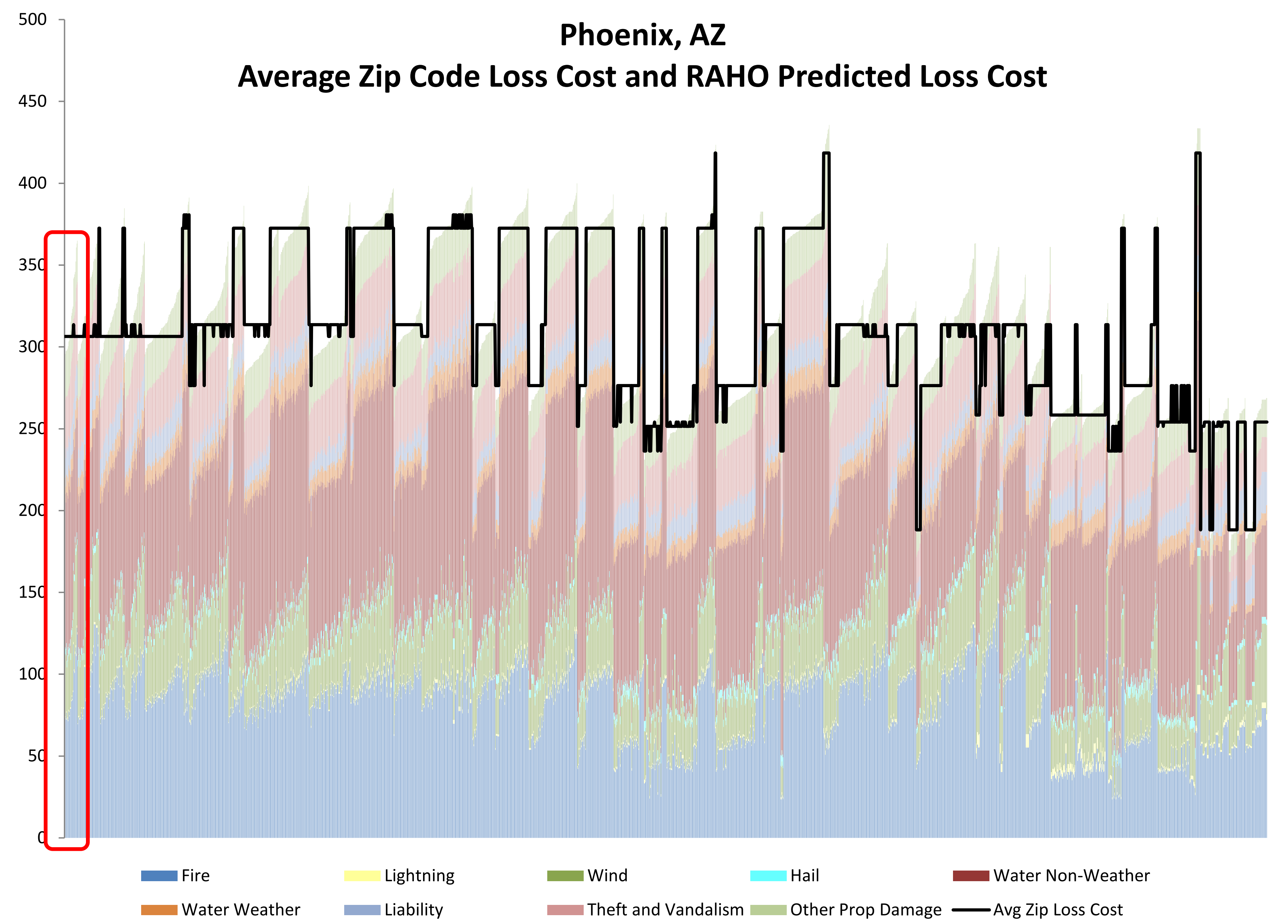
Zip Codes: 80



RAHO: 1309

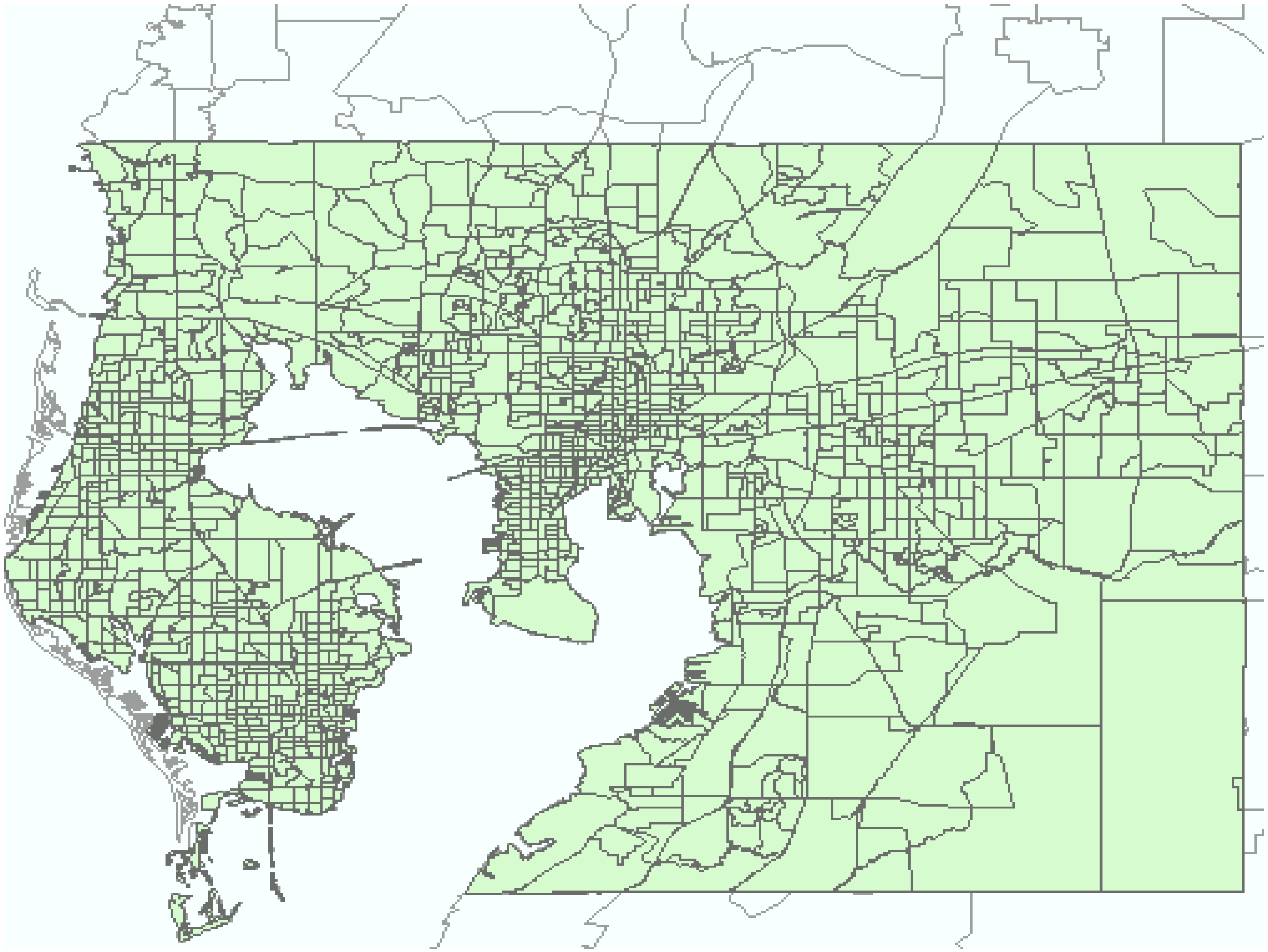


* Loss cost are calculated @ Territory Representative Risk

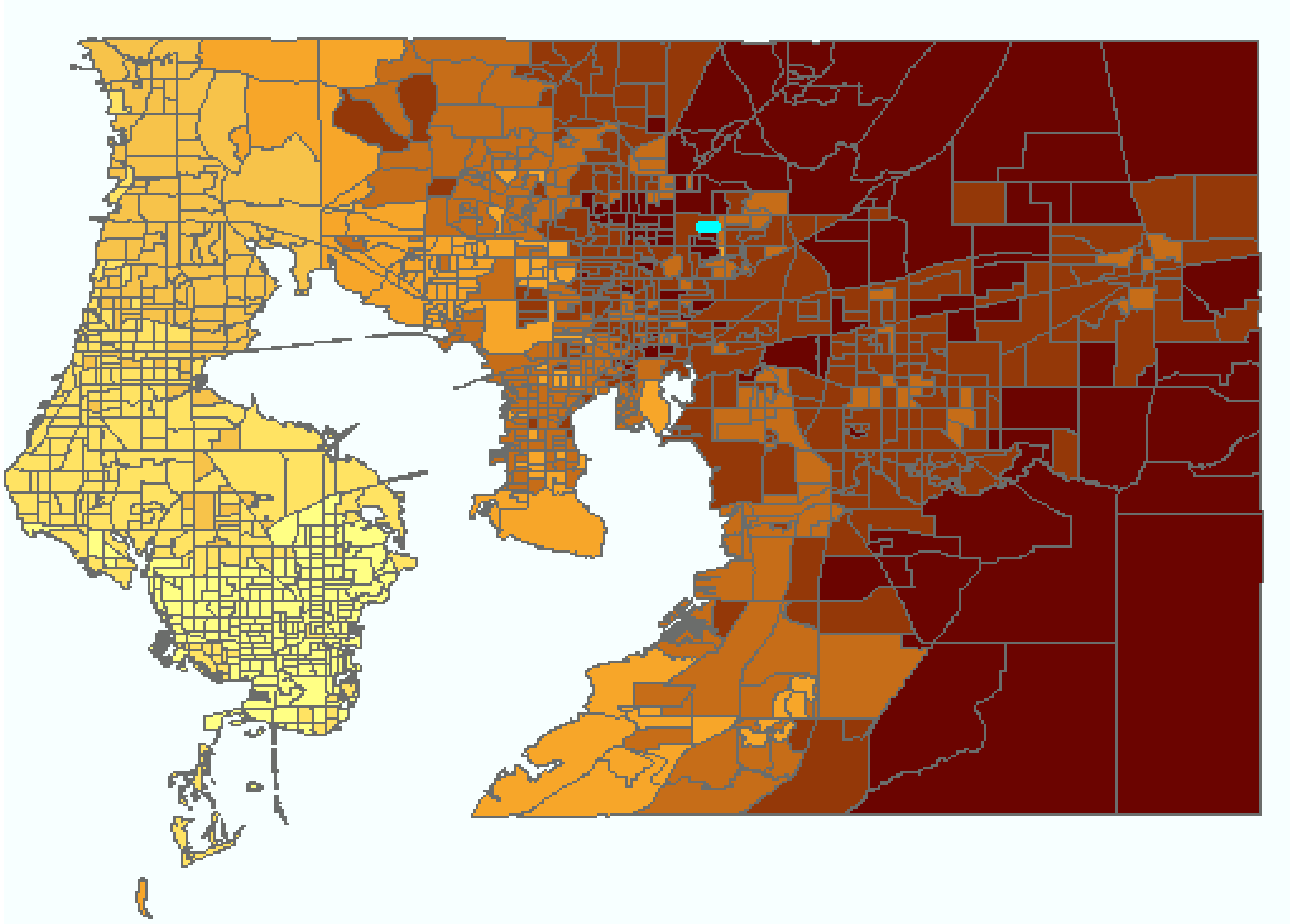


*** Loss cost are calculated @ Territory Representative Risk**

Tampa Bay, FL Area



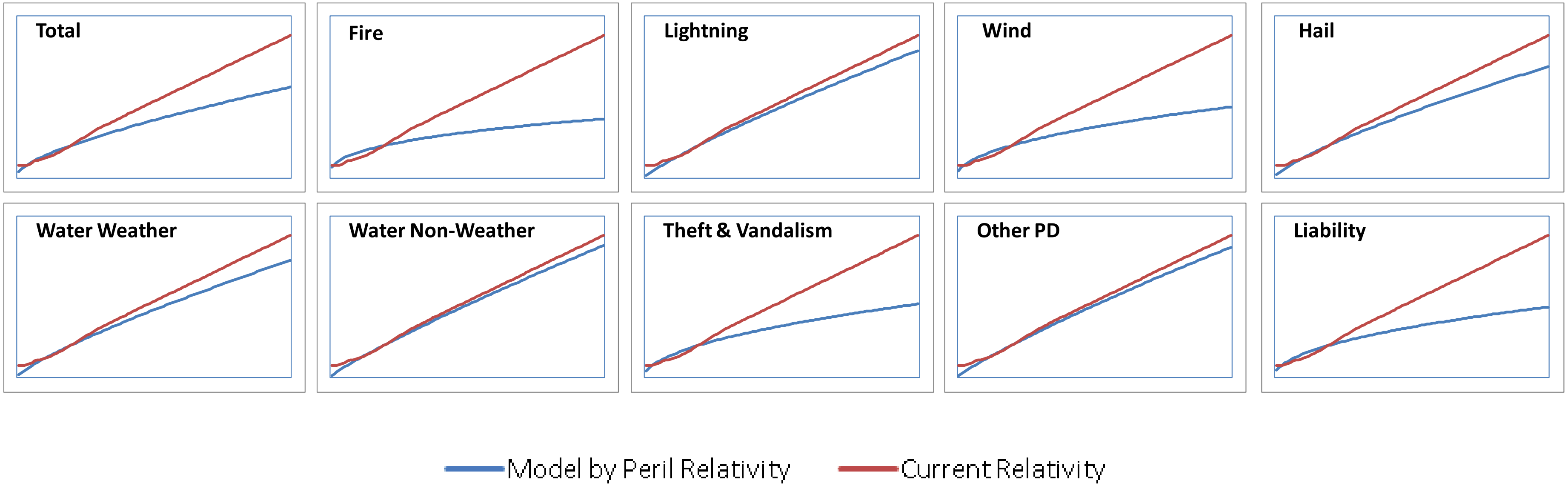
Tampa Bay Area
Detailed Loss Costs (Non-Hurricane)



Opportunities for Enhanced Segmentation

- **Use sum-of-peril loss cost estimates**
  - Build new territories
  - Refine existing territories
- **Use peril-specific models to break apart all-peril rating**
  - Geographic exposures and rating variables
- **Using components as input to models**
  - Incorporate new predictive data with simpler sourcing, preparing, and selecting of variables

Rating Variable Impact by Peril



- **Significant variation by peril**
- **Enhanced accuracy of loss prediction**

ISO

Rating Variable Relativities by Peril

- **Relativities that vary by peril provide lift**
- **Adds accuracy and complexity**
  - All-peril relativities can be derived from peril-based relativities according to peril mix within the area
  - Local Prediction by peril results in varying peril loss costs at the address level
- **Effectively produces all-peril amount relativities that vary at the address level**

Questions?

David Cummings
dcummings@iso.com