Predictive Modeling for Homeowners

David Cummings
VP & Chief Actuary
ISO Innovative Analytics

ISO Innovative Analytics

1

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

2

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

3

ISO Risk Analyzer® - Homeowners Framework

Traditional Rating Plan **New By Peril Rating**

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

- Prior Claims
- Demographics
- Credit

Environmental Module

Rating Factors

Building Characteristics

Occupant

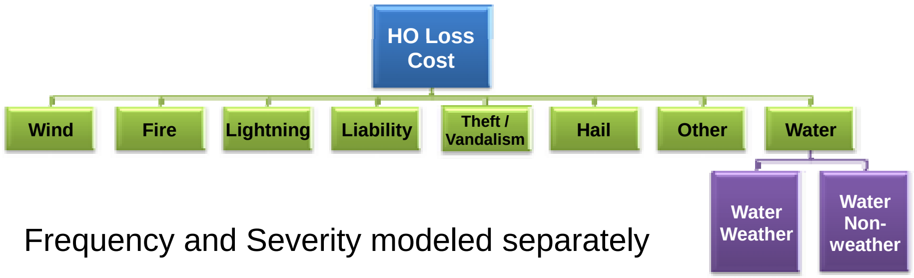
Total Policy Risk
Interactions of all indicators

ISO

4

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
- Rating factors from Risk Analyzer used to modify the loss costs by peril to account for the effect of amount of insurance, deductible and age of construction.

5

The Environment is the Exposure

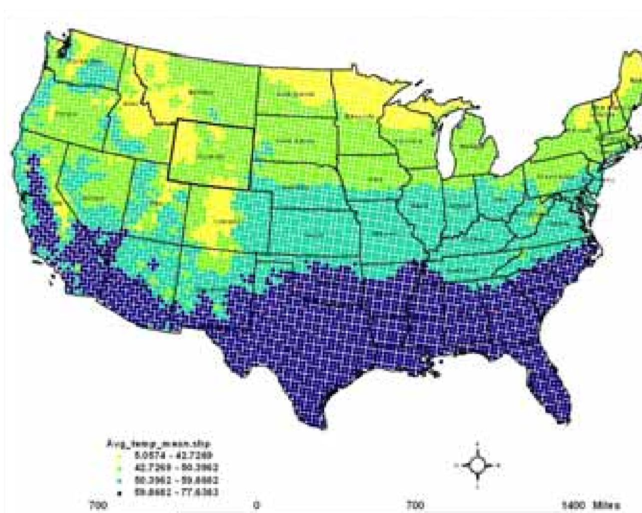
Rating Factors complement the environment predictions

6

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

Mean of daily average temperature in the last 27 years

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)

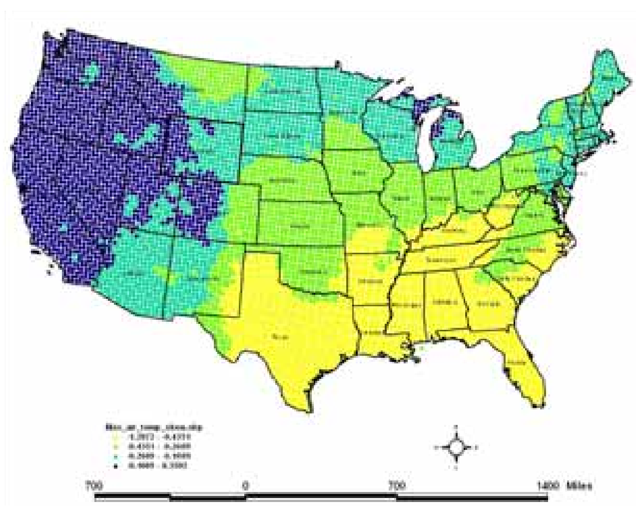
2 person-years work

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)

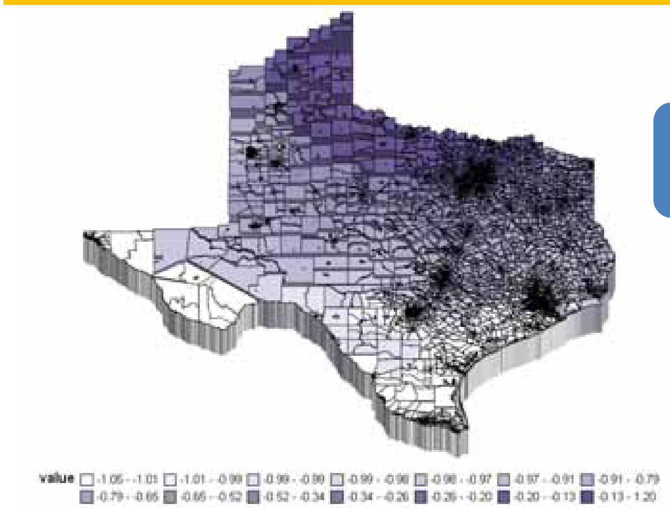
External Data – Weather



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

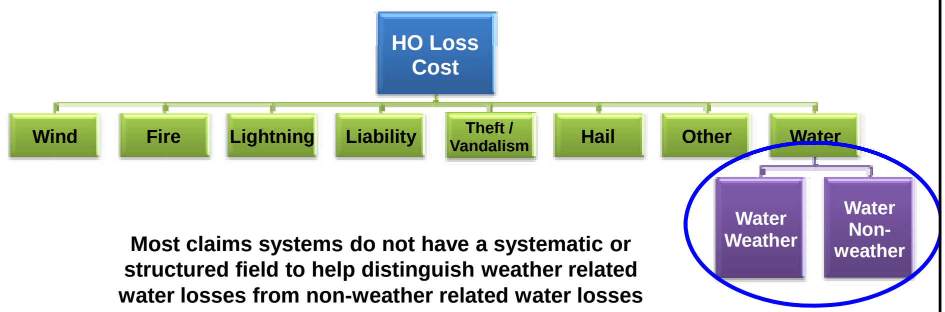
Using SAS/Graph

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)

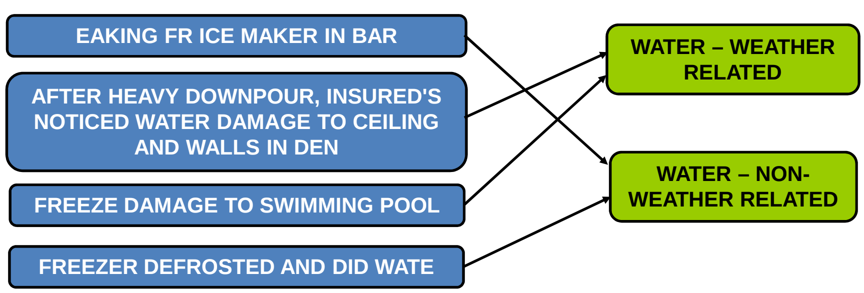
Decomposing Water Losses



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



Public Protection Class (PPC)

- Derived from detailed review of local fire protection capabilities
- Applies within fire district boundaries, plus considerations of available water supply and fire station distance
- By-Peril Modeling allows PPC to be used differently than current Loss Costs

Current ISO Loss Costs

- Single factor applies to all-perils loss cost
- Only geographic refinement below Territory

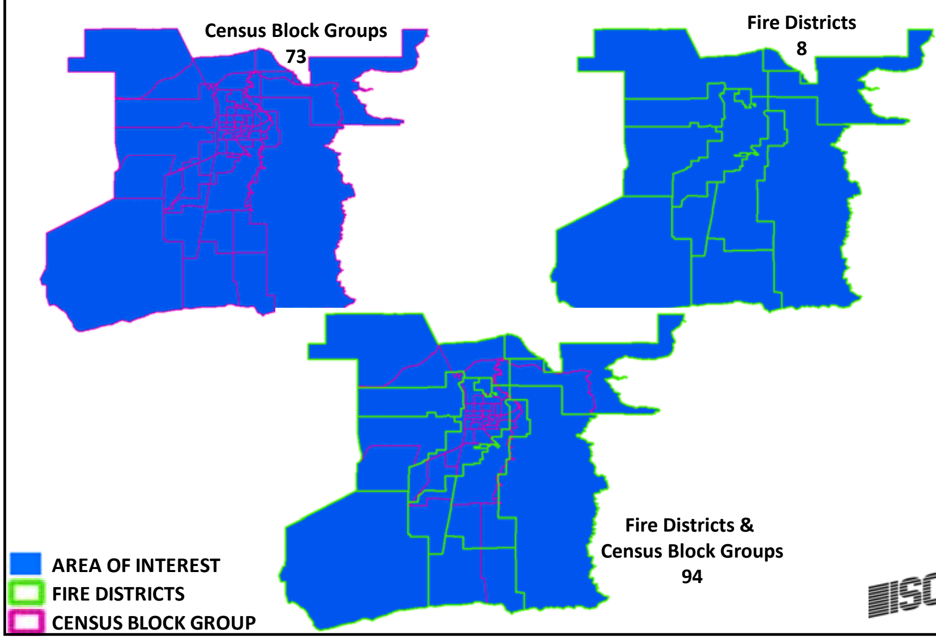
By-Peril Modeling

- Input variable in peril models
- Applies to perils where statistically significant
- Multivariate analysis with other geographic variables

ISO

15

Geographic Units

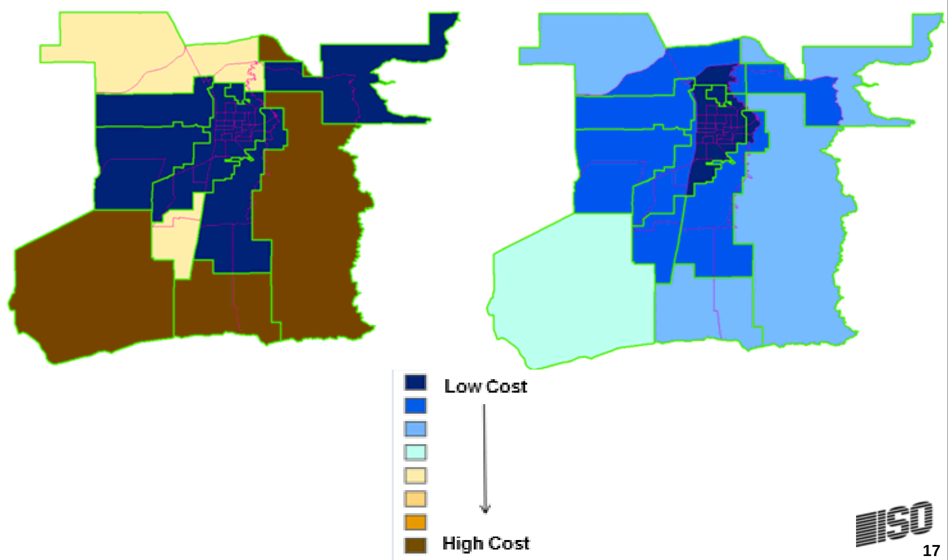


ISO

16

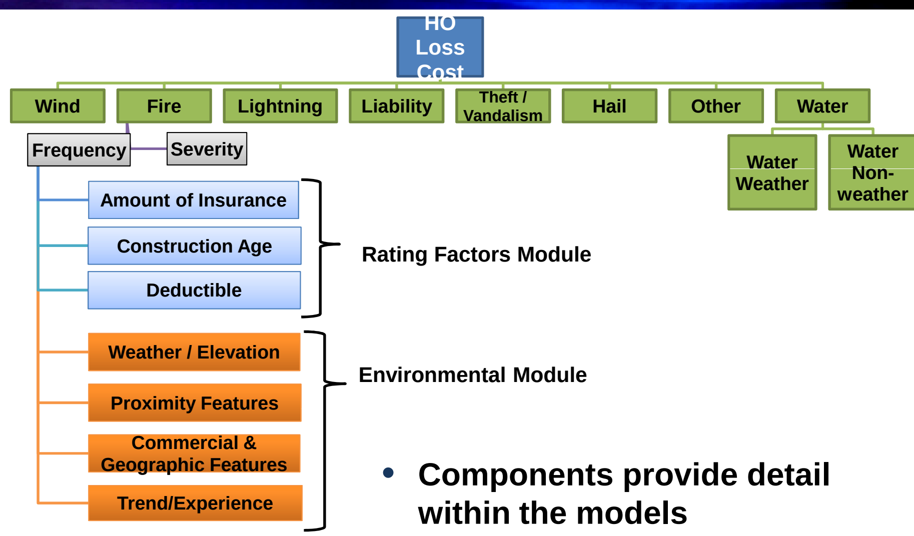
ISO
TOTAL LOSS COST
WITH PPC

BY-PERIL MODEL
TOTAL LOSS COST
WITH PPC



17

Components



- **Components provide detail within the models**
  - Categorized summations of underlying variables and model parameters

18

Dealing with Data for By-Peril Modeling

- **Accurate by-peril Homeowners models require extensive data resources**
  - Low frequency line – split further by peril
  - Severity is volatile and differs significantly by peril
- **Components create re-usable data features**
  - Derived from modeling on larger datasets
  - Can be used directly as inputs into models on smaller datasets – Ensuring stable results without overfitting
- **Components enable efficient modeling**
  - Customized lift while short circuiting variable selection

ISO
19

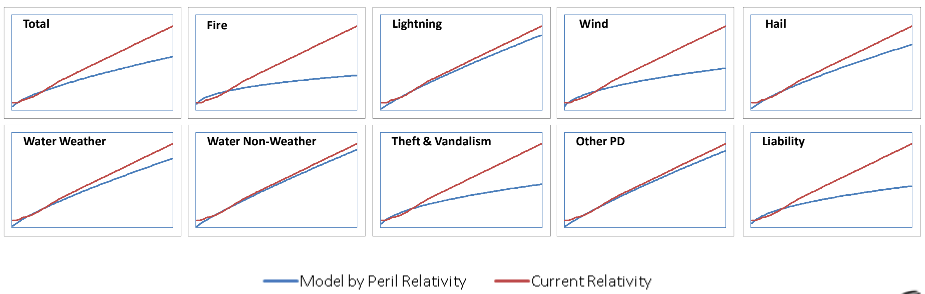
Example of Variables
Environmental Components

- **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
      - Liability amount

ISO
20

By-Peril Rating Factors

- **Modeled simultaneously with geographic variables**
  - Amount of Insurance
  - Deductible
  - Age of Construction
- **Produces a set of countrywide tables by peril for each rating factor**



ISO

By-Peril Rating Factors + Environmental Factors

- **Why are by peril rating factors more accurate?**
  - By-peril rating factors allow for a more explicit recognition of the impact of perils *varying by location*
  - By-peril rating factors more dynamically react to changing peril contributions over time

Peril	Amount of Insurance Factor	Location A	Location B	Location C
Fire	1.5	30%	25%	50%
Wind	1.2	20%	25%	15%
Water	1.0	40%	25%	20%
Other	2.0	10%	25%	15%
All-Perils Factor	1.37	1.29	1.43	1.39

ISO

By-Peril Rating Factors + Environmental Factors

- **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 rating factor relativities that vary at the address level**

ISO

23

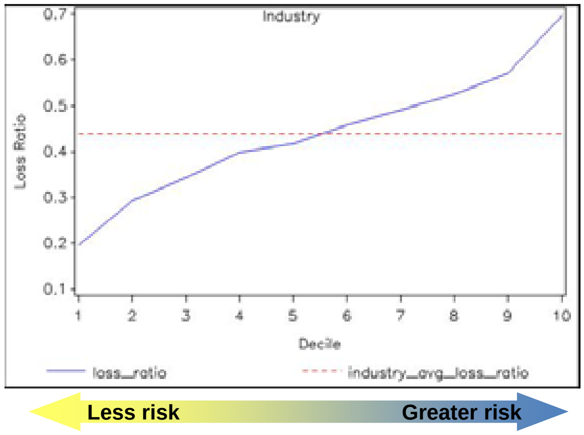
Model Testing

- **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**

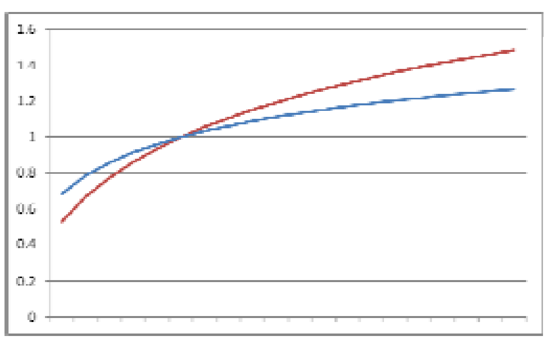
ISO

24

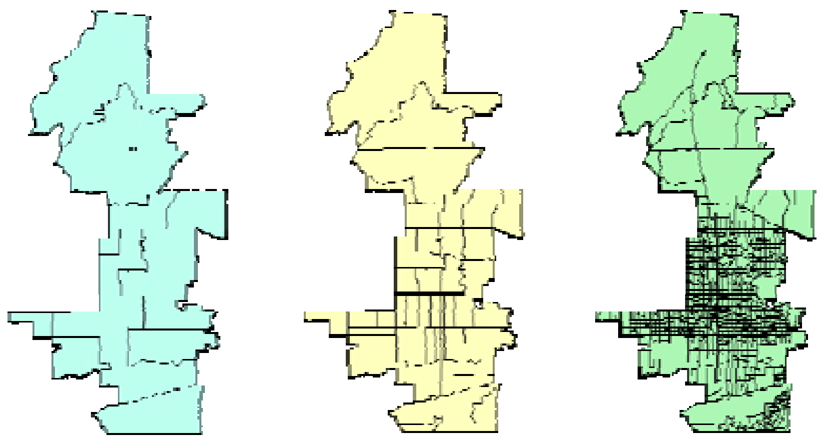
Industry Total Loss Cost
Loss Ratio by Premium Decile



Using Components to gain customized lift

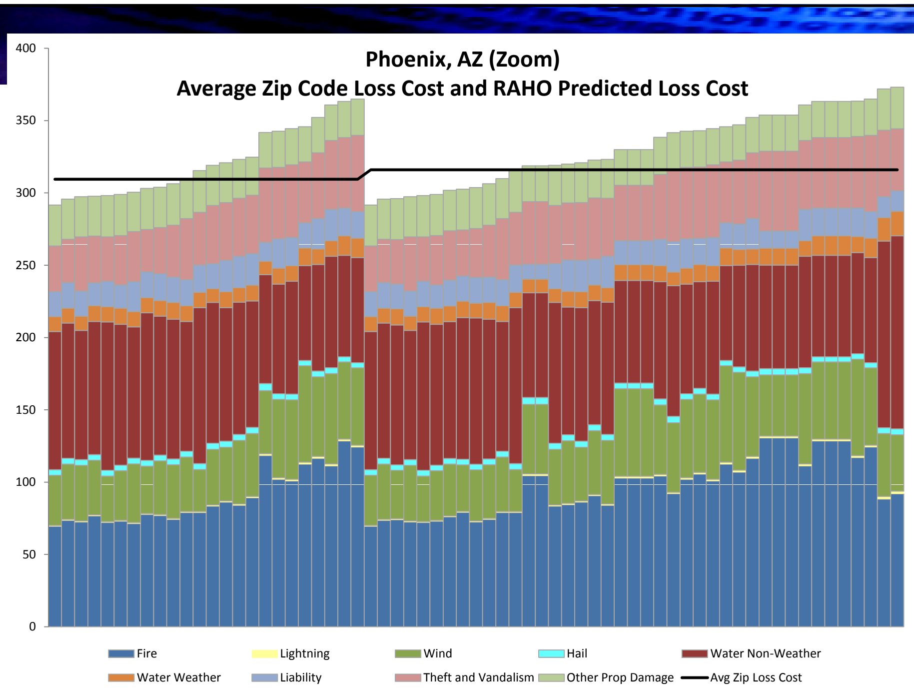


Phoenix, AZ Geographic Area



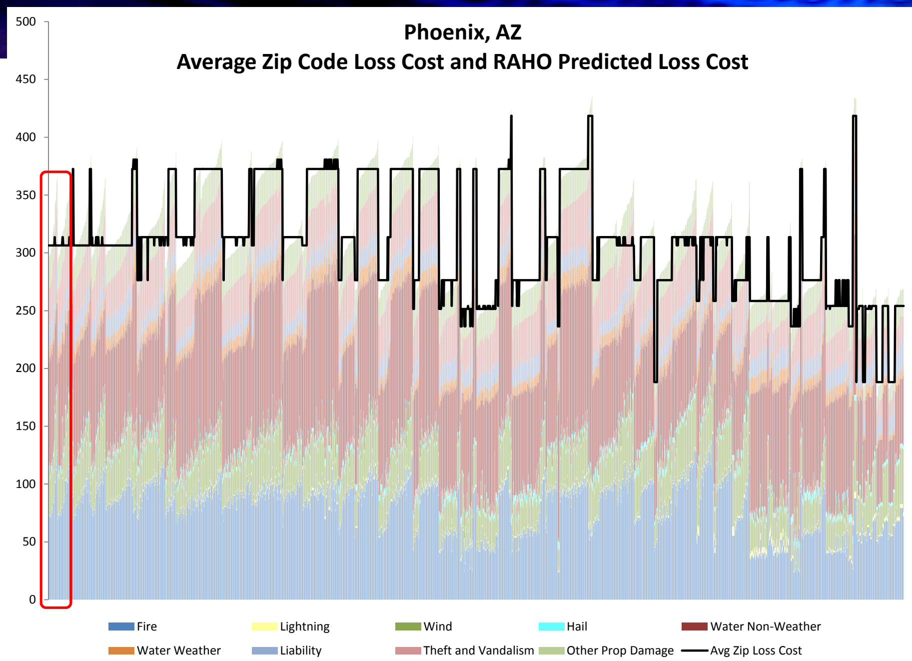
ISO Territories: 9 | Zip Codes: 80 | RAHO: 1309

27



* Loss cost are calculated @ Territory Representative Risk

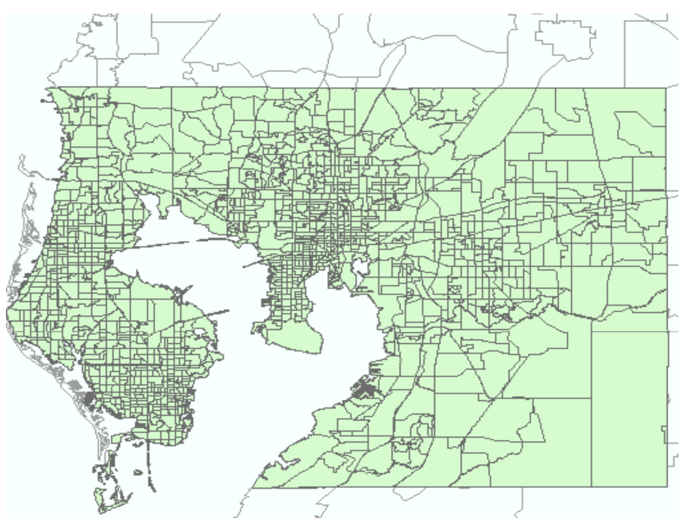
28



* Loss cost are calculated @ Territory Representative Risk

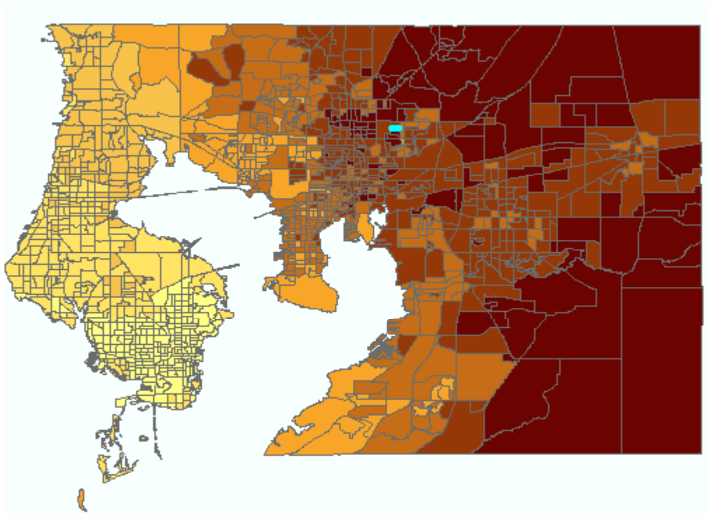
29

Tampa Bay, FL Area



30

Tampa Bay Area
Detailed Loss Costs (Non-Hurricane)



31

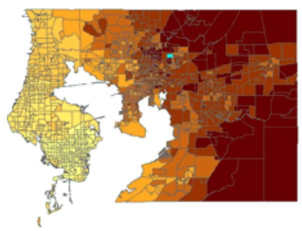
Opportunities for Enhanced Segmentation

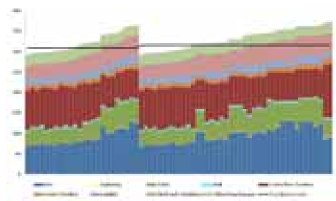
- **Use sum-of-peril loss cost estimates**
  - Build new territories
  - Refine existing territories
- **Use peril-specific models to break apart all-peril rating**
  - Geographic exposures and rating variables
- **Using components as input to models**
  - Incorporate new predictive data with simpler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

32

Questions?

David Cummings
dcummings@iso.com





33