PM-12: Handling High-Dimensional Variables: Territory Analysis

CAS Ratemaking and Product Management Seminar

Klayton N. Southwood, FCAS, MAAA

March 16, 2010

TOWERS WATSON

CAS antitrust notice

ANTITRUST Notice

The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding — expressed or implied — that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

Risk Classification

Definition — A grouping of risks with similar risk characteristics so that differences in expected costs may be recognized

Purpose — Means by which data can be gathered so as to measure and quantify a specific risk characteristic's relation to the propensity for loss

Example — Territorial classes are a means to gather data so as to measure and quantify geographic risk factors relative to the propensity for loss

Homogeneity

Definition — A risk classification is homogeneous if all risks in the class have the same or a similar expected degree of risk with respect to the risk factor being measured

Purpose — Homogeneity of the class increases the credibility of the loss data generated by the class

Example — A territory is considered homogeneous if all risks in the territory represent the same, or approximately the same, level of geographical risk (all else being equal)

Statistical Test of Homogeneity

Within Variance = Based on the squared difference between each zip code pure premium in the cluster and the average pure premium for the specific cluster being tested

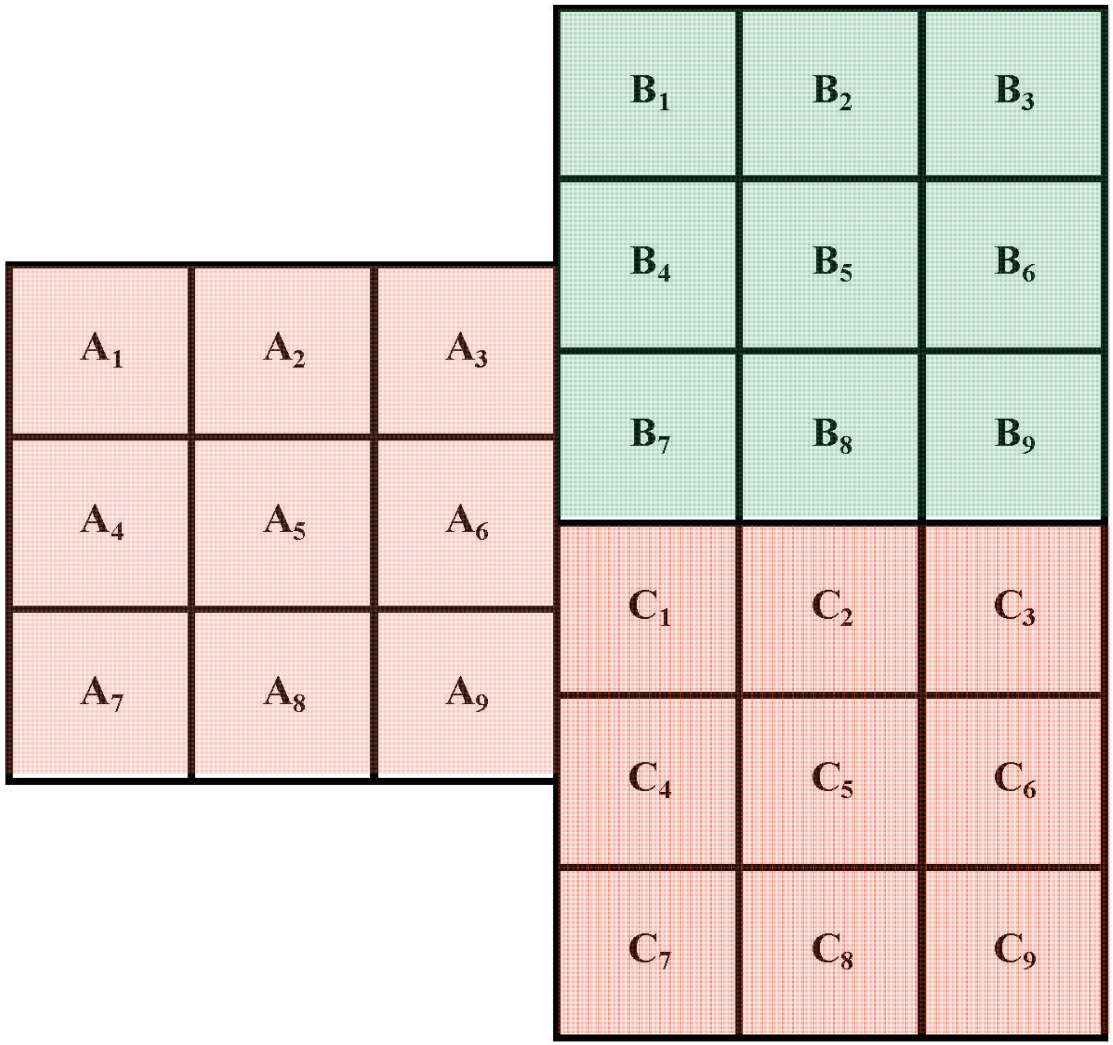
Between Variance = Based on the squared difference between each cluster's pure premium and the statewide average pure premium

Total Variance = Within Variance + Between Variance

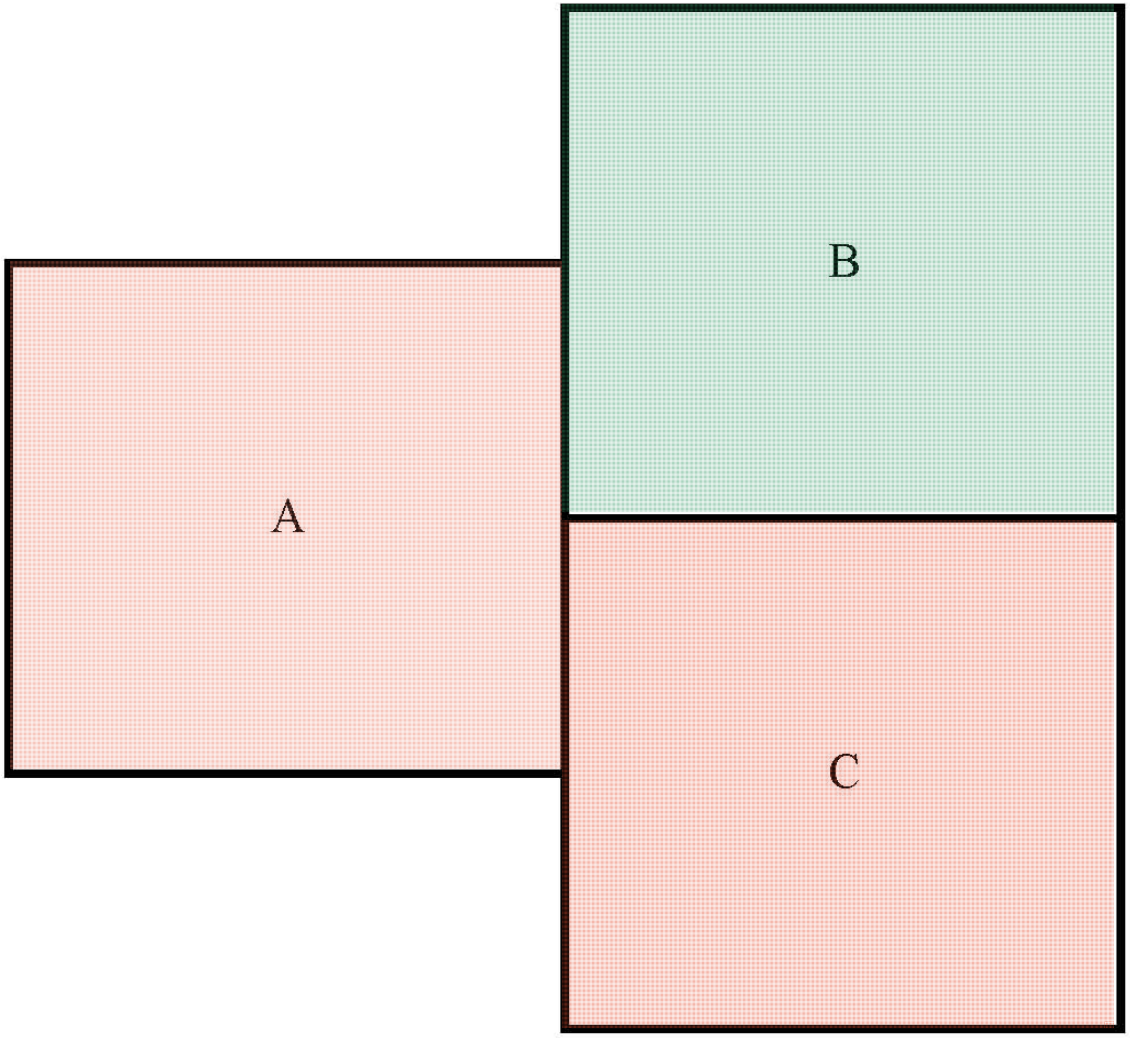
Within Variance Percentage = Within Variance divided by Total Variance

Goals: $\frac{\text{Low Percentage of Total Variance Within}}{\text{High Percentage of Total Variance Between}}$

Building Blocks



Territorial Risk Classes



Basis to group areas

County

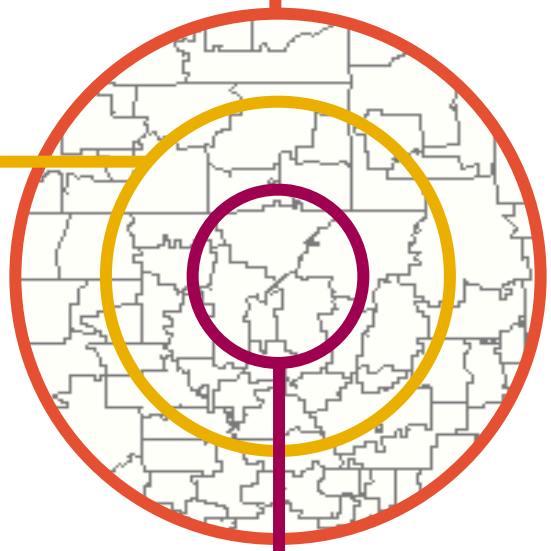
- Largely stable over time
- Broad area

ZIP Code

- Narrowly defined — may be beneficial to define territories
- Useful for online rating
- Main disadvantage is need to deal with change over time

Geo-Coding

- Finest detail
- Static over time
- No predefined grouping



Loss index normalized pure premium

Normalized Zip Code Pure Premium

=

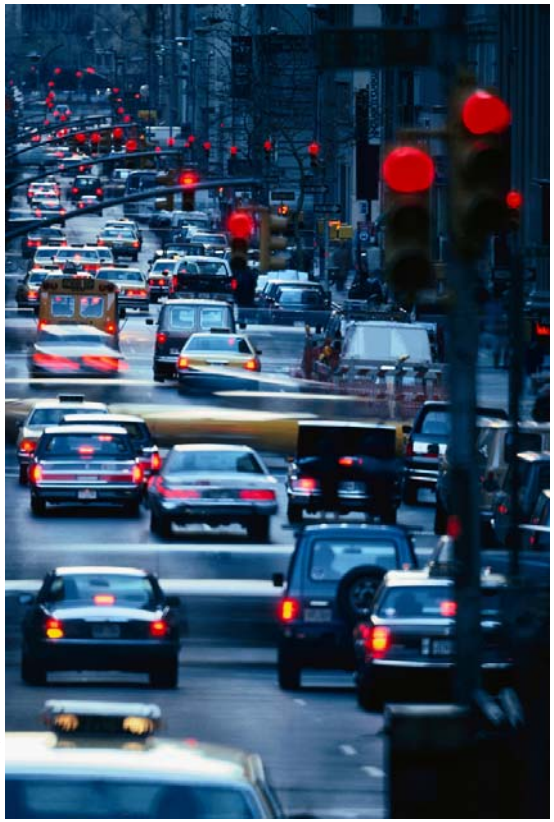
Actual Zip Code Pure Premium

X

$$\frac{\text{State Avg. Prem.}}{\text{State Avg. Base}} \div \frac{\text{Zip Avg. Prem.}}{\text{Zip Base}}$$

Loss index econometric model — Private passenger auto

- Population density
- Vehicle density
- Accidents per vehicle
- Injuries per accident
- Thefts per vehicle



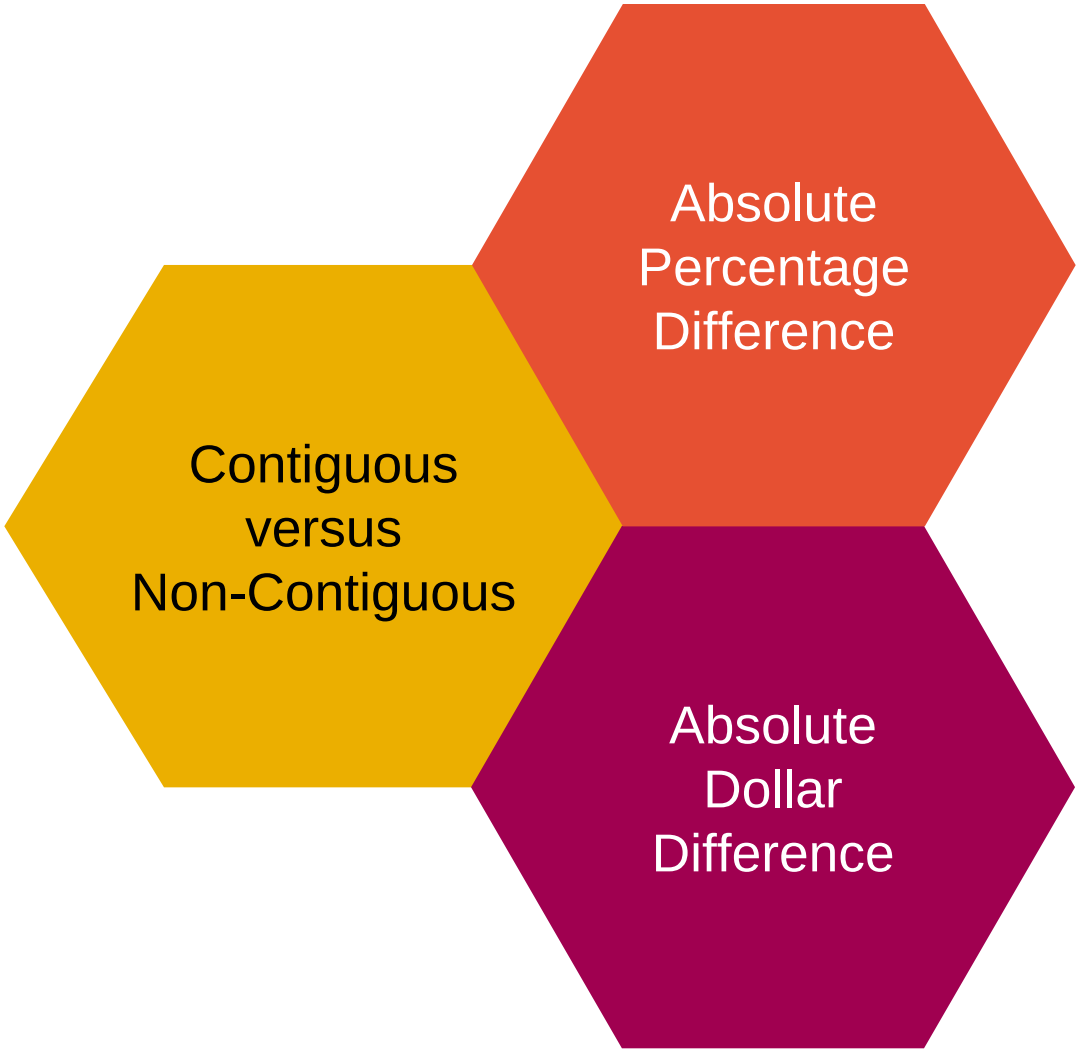
Loss index econometric model — Business owners liability

- Departure from normal temperature
- Number of days maximum temperature is below freezing
- Total precipitation
- Population density
- Population growth
- Percent of population using public transportation

Credibility

- No “right” answer
- We commonly use:
  - 3,000 claims
  - With complement applied to:
    - Neighborhood pure premium
    - Within two miles
    - One mile extensions

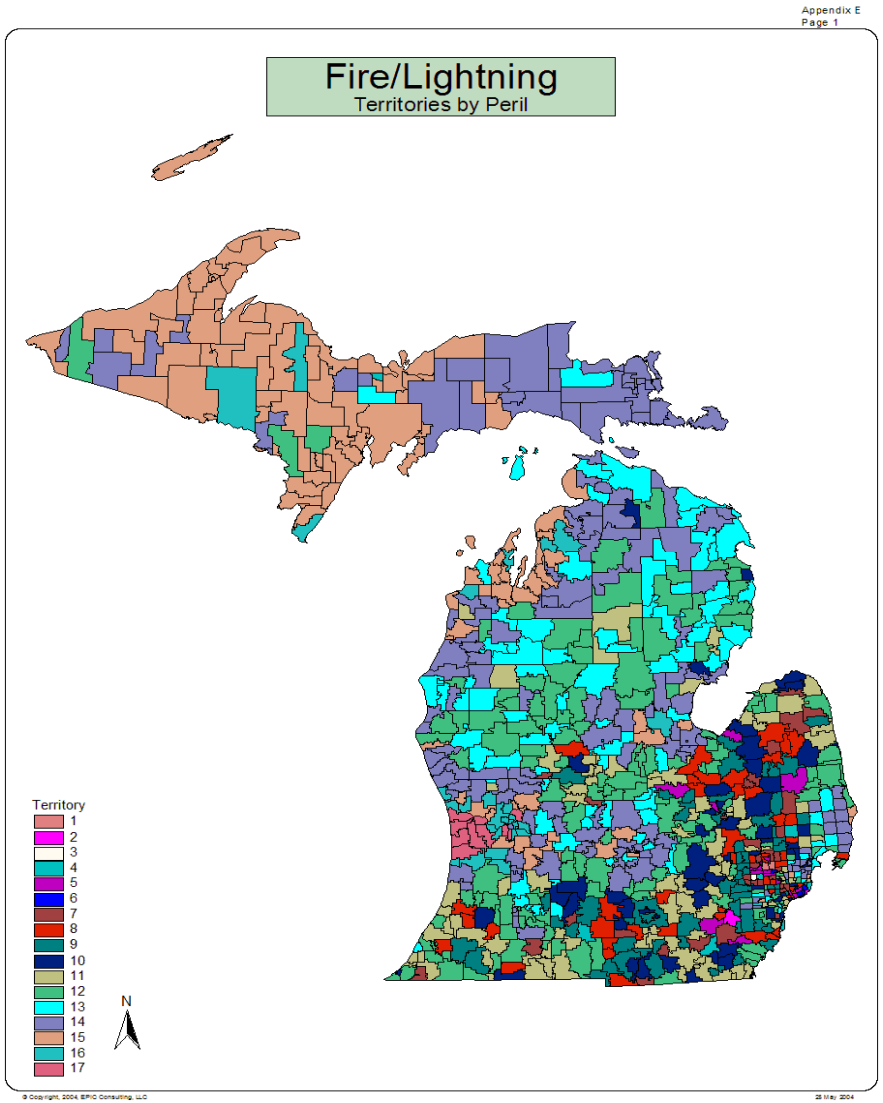
Clustering



Michigan Industry Homeowners

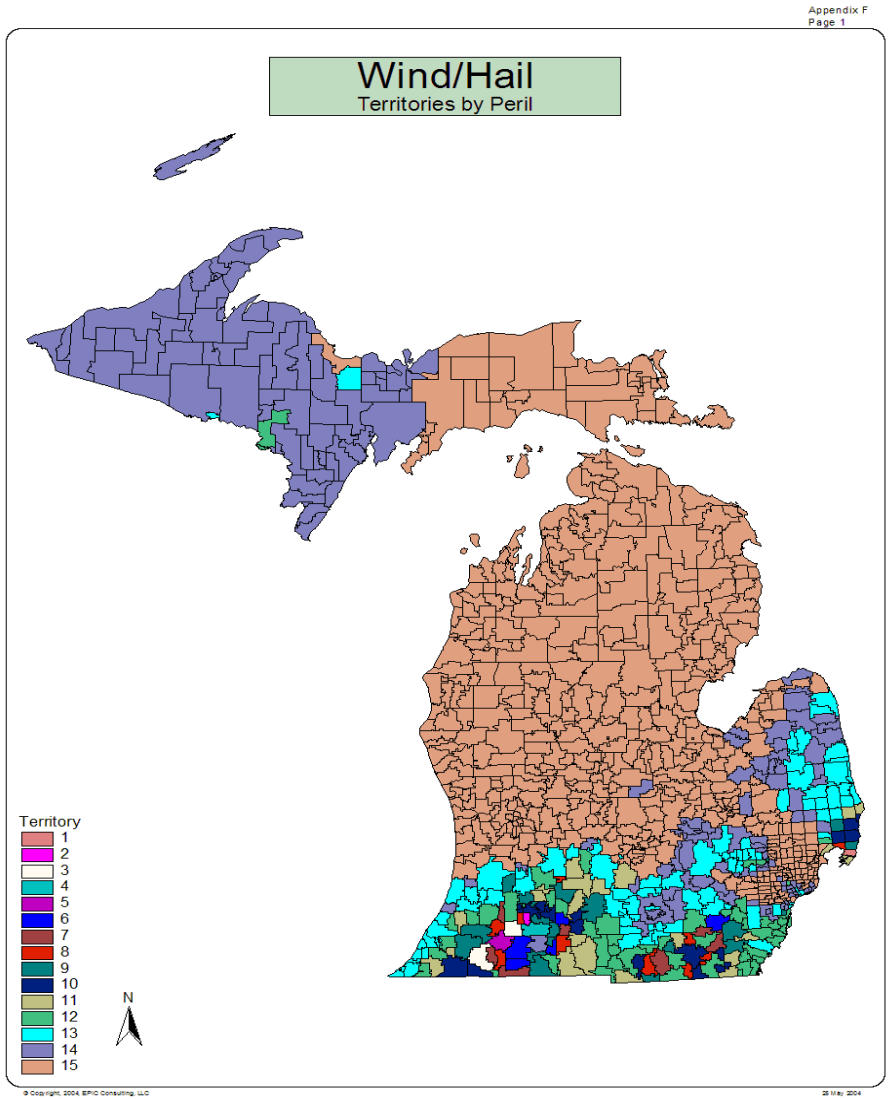
Industry homeowners — Fire (non-contiguous)

Michigan



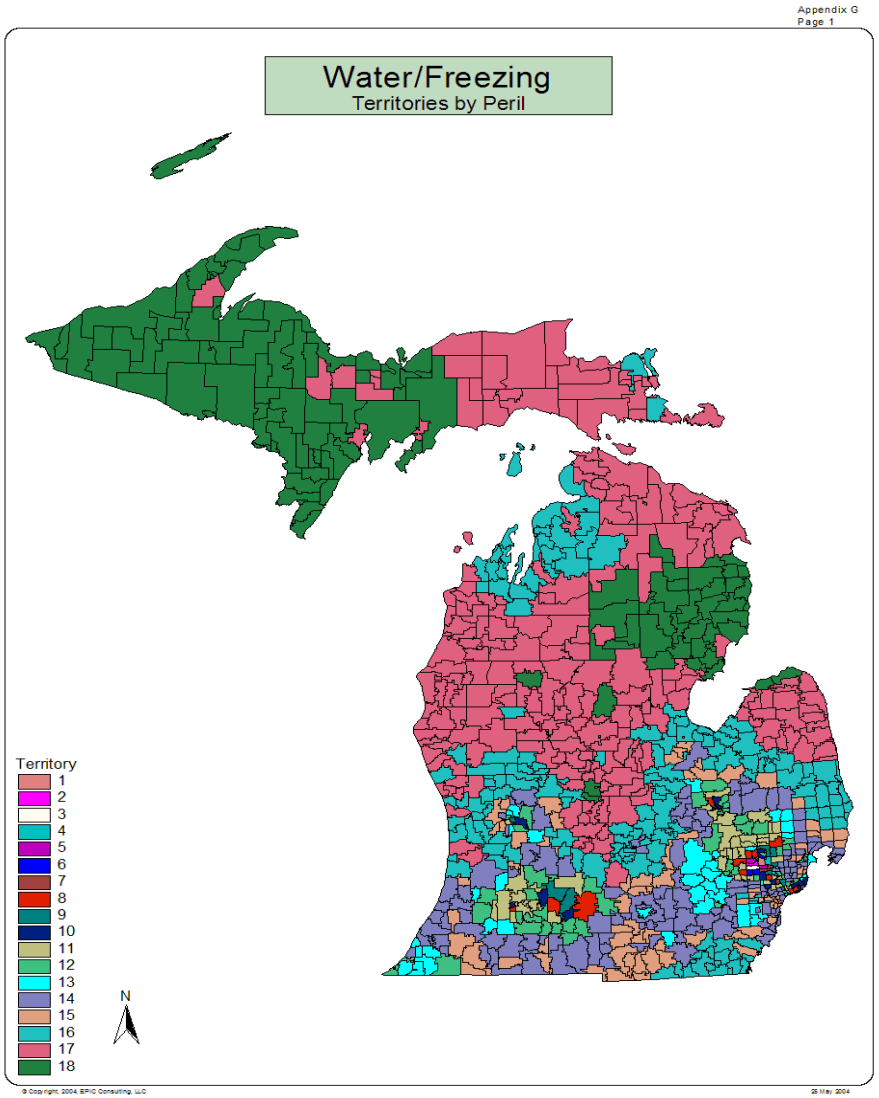
Industry homeowners — Wind/hail (non-contiguous)

Michigan



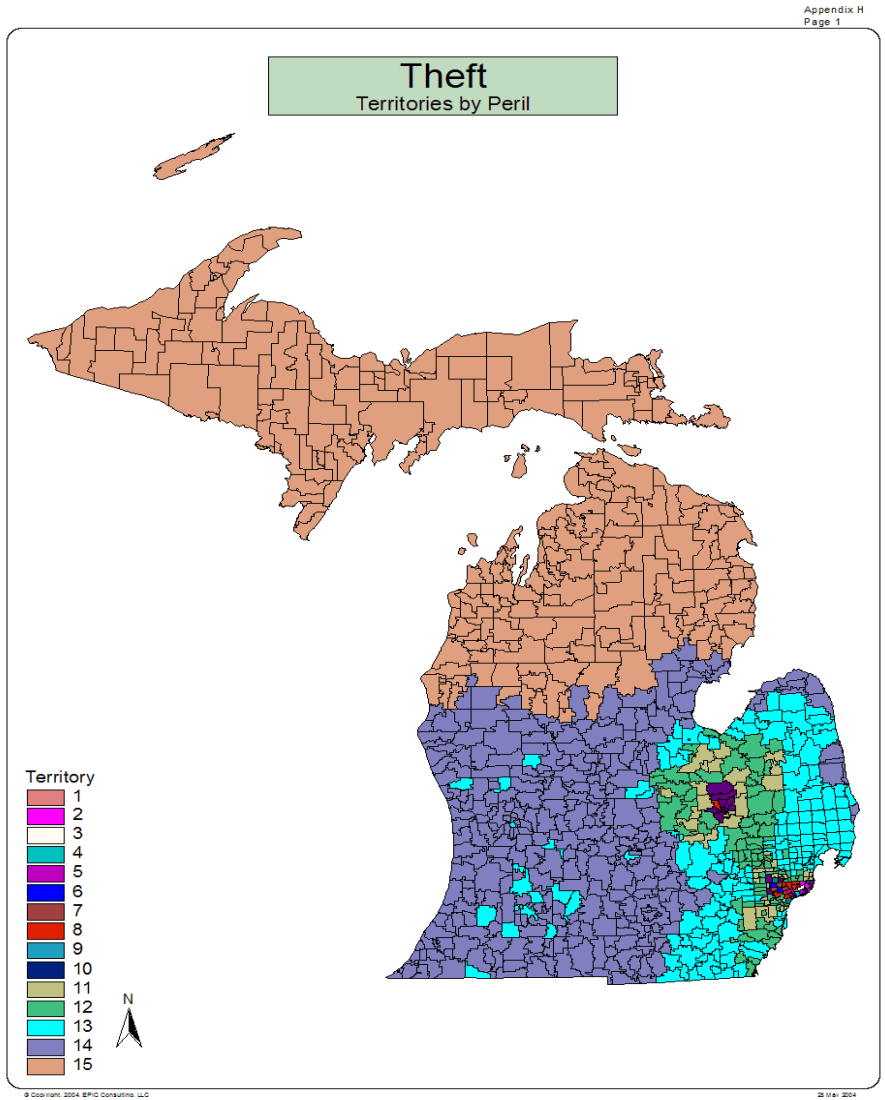
Industry homeowners — Water/freezing (non-contiguous)

Michigan



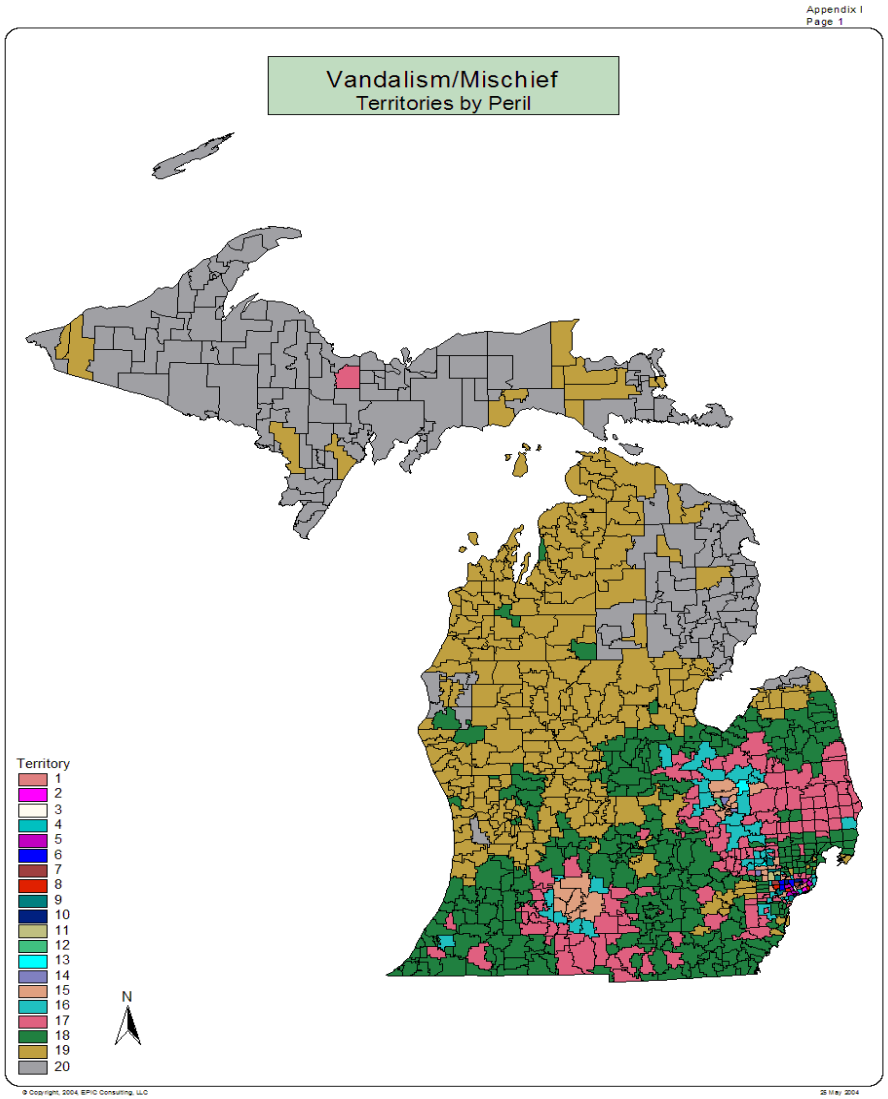
Industry homeowners — Theft (non-contiguous)

Michigan



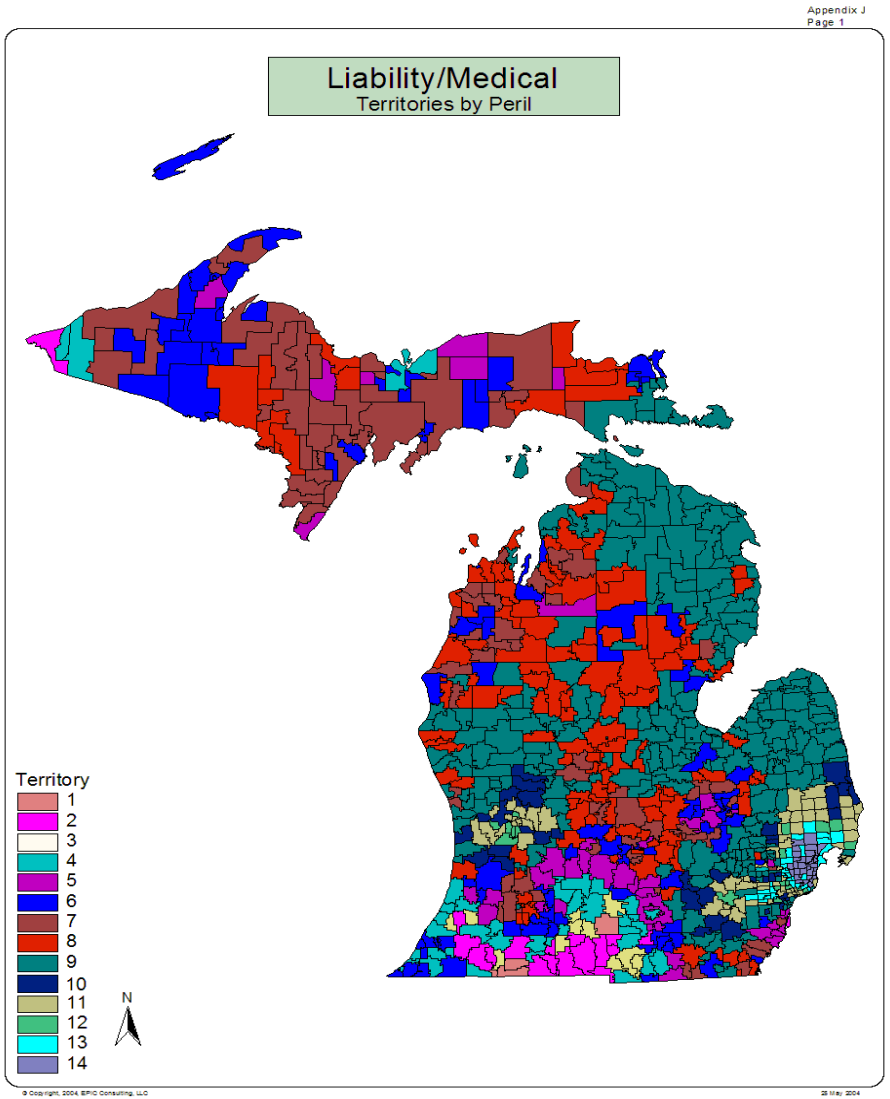
Industry homeowners — Vandalism (non-contiguous)

Michigan



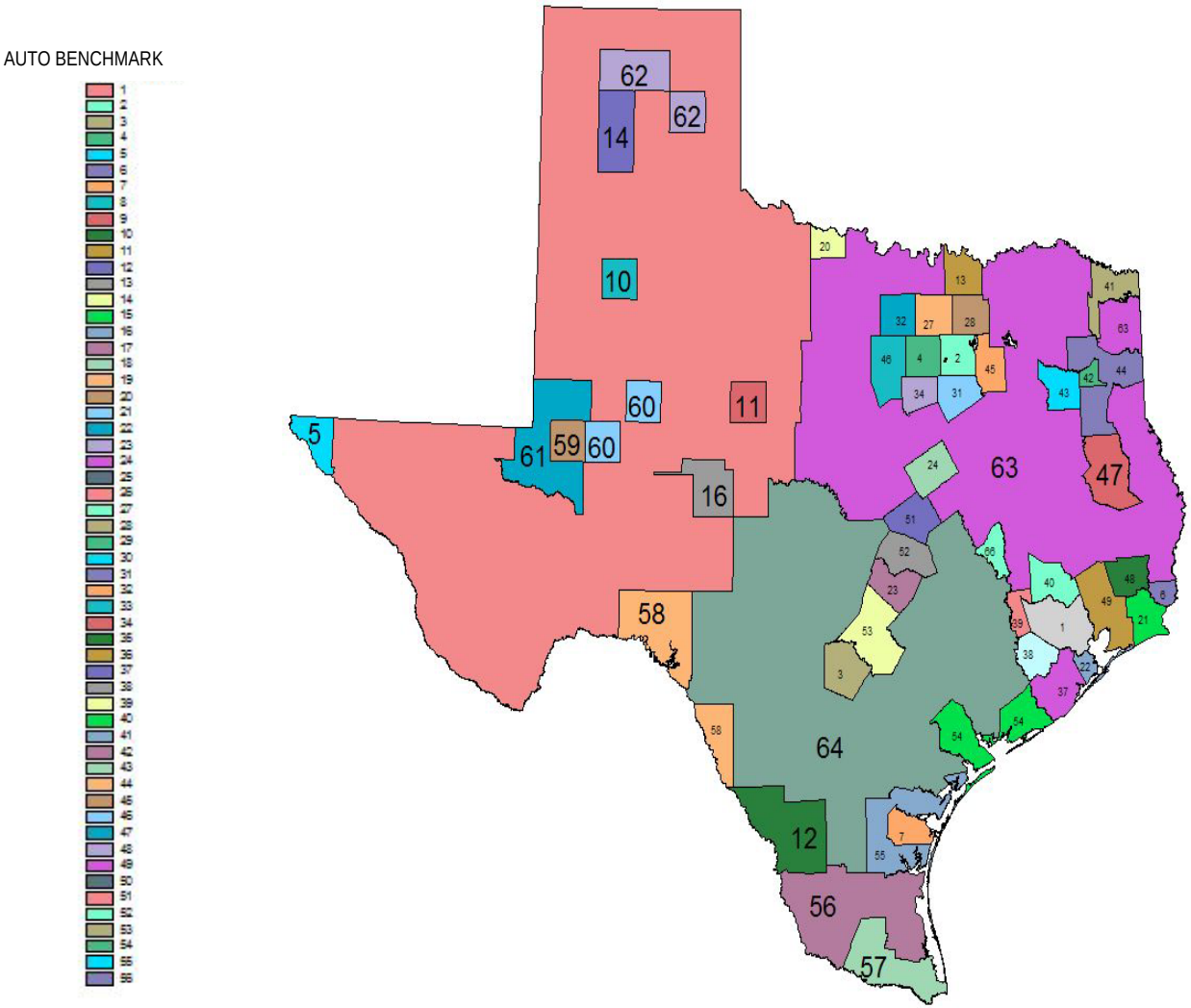
Industry homeowners — Liability (non-contiguous)

Michigan



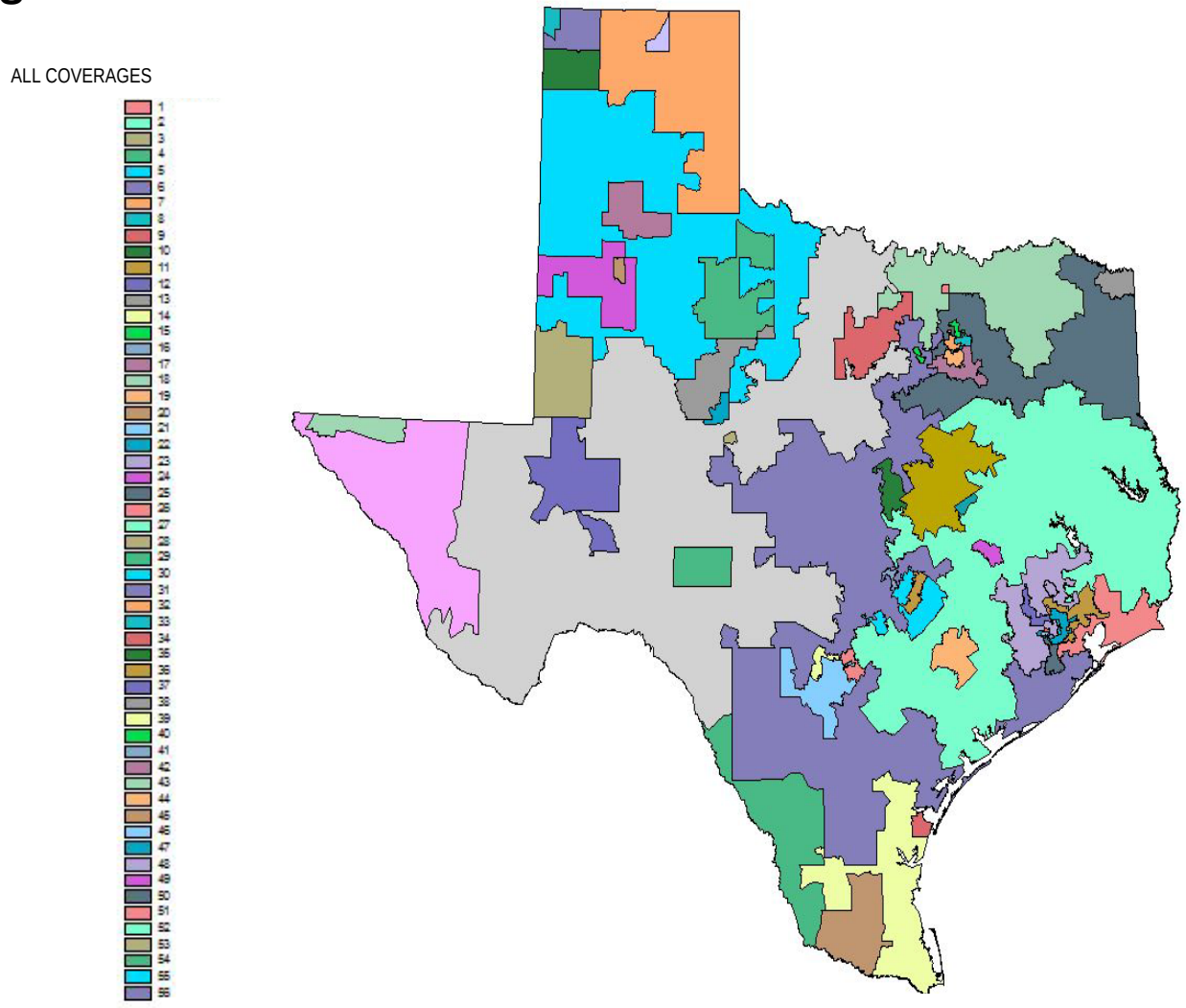
Texas Auto Benchmark

Texas auto benchmark



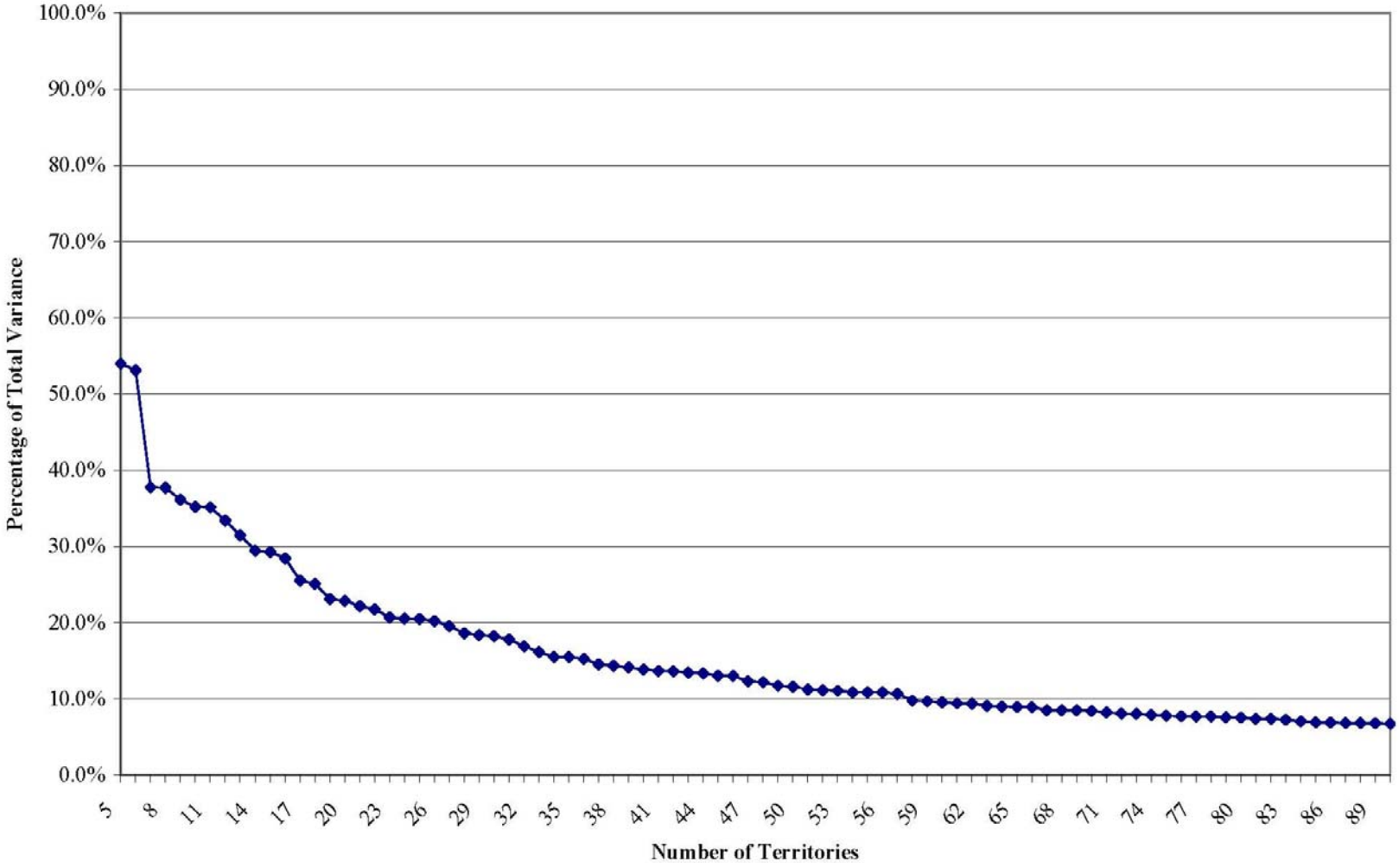
Indicated auto territories — All coverages (contiguous)

Texas



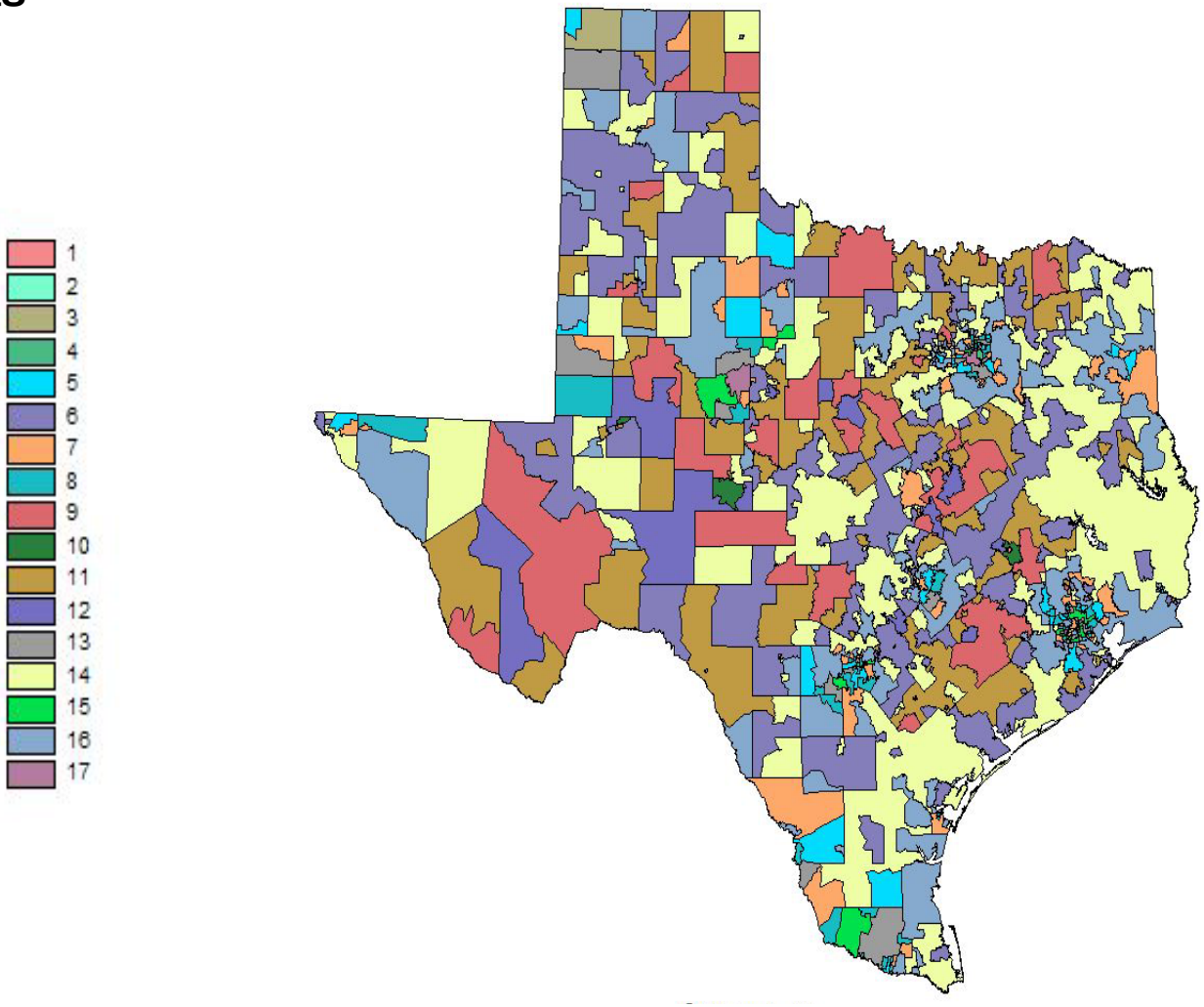
Within territory variance as a percentage of total variance — All coverages (contiguous)

Texas



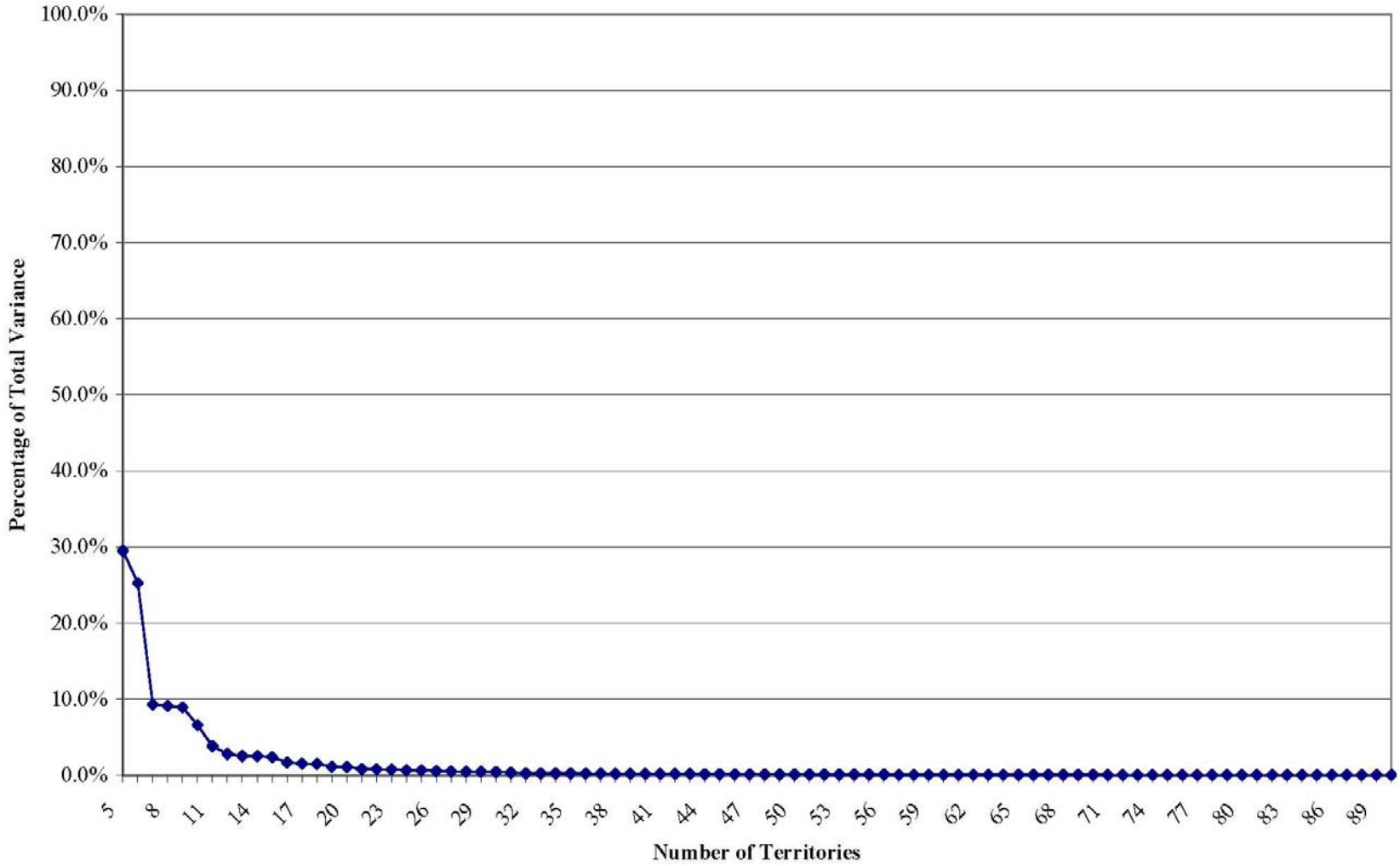
Indicated auto territories — All coverages (non-contiguous)

Texas



Within territory variance as a percentage of total variance — All coverages (non-contiguous)

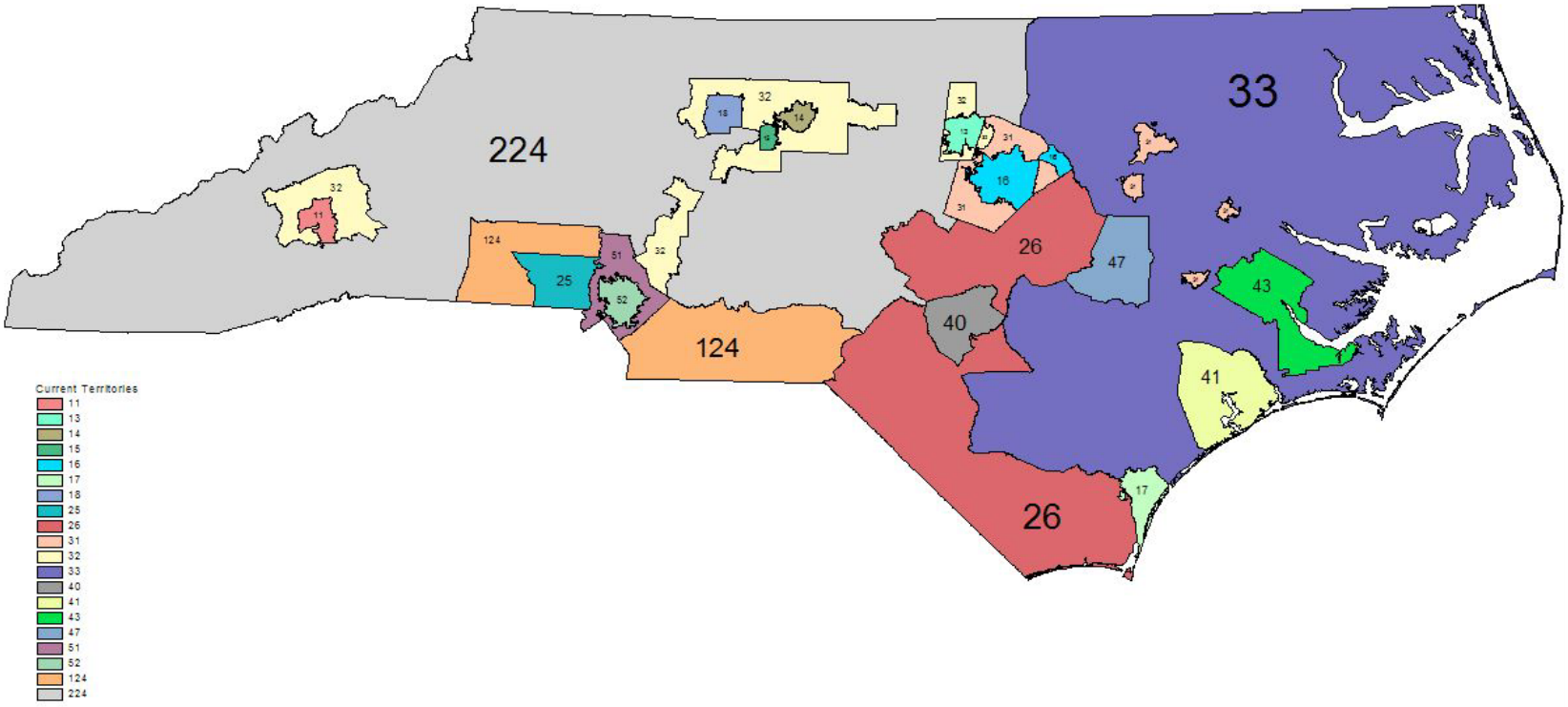
Texas



North Carolina

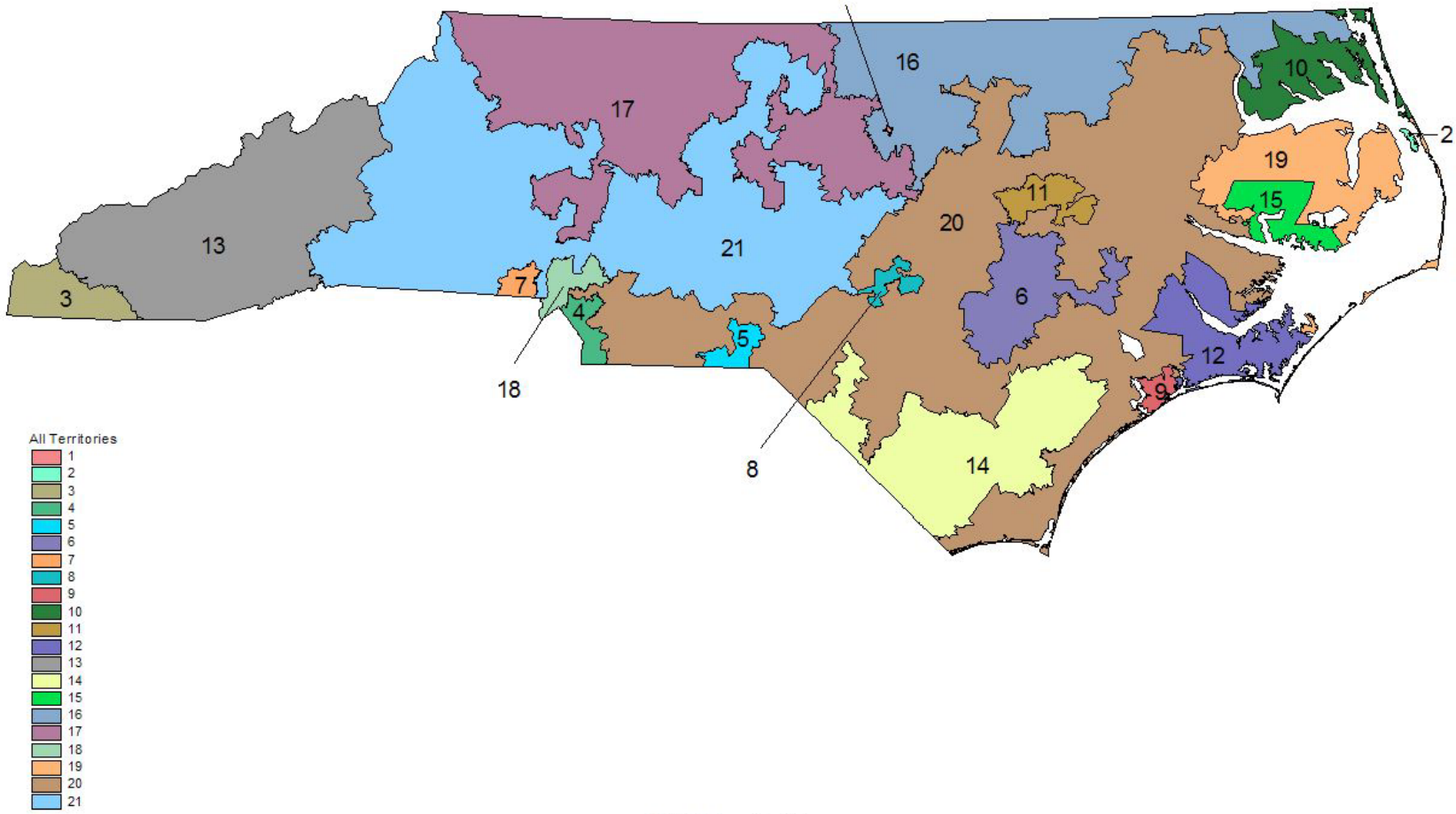
Current auto territories — All coverages

North Carolina



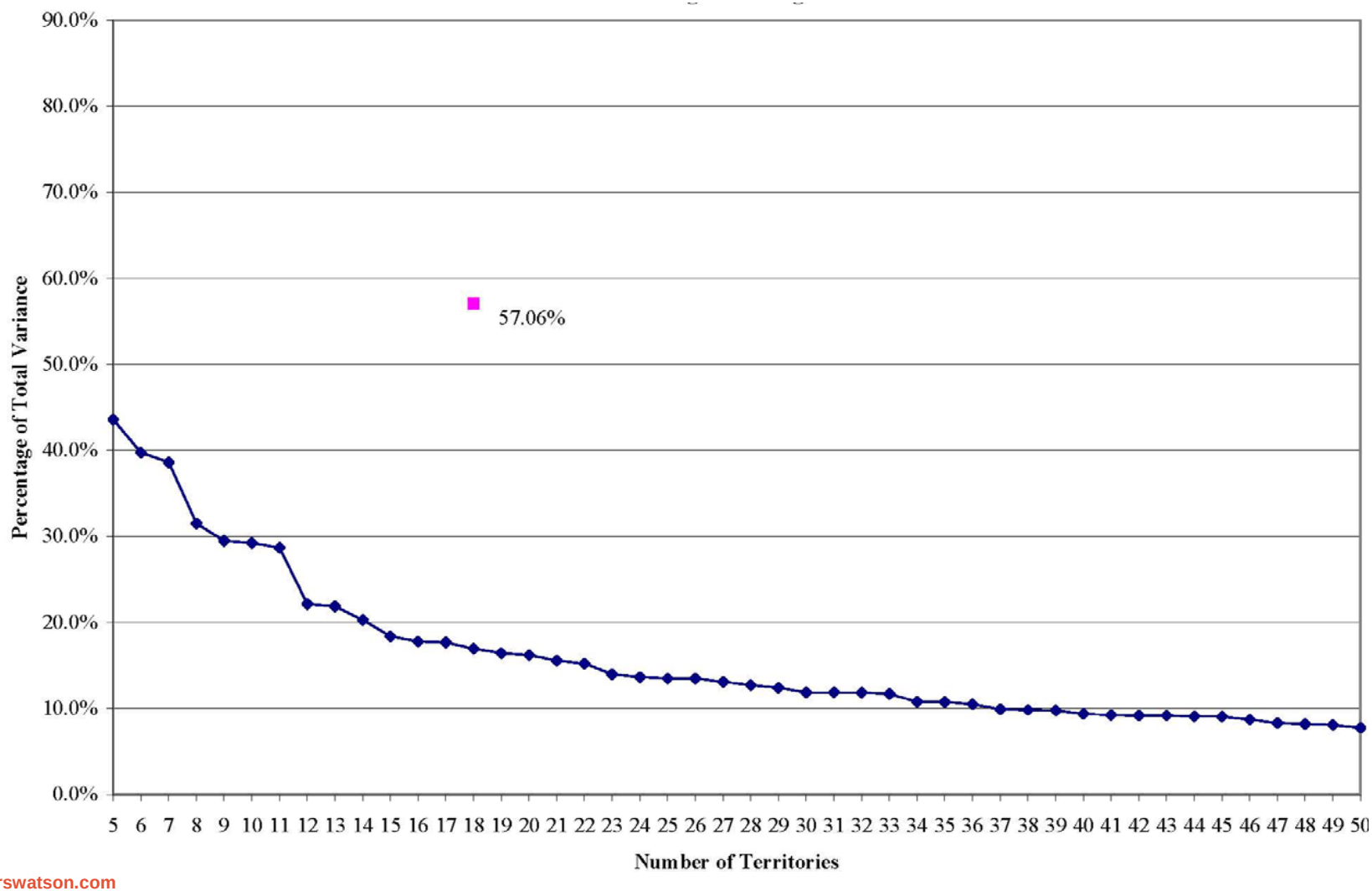
Indicated auto territories — All coverages (contiguous)

North Carolina



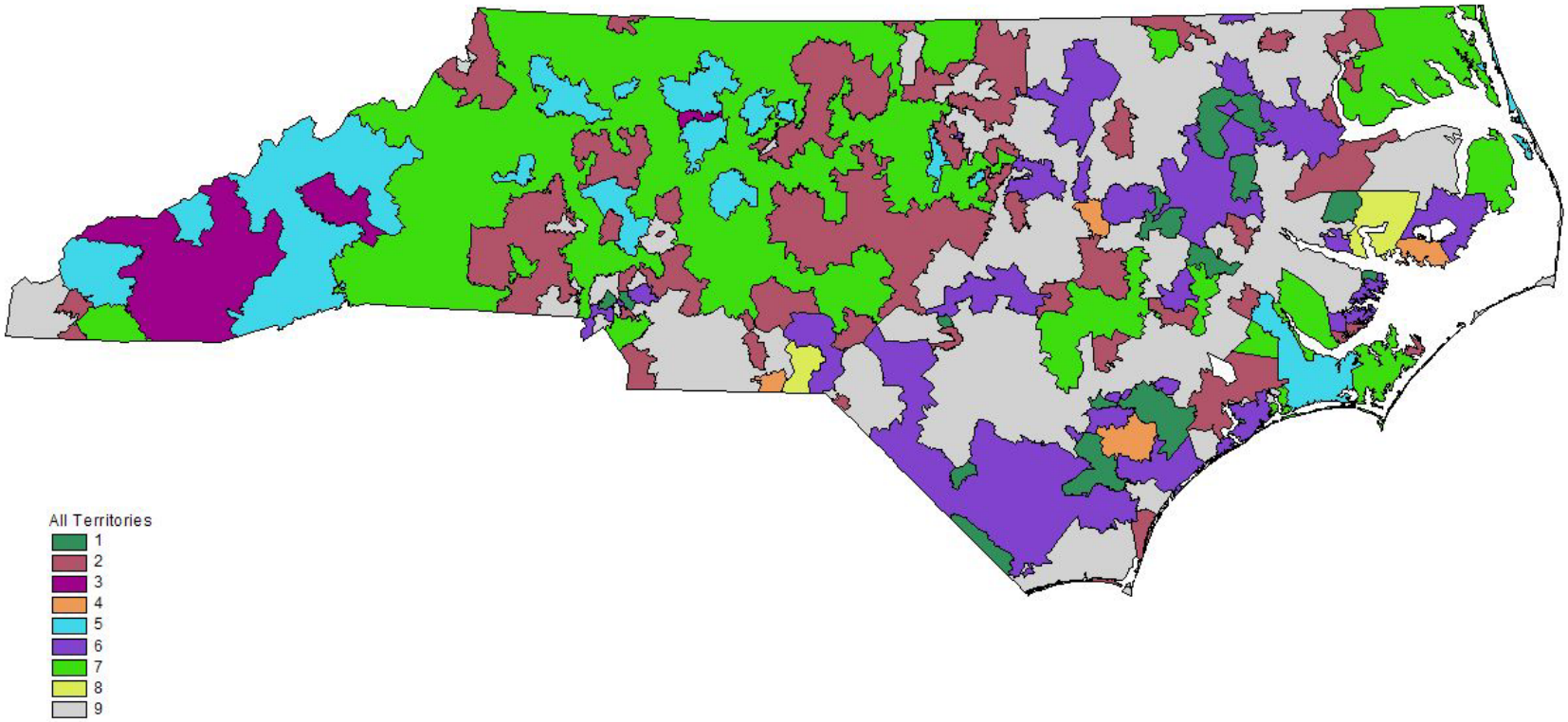
Within territory variance as a percentage of total variance — All coverages (contiguous)

North Carolina



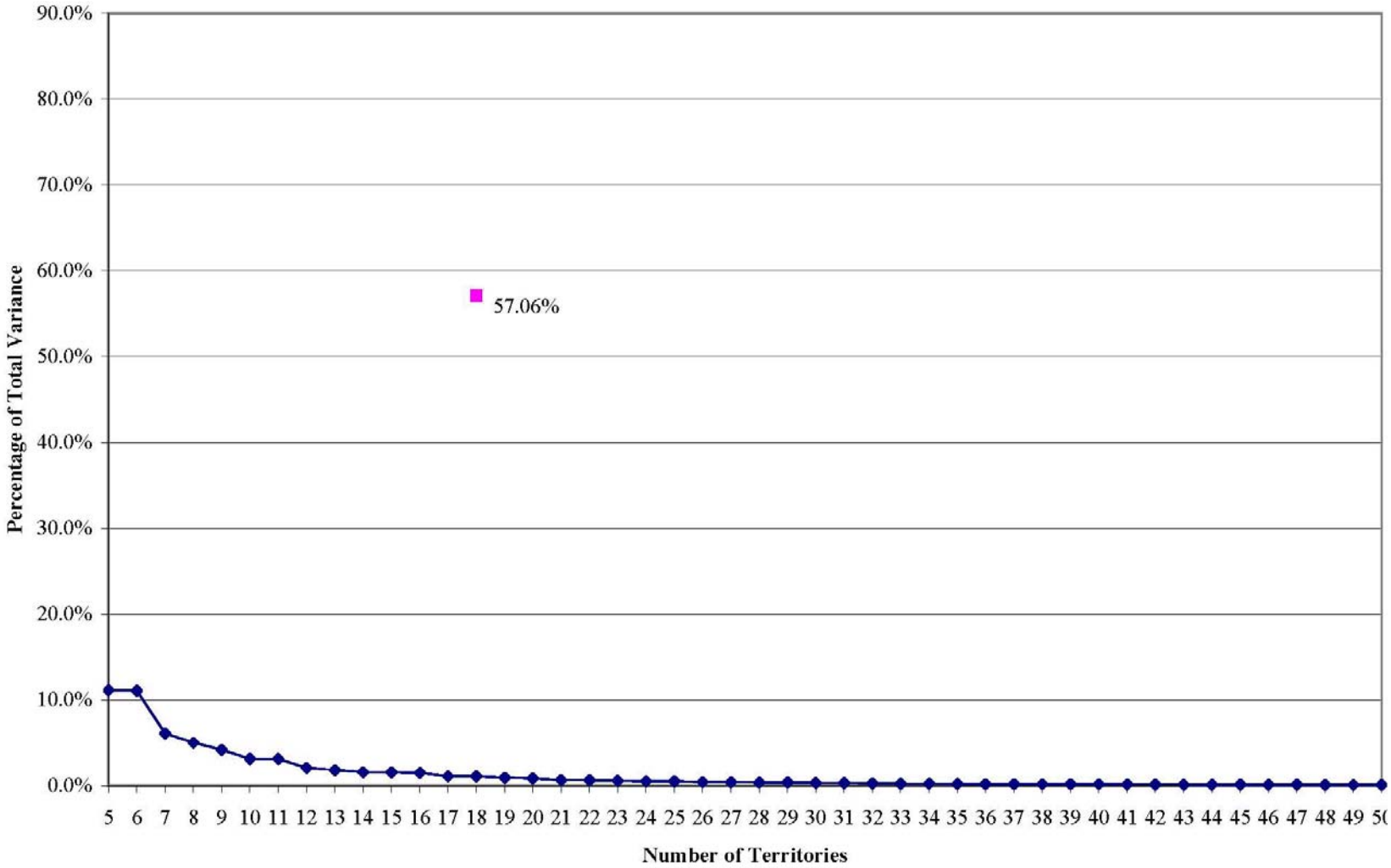
Indicated auto territories — All coverages (non-contiguous)

North Carolina



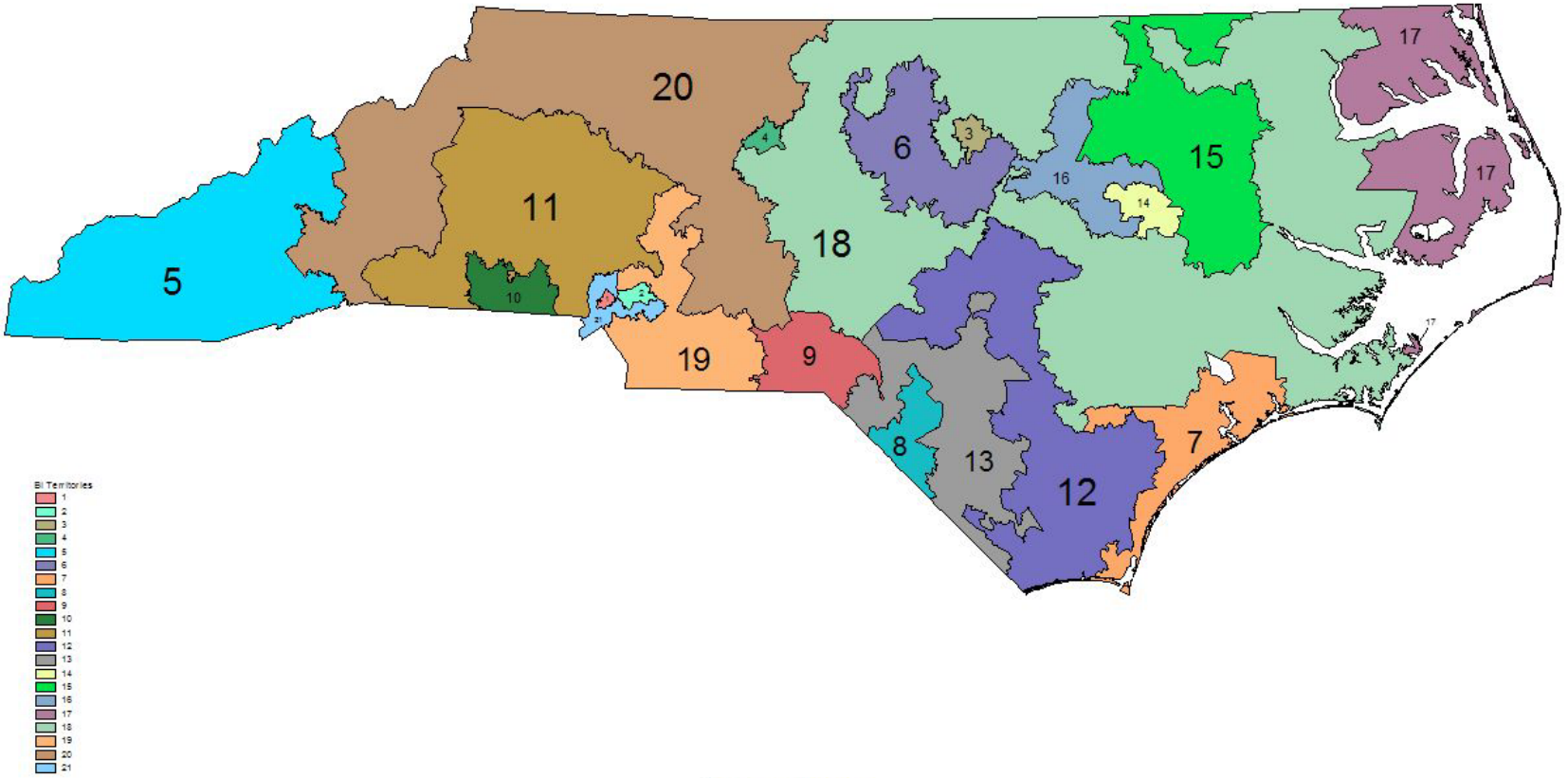
Within territory variance as a percentage of total variance — All coverages (non-contiguous)

North Carolina



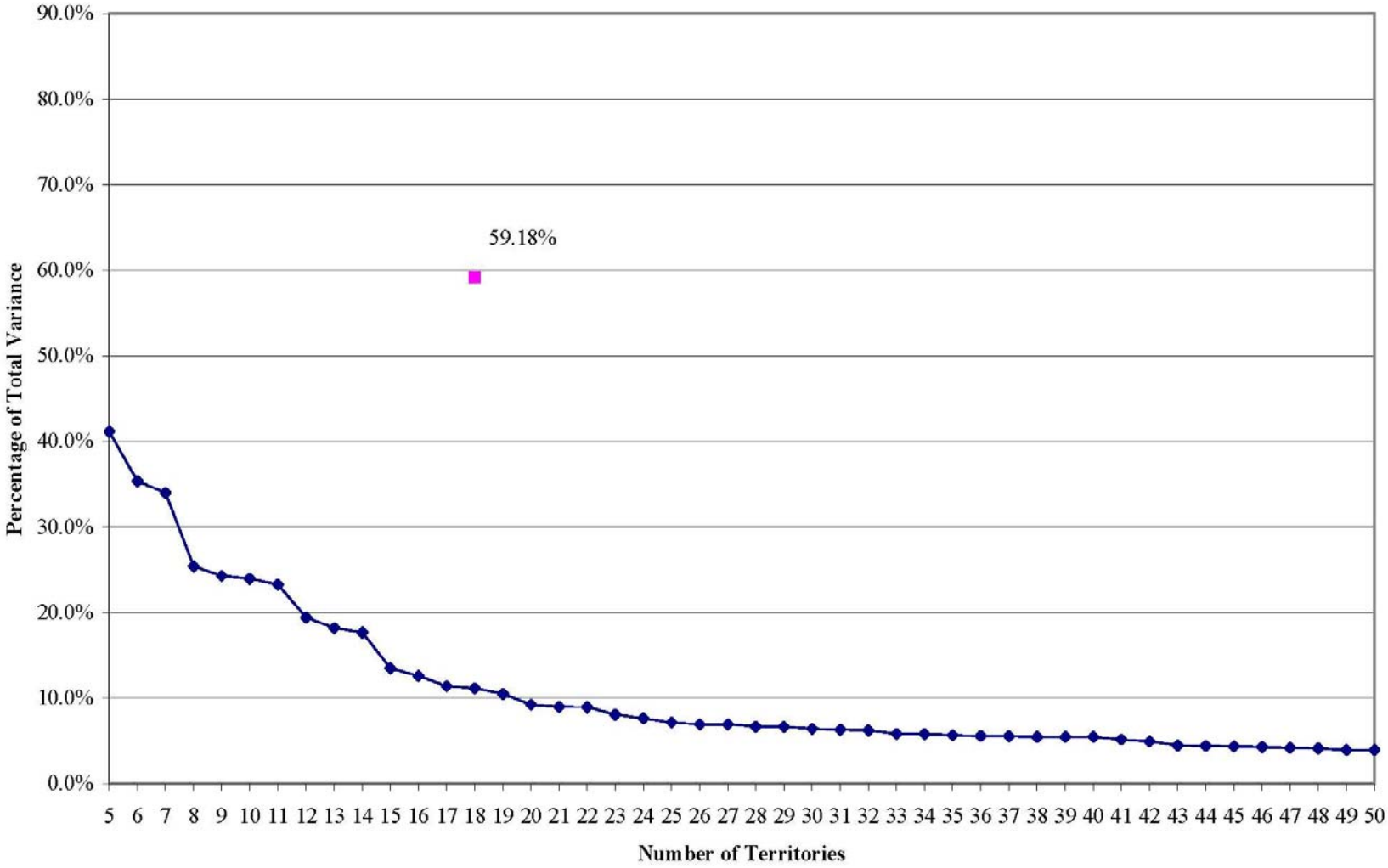
Indicated auto territories — Bodily injury (contiguous)

North Carolina



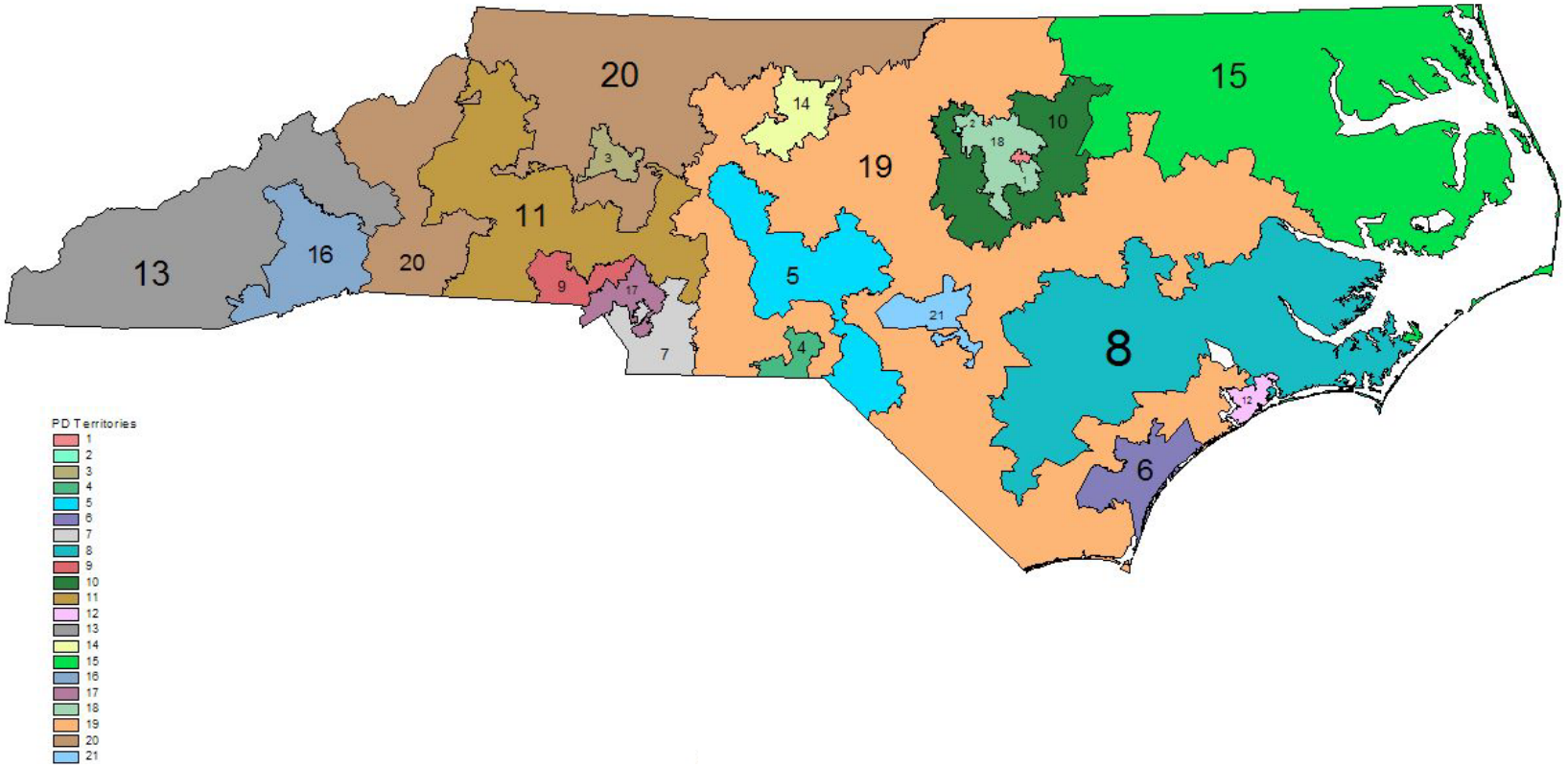
Within territory variance as a percentage of total variance — Bodily injury (contiguous)

North Carolina



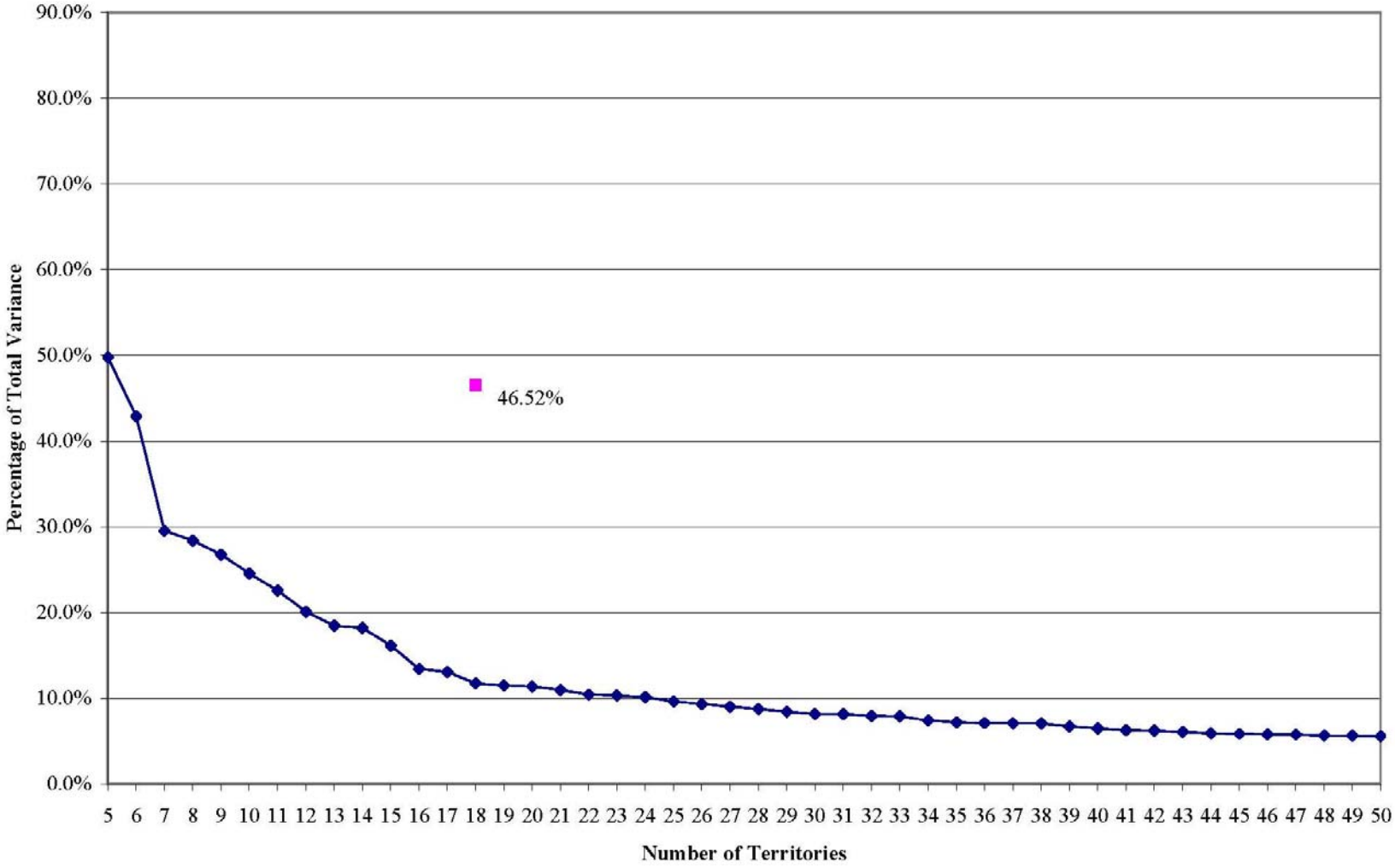
Indicated auto territories — Property damage (contiguous)

North Carolina



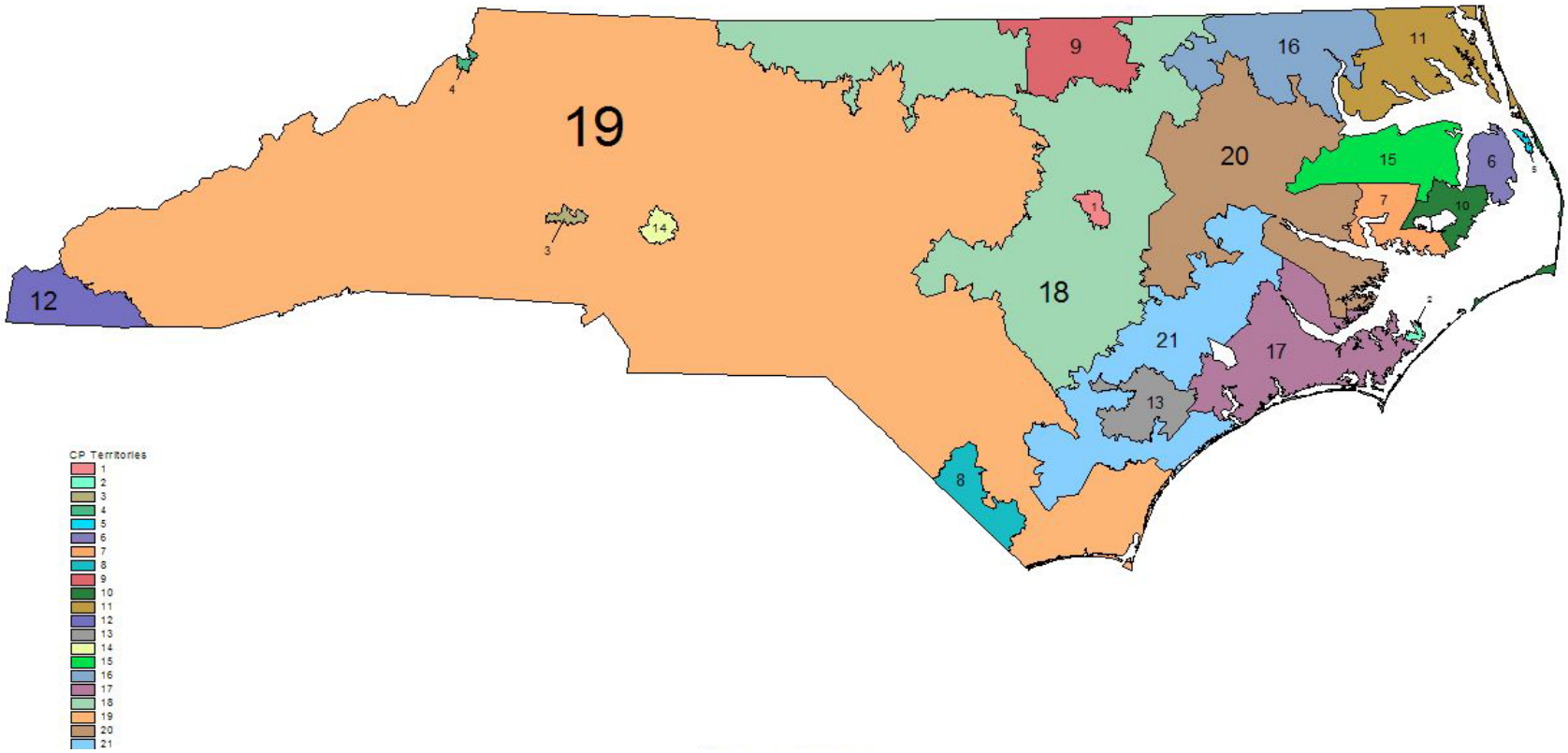
Within territory variance as a percentage of total variance — Property damage (contiguous)

North Carolina



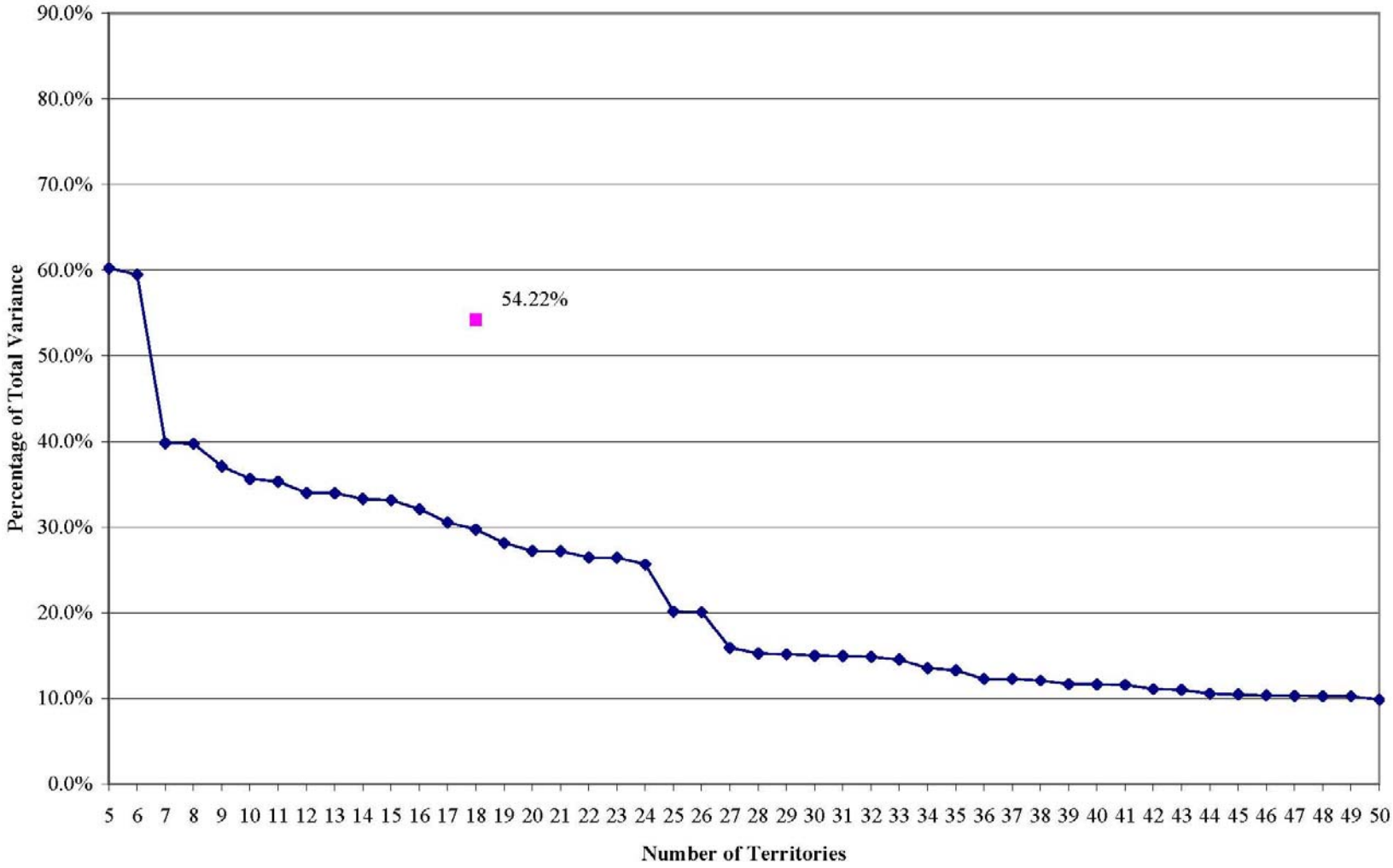
Indicated auto territories — Comprehensive (contiguous)

North Carolina



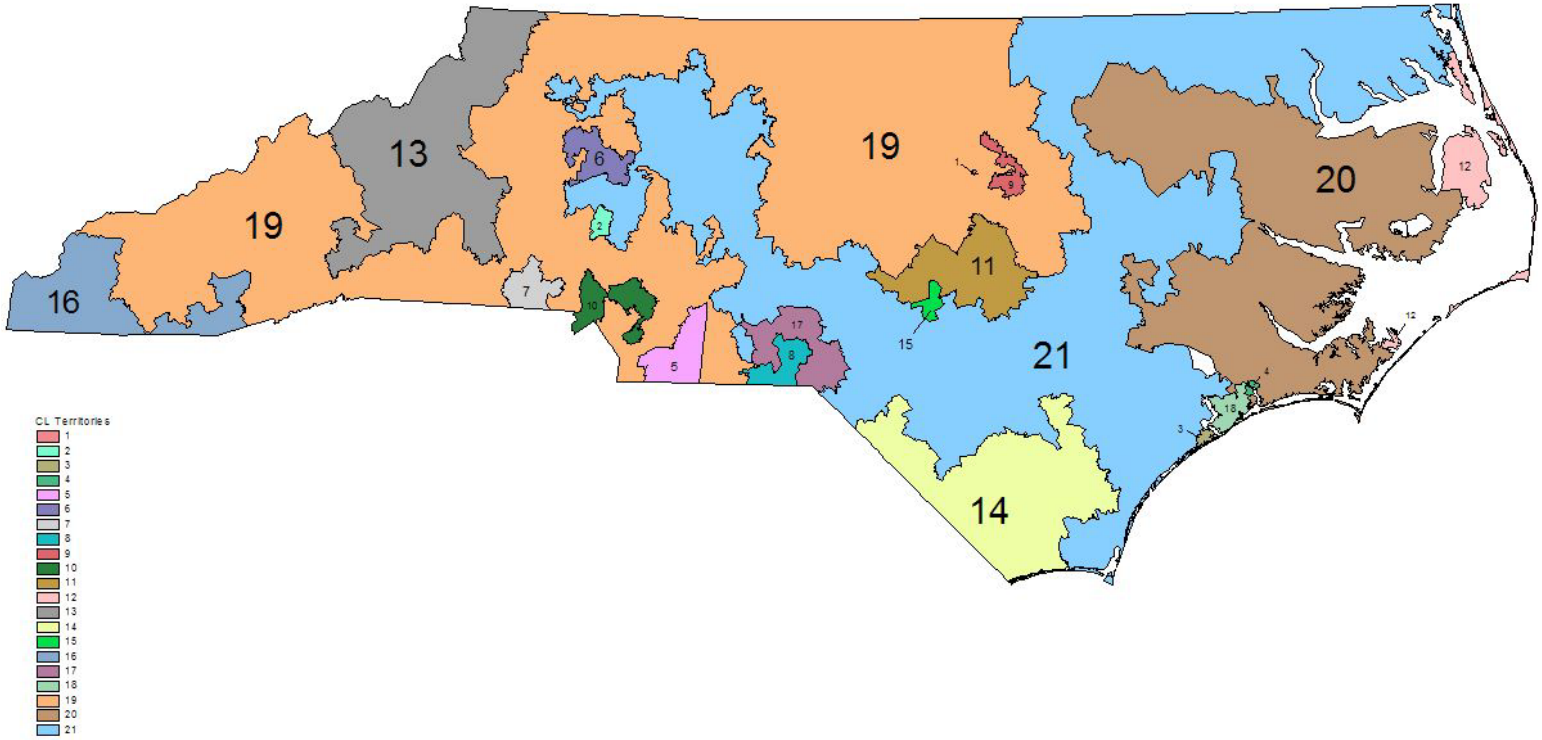
Within territory variance as a percentage of total variance — Comprehensive (contiguous)

North Carolina



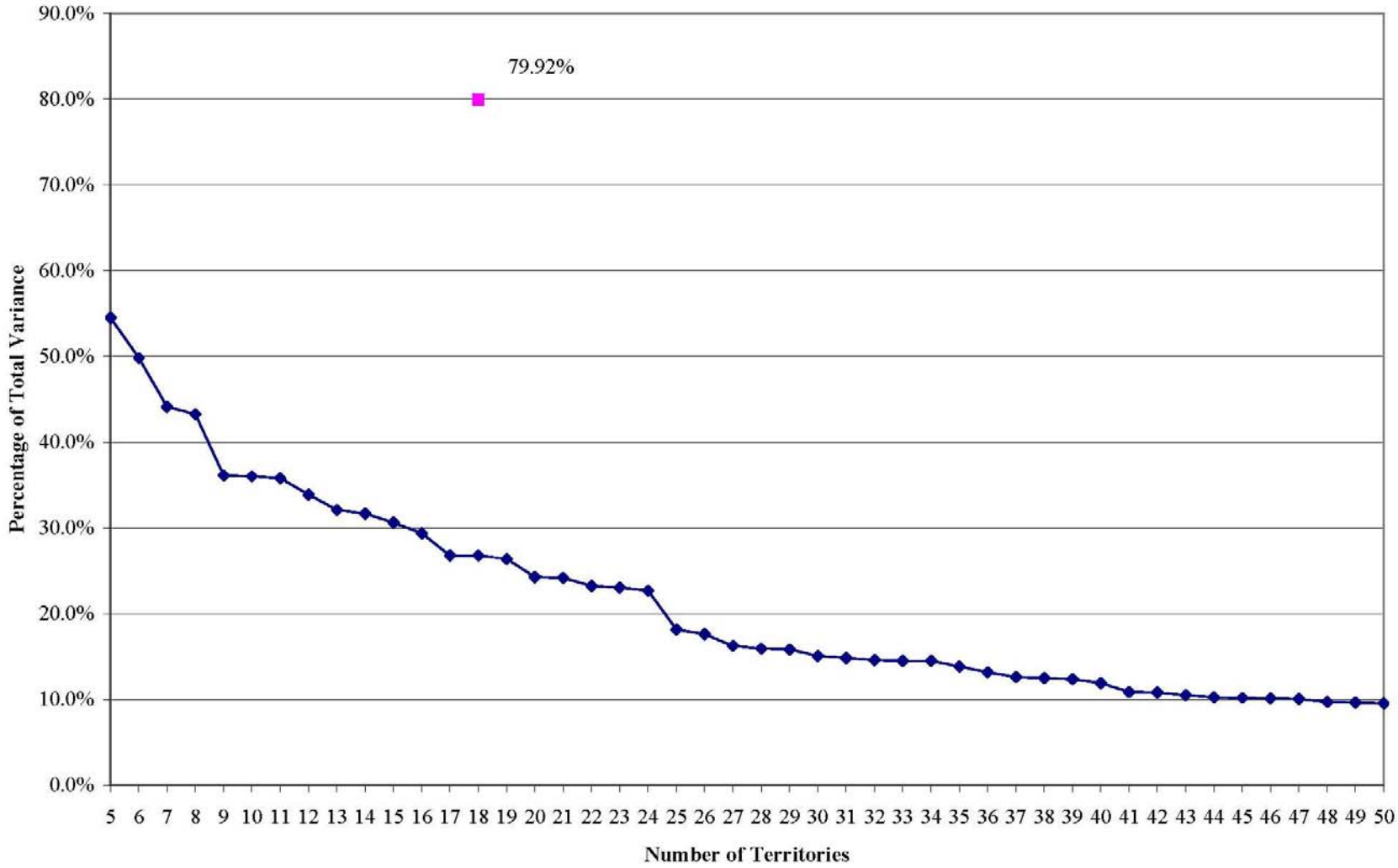
Indicated auto territories — Collision (contiguous)

North Carolina



Within territory variance as a percentage of total variance — Collision (contiguous)

North Carolina



Stability and implementation considerations

Predictive Stability

- Choice of perils included in data
- Number of years of data

Implementation Considerations/Rating Stability

- Limit movement between zones
- Use of capping
- Use of confidence intervals to help analyze changes

Predictive power and stability

Predictive Power — Test #1

- 1993 –1994 versus 1995 – 1996
- Correlation coefficient
- Tested boundaries based on 1994 – 1996
- Non-contiguous better

Predictive Power — Test #2

- 1993 – 1995 versus 1994 – 1996
- Tested boundaries based on 1994 – 1996
- Within variance only marginally better for 1994 – 1996 data

Stability

- 1993 – 1995 clusters versus 1994 – 1996 clusters
- Compared indicated boundaries and relativities
- Little dislocation