Electoral Redistricting with Moment of Inertia and Diminishing Halves Models

Andrew Spann, Dan Gulotta, Daniel Kane

Presented July 9, 2008

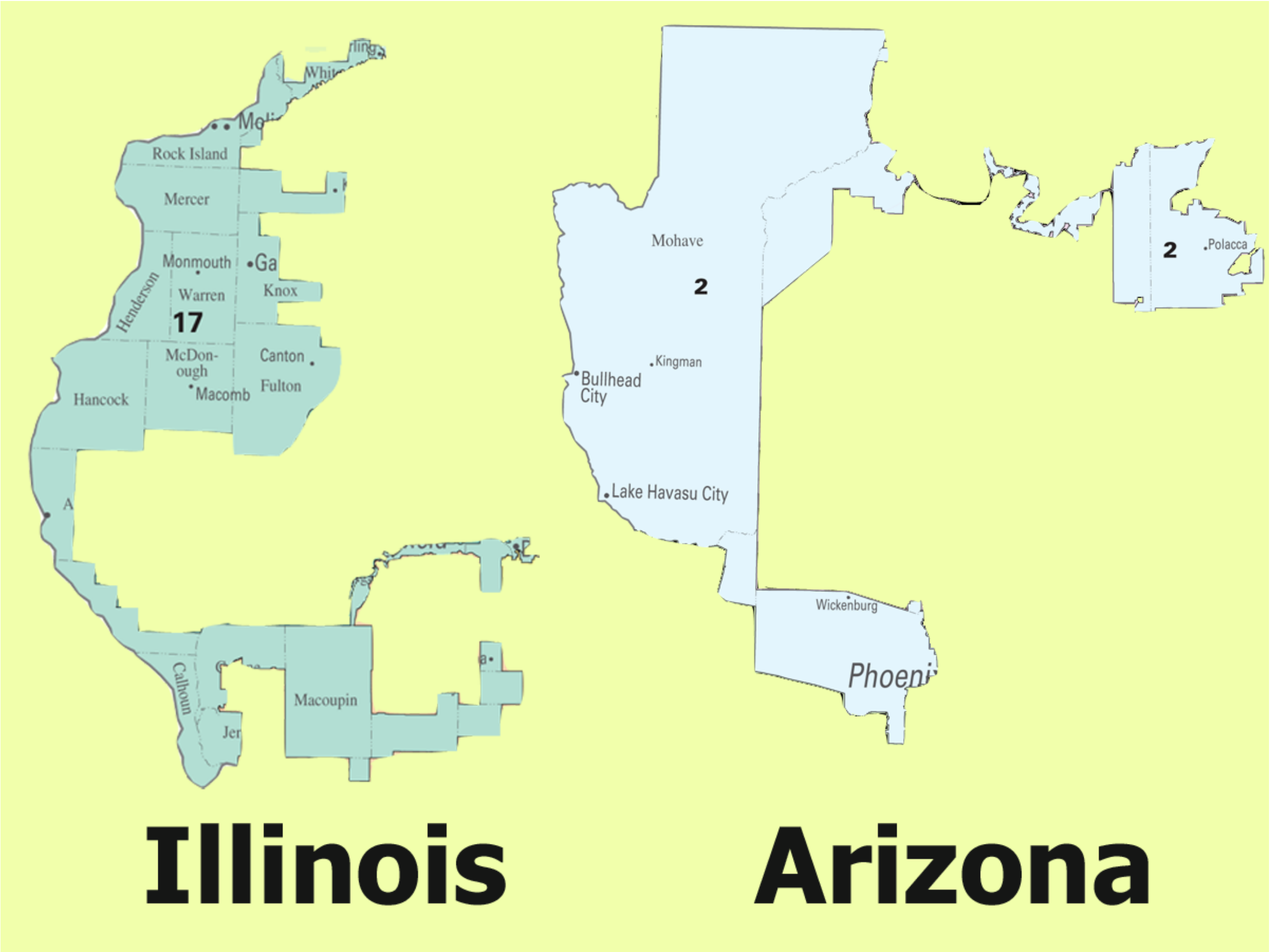
Outline

1. Introduction
2. Motivation
3. Simplifying Assumptions
4. Moment of Inertia Method
5. Diminishing Halves Methods
6. Quantitative Compactness Analysis
7. Conclusion

Problem Statement: Congressional Apportionment

- We wish to draw congressional districts for a state.
- Goal: Algorithm that avoids Gerrymandering.
- Want to create "simplest" shapes.
- Definition of "simple" left to problem solvers.
- Only rule is that districts have equal population.

Gerrymandering Examples



Adapted from *National Atlas of the United States*.

Motivation

Many possible criteria suggested in literature.

- Equality of district size
- Compactness
- Contiguity
- Similarity to existing borders
- Targeted homogeneity/heterogeneity

Instead of specifying many properties, we wish to explicitly specify as few as possible. Additional properties become an **emergent behavior**.

Motivation

Our algorithms will use only the criteria of

1. **Equal population districts**
2. **Compactness**

Examine map afterwards to determine emergent properties.

Simplifying Assumptions

We make the following simplifying assumptions in our model:

- A 2% error tolerance from the mean in size of district population is acceptable.
- Euclidean geometry: variations in longitudinal spacing negligible.
- County borders not sacred (ratio of counties:districts in many states necessitates cutting between borders).

Extracting test data

- Perl script extracts US Census data at the census tract level.
- Discretizes problem into points with a latitude, longitude, and population.
- For New York, 6398 tracts of nonzero population, median 2518 people.

Test Data

State	Population	Districts	Non-empty Census Tracts
TX	20,851,820	32	7530
NY	18,976,457	29	6398
IL	12,419,293	19	8078
AZ	5,130,632	8	1934

Algorithms for Fair Apportionment

We will now compare two methods for apportioning districts

1. Moment of Inertia Method
2. Diminishing Halves Method (Recursive Splitting)

Moment of Inertia Method

Intuition: **Minimize the expected squared distance between all pairs of people in a district.**

- Equivalent to physics concept of moment of inertia.
- One of the first methods to appear in literature (*Hess 1965*)
- Results in districts that are not only connected, but also convex, which limits the possibilities of oddly shaped districts

Moment of Inertia Method

- Moment of Inertia given by $\sum_i P_i(X_i - \bar{X})^2$.
- Equals $\mathtt{min}_Y \sum_i P_i(X_i - Y)^2$.
- Define districts and centers.
- Given centers optimize districts by computing "size" parameters.
- Given districts optimize centers as centers of mass.
- Gives iterative method to converge to near optimal solution.

Diminishing Halves Methods

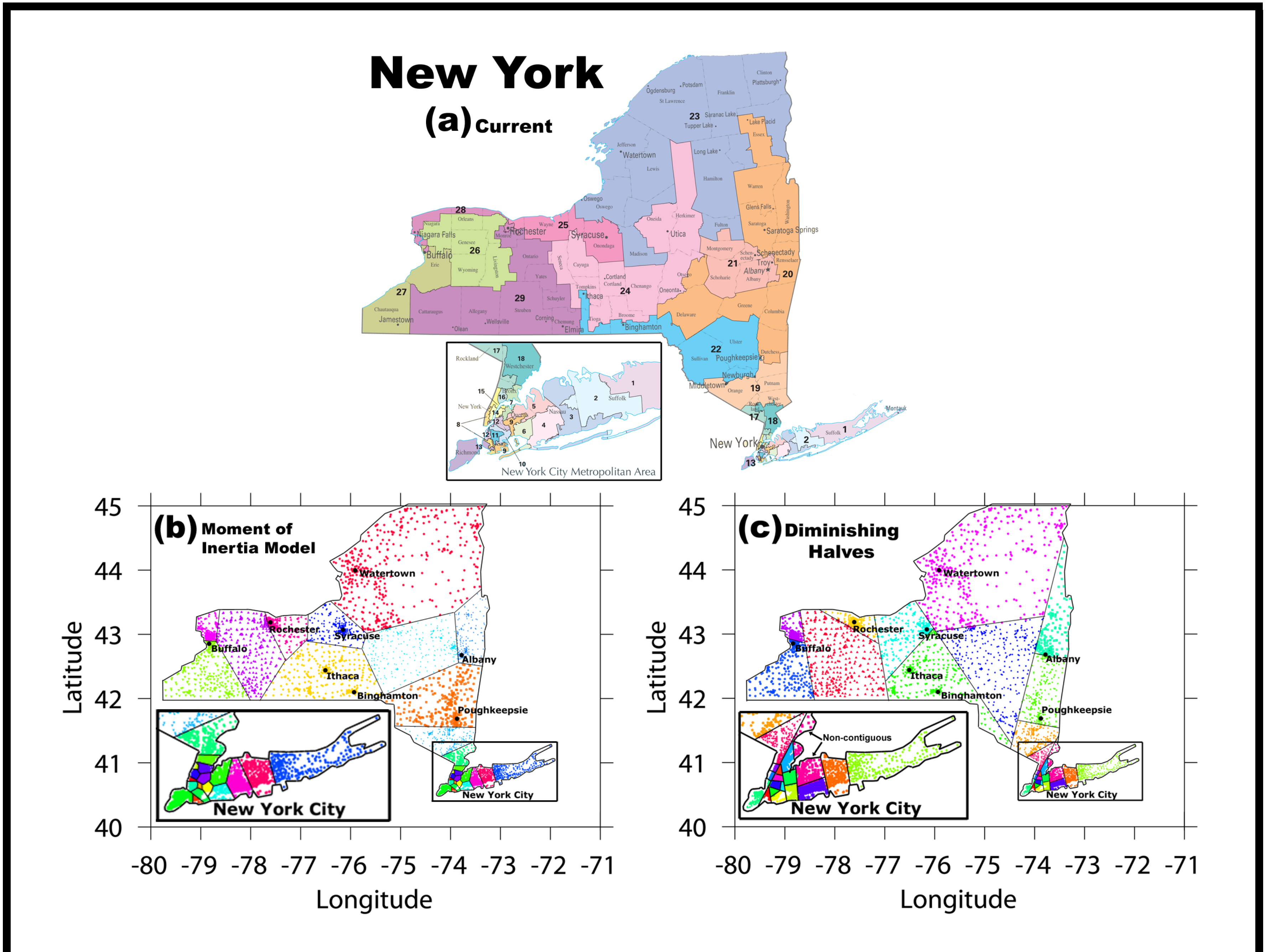
Intuition: **Divide the state into two halves of equal population. Then divide those halves recursively.**

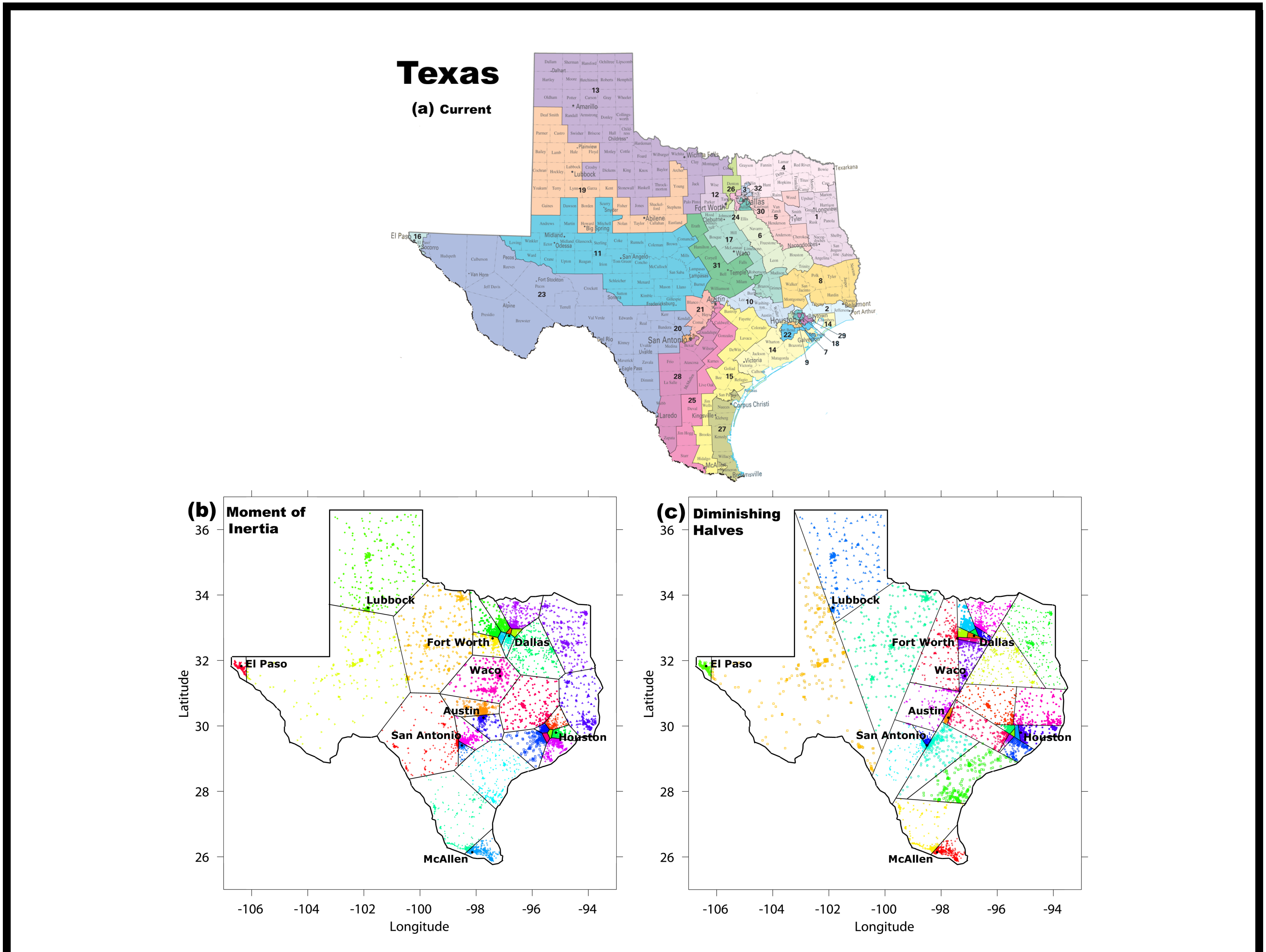
- Suggested by *Forrest 1964*.
- Many different ways to divide state in half.
- After experimentation, we choose the following:
  1. Treat the population data as regression coordinates.
  2. Compute the best fit squared distance (bfsd) line to the data.
  3. Use a line with slope perpendicular to this bfsd line to divide the state.

Diminishing Halves Methods

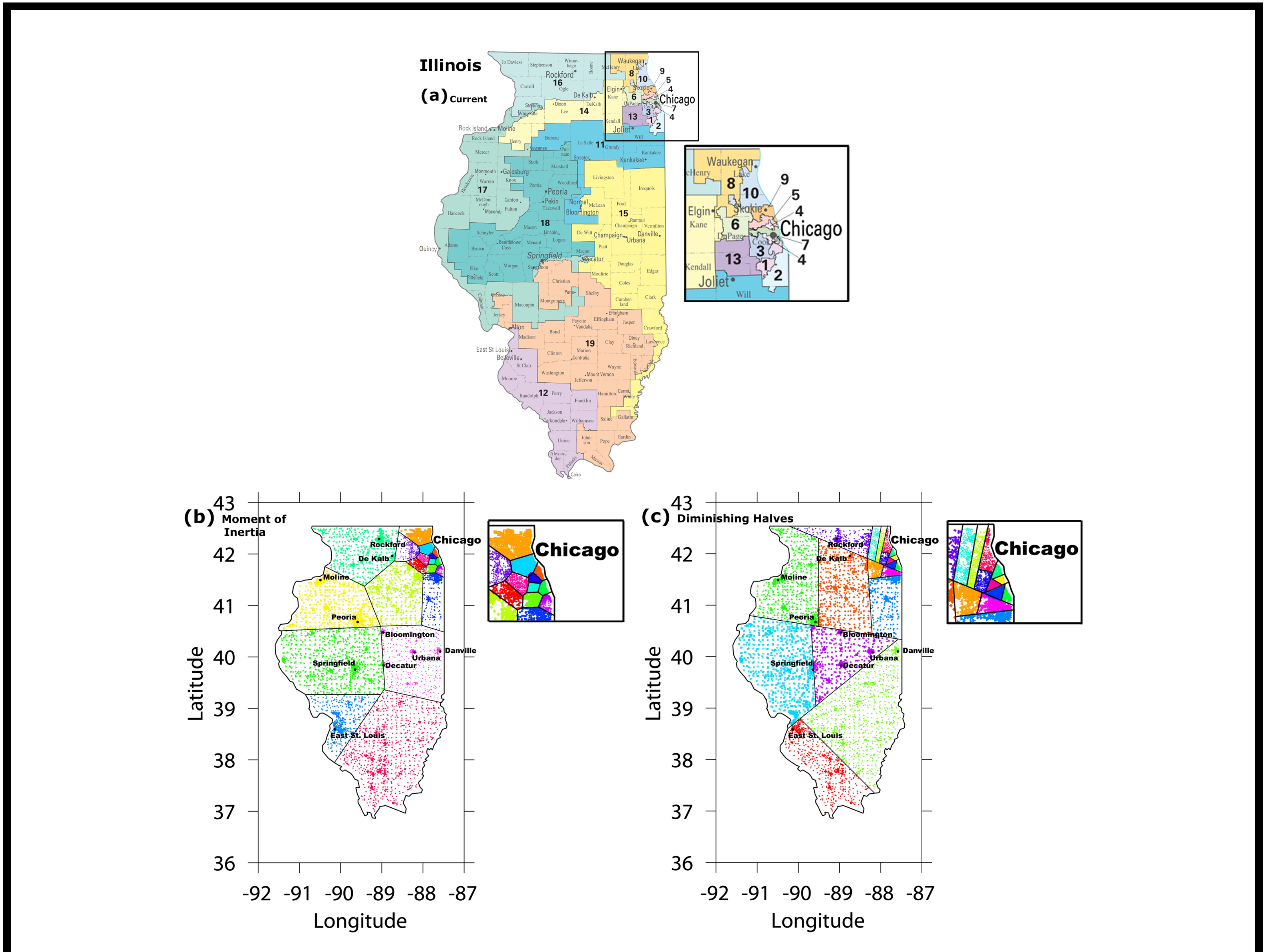
Compute best fit squared distance line based on orthogonal distance to data.

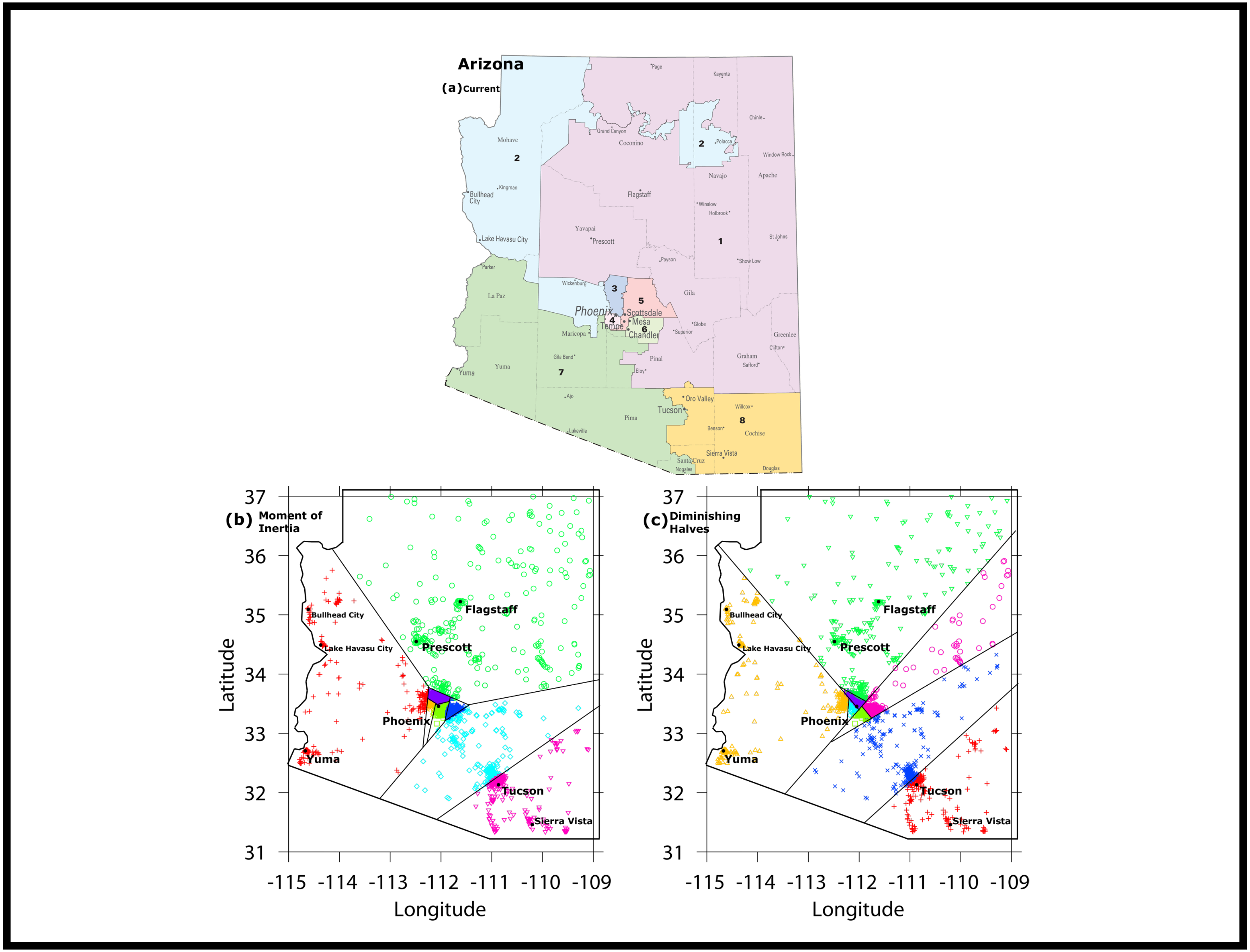
- Best fit line for given slope contains center of mass.
- Line of form $(X - \overline{X})\sin\theta + (Y - \overline{Y})\cos\theta = 0$.
- Line minimizes
  $\mathrm{E}\left[\left((X - \overline{X})\sin\theta + (Y - \overline{Y})\cos\theta\right)^2\right] =$
  $\sin^2\theta \mathrm{Var}[X] + 2\sin\theta\cos\theta \mathrm{Cov}(X,Y) + \cos^2\theta \mathrm{Var}[Y]$.
- Optimal angle satisfies $\tan(2\theta) = \frac{-2\mathrm{Cov}(X,Y)}{\mathrm{Var}[X] - \mathrm{Var}[Y]}$.


New York
(a) Current
New York City Metropolitan Area
(b) Moment of Inertia Model
(c) Diminishing Halves
Watertown
Rochester
Syracuse
Buffalo
Albany
Ithaca
Binghamton
Poughkeepsie
New York City
Non-contiguous
Latitude
Longitude
45
44
43
42
41
40
-80 -79 -78 -77 -76 -75 -74 -73 -72 -71



Texas
(a) Current
(b) Moment of Inertia
(c) Diminishing Halves
Lubbock
El Paso
Fort Worth
Dallas
Waco
Austin
San Antonio
Houston
McAllen
Latitude
Longitude


Illinois
(a) Current
Chicago
(b) Moment of Inertia
(c) Diminishing Halves
Rockford
De Kalb
Moline
Peoria
Bloomington
Danville
Urbana
Springfield
Decatur
East St. Louis
Latitude
Longitude


Arizona
(a) Current
(b) Moment of Inertia
(c) Diminishing Halves
Flagstaff
Prescott
Bullhead City
Lake Havasu City
Phoenix
Yuma
Tucson
Sierra Vista
Latitude
Longitude

Quantitative Measures of Compactness

- **Inverse Roeck Test**. Let C be the smallest circle containing the region, R. Measure $\frac{Area(C)}{Area(R)}$.
- **Schwartzberg Test**. Compute $\frac{Perimeter}{\sqrt{4\pi Area}}$.
- **Length-Width Test**. Inscribe the region in the rectangle with largest length-to-width ratio. Calculate $\frac{Length\ of\ Rectangle}{Width\ of\ Rectangle}$.

Smaller numbers are better. Circles give optimal value 1.

See *Young 1988* for a review.

Compactness Results

Districts	Inv. Roeck	Schwartzberg	Length-Width
NY (MoI)	2.29 ± 0.66	1.64 ± 0.62	1.91 ± 0.61
NY (DH)	2.50 ± 0.87	1.74 ± 0.69	1.91 ± 0.77
TX (MoI)	2.04 ± 0.64	1.14 ± 0.09	1.72 ± 0.57
TX (DH)	2.76 ± 1.66	1.27 ± 0.20	2.30 ± 1.73
IL (MoI)	1.90 ± 0.36	1.28 ± 0.26	1.55 ± 0.39
IL (DH)	2.49 ± 0.99	1.35 ± 0.24	2.01 ± 0.96
AZ (MoI)	2.18 ± 0.56	1.17 ± 0.08	1.77 ± 0.51
AZ (DH)	2.69 ± 0.91	1.29 ± 0.15	2.07 ± 0.79

MoI = Moment of Inertia, DH = Diminishing Halves.

Smaller numbers correspond to more compact districts.

Conclusion

- Obtain nice districts despite few specified criteria.
- Both methods give regions that are not only connected but also **convex**.
- Processing data at the census tract level is feasible.
- Moment of Inertia method gives better results than our Diminishing Halves.
- Moment of Inertia derived from physics intuition.