Electoral Redistricting with Moment of Inertia and Diminishing Halves Models

Andrew Spann
Daniel Kane
Dan Gulotta
Massachusetts Institute of Technology
Cambridge, MA

Advisor: Martin Z. Bazant

Summary

We propose and evaluate two methods for determining congressional districts. The models contain explicit criteria only for population equality and compactness, but we show that other fairness criteria such as contiguity and city integrity are present, too.

The Moment of Inertia Method creates districts whose populations are within 2% of the mean district size, minimizing the sum of the squares of distances between the district's centroid and each census tract (weighted by population size). We prove that this model gives convex districts.

In the Diminishing Halves Method, the state is recursively halved by lines perpendicular to best-fit lines through the centers of census tracts.

From U.S. Census 2000 data, we extract the latitude, longitude, and population count of each census tract. By parsing data at the tract level instead of the county level, we model with high precision. We run our algorithms on data from New York as well as Arizona (small), Illinois (medium), and Texas (large).

We compare the results to current districts. Our algorithms return districts that are not only contiguous but also convex, aside from borders where the state itself is nonconvex. We superimpose city locations on district maps to check for community integrity. We evaluate our proposed districts with various quantitative measures of compactness.

The initial conditions do not greatly affect the Moment of Inertia Method. We run variants of the Diminishing Halves Method and find that they do not

The UMAP Journal 28 (3) (2007) 281–299.

improve over the original. Based on our results, district shapes should be convex, and city boundaries and contiguity can be emergent properties, not explicit considerations. We recommend our Moment of Inertia Method, as it consistently performed the best.

Assumptions and Justifications

About States

- **The Earth's geometry is Euclidean**. No state is so big that the spherical shape of the earth significantly distorts distance calculations obtained from Euclidean geometry.
- **County lines are not inherently more significant than other boundaries**. Some states attempt to not split counties when determining districts, and other states give only slight consideration to county lines. Since several of New York's counties are too big to use as discrete units for dividing representatives, and representing county boundaries in the model is difficult, we instead use the census tract as our base unit of population.
- **Deviation from the current district division is not a major factor**. There are no inherent transitional problems with switching to a completely new division if it can be shown to have a higher degree of fairness.
- **District populations may vary by as much as 2% from the average value**. We use the 2% allowance to get around problems with our data on populations not being fine enough. The error could be made smaller if census blocks were used instead of census tracts.

About Census Data

- **Census data are always accurate**. There are no other reasonable data.
- **Census tracts individually satisfy fair apportionment criteria**. No U.S. census tract is gerrymandered; there is no political benefit to doing so.
- **All population in a census tract can be approximated as located at a single point**. For data input to our program, we read in latitude and longitude for each census tract. We assume that the entire population of a census tract is located at this point. Since we have 6,398 census tracts for New York State, none of which have more than 4% of the population for a congressional district (and most of which are considerably smaller), this does not provide a very severe discretization problem.

Literature Review

Gerrymandering has attracted scholarly attention for decades.

Attempts to assign districts with computers began in the 1960s and 1970s with models by Hess et al. [1965], Nagel [1965], and Garfinkel and Nemhauser [1970]. These methods typically represent population as a series of weighted (x, y) coordinates and attempt to draw equal-population districts based on compactness and contiguity. The methods criteria for compactness vary, and a collection of compactness metrics is reviewed in Young [1988]. Computer resources were limited in this era, and Garfinkel and Nemhauser even report being unable to compute a 55-county state. For a more detailed review of early papers, see Williams [1995].

Many versions of the redistricting problem are NP-hard [Altman 1997]. Recent papers have tried graph theory [Mehrotra et al. 1988], genetic algorithms [Bacao et al. 2005], statistical physics [Chou and Li 2006], and Voronoi diagrams [Galvao et al. 2006]. There are also papers such as Cirincione et al. [2000], which are intellectual and stylistic descendants of the old papers but using modern computer resources that enable finer population blocks and tighter convergence criteria.

We use a moment-of-inertia model similar in formulation to Hess et al. [1965] but with differences in optimization.

Criteria for Fair Districting

We list factors considered in districting and explain which ones we choose. Later, we describe the specific expressions of these criteria in our two models and their mathematical consequences. Core criteria are:

- *Equality of population*. The population difference between two districts can vary only by at most a certain number of people, usually on the order of 5%.
- *Contiguity*. Each district must be topologically simply connected.
- *Compactness*. There are differing opinions on how to quantitatively define compact, but all agree that small wandering branches are bad.

A criterion not emphasized in the literature is *convexity*, a stronger form of contiguity: Any two points in the region can be connected by a straight line segment contained within the region. This disallows holes or extraneous arms that contribute to most poorly-shaped districts. The worst case for a convex region is a district containing sharp angles or that is very elongated.

Other criteria [Nagel 1965; Williams 1995] serve one of two purposes:

- *Targeted homogeneity or heterogeneity*. Nagel explicitly expresses a desire to use predicted voting data to create "safe districts" and "swing districts," where the outcomes of elections are more predictable or less predictable, respectively. The stated reasons for this involve balancing the state's districts so that some parts of a state have experienced candidates who are stable to long term change and other parts more responsive. Other papers discuss clustering groups based on race, economic status, age, or other demographic data into a district where statewide minorities have a local majority.

- *Similarity to boundaries or precedent*. Whenever possible, people in the same city should have the same representative. Likewise, it can be viewed as unfair to a representative if the people represented change too quickly. It also makes sense for districts to follow rivers, lakes, mountains, and other natural boundaries where appropriate. Usually, boundary of precedent objectives are accomplished by keeping county boundaries intact whenever possible.

These optional criteria conflict with the earlier core criteria.

The explicit criteria should be as minimalist as possible, so that more-complicated measures of good districting emerge rather than be forced. Additionally, with complicated objectives, politicians could gerrymander by tweaking parameters of the objective function.

We explain why we do not include the two optional criteria listed above:

- We do not consider targeted homogeneity or heterogeneity criteria because we consider it highly unethical to write a computer program to draw districts that benefit a particular candidate or party, even if the stated reasons appear well-intentioned. The goal of computer assignment of districts is to eliminate all manipulations of this form, so including criteria of this form in the objective function is unacceptable.
- Although the use of existing county or natural boundaries might work well for small states with a high ratio of counties to congressional districts, the county borders of New York are ill-suited for this purpose. New York has only 62 counties but 29 representatives. Following county borders whenever possible but splitting counties where reasonable involves much more work in preprocessing data to incorporate county information and still places pressure on creating noncompact districts.

We formulate a methodology that involves only *equality of population* and *compactness*.

Moment of Inertia Method

Description

By equality of population, we mean that no district's population should differ by more than 2% from the mean population per district in the state. There does not appear to be any clear court-mandated tolerance for population difference [Williams 1995], so we simply pick a reasonable number that is within the feasibility of computation based on the discretized units of census tracts. We could tighten the bounds further if we were willing to tolerate an increase in computational time and use smaller divisions, such as census blocks.

Young [1988] lists eight different measures of compactness, none of which is perfect. The most intuitive definition is to minimize the expected squared distance between all pairs of two people in a district. This has the nice physical interpretation of being analogous to the moment of inertia (if the distance is Euclidean). Papers such as Galvao et al. [2006] minimize inertia based on travel-time distance (adjusted for roads, lakes, etc) rather than absolute distance; but we consider only absolute distance, which is easier to find. Also, if district borders are affected by travel time, then it is possible to gerrymander by constructing strategic roads or bridges.

Response to Prior Literature Commentary

Young [1988] finds two problems with moment of inertia as a measure of compactness:

- It gives good ratings to "misshapen districts so long as they meander within a confined area."
- There is a significant bias based on the area of the district (the moment of inertia is uses squared distances).

In response to the first objection, we get districts that are not only contiguous but also convex (except where they meet nonconvex state lines). We draw districts where it must be possible to travel between any two points in a district in a straight line without leaving the district. This eliminates the first of Young's concerns, since the cited examples of misshapen districts, such as spirals, that cause moment of inertia to predict poorly all have the property of being nonconvex.

The concerns about bias toward large-area districts is perhaps more serious. If the complaint is true, then the moment of inertia compactness criterion could lead to stretched or awkward urban districts so as to smooth out larger neighboring districts. In our experimental runs, this problem was not severe.

Mathematical Interpretation

We describe the mathematics of the moment-of-inertia criterion and its objective function. We derive an important result: Any local minimum of our objective function should consist of a collection of convex districts (except where the state border is nonconvex).

We use the average squared distance between two people in the same district as a measure of the misshapenness of that district. We assume a Euclidean metric. Let $\mathrm{E}[x]$ and $\mathrm{Var}\, x$ represent the expectation and variance of a random variable x. Let the coordinates of two randomly chosen people in the district be (x_1, y_1) and (x_2, y_2), and let the coordinates of an arbitrary randomly chosen person be (x, y). Then our measure is

$$\mathrm{E}\left[(x_1 - x_2)^2 + (y_1 - y_2)^2\right] = 2\operatorname{Var} x + 2\operatorname{Var} y = 2\,\mathrm{E}\left[|(x, y) - (\overline{x}, \overline{y})|^2\right],$$

where $(\overline{x}, \overline{y})$ is the center of mass of people in the district. Furthermore, this quantity is increased if $(\overline{x}, \overline{y})$ is replaced by another point.

Let there be N people in the state to be divided into k districts. Our objective is equivalent to partitioning the people into k sets $S_1, \dots, S_k$ of equal size, and picking points $p_1, \dots, p_k$ to minimize

$$\sum_{i=1}^{k} \sum_{x \in S_i} d(x, p_i)^2,$$

where d is Euclidean distance. Taking the points p_i to be fixed, we find that even if we allow ourselves to split a person between districts (which we do not do in the actual program), we can recast this as a linear programming problem. We let $m_{x,i}$ be the proportion of x that is in district i. We then have

$$m_{x,i} \geq 0; \qquad \textbf{(1)}$$

and for any x,

$$\sum_i m_{x,i} = 1. \qquad \textbf{(2)}$$

The restriction of district sizes says that for any i, we must have

$$\sum_x m_{x,i} = \frac{N}{k}, \qquad \textbf{(3)}$$

where N is the total population of the state. The objective function is

$$\sum_{x,i} m_{x,i} d(x, p_i)^2.$$

A global minimum exists since $0 \leq m_{x,i} \leq 1$, implying that our domain is compact. By linear programming duality, at the point that minimizes the objective, the objective function can be written as a positive linear combination of the tightly satisfied constraints in the solution. For this linear combination, let C_i be the coefficient of **(3)**, D_x the coefficient of **(2)**, and $E_{x,i}$ the coefficient of **(1)**. We have that C_i and D_x are arbitrary, but $E_{x,i} \geq 0$ with equality unless $m_{x,i} = 0$. Comparing the $m_{x,i}$ coefficients of our objective and this linear combination of constraints, we get that

$$d(x, p_i)^2 = C_i + D_x + E_{x,i}.$$

If $m_{x,i} \geq 0$, then $E_{x,i} = 0$, hence that $E_{x,i} \leq E_{x,j}$. In particular, person x can be only in the district i for which $E_{x,i} = d(x, p_i)^2 - C_i - D_x$ is minimal. Equivalently, they are in the district i for which $d(x, p_i)^2 - C_i$ is minimal. Therefore, for the optimal solution, there are numbers C_i and the ith district is the set of people $\{x : d(x, p_j)^2 - C_j \text{ is minimized for } j = i\}$. Furthermore, these regions are uniquely defined up to exchanging people at the boundaries.

The next thing to note is that the ith district is defined by the equations

$$d(x, p_i)^2 - C_i \leq d(x, p_j)^2 - C_j. \tag{4}$$

Rotating and translating the problem so that $p_i = (0, 0)$ and $p_j = (a, 0)$, and letting $x = (x, y)$, **(4)** reduces to

$$x^2 + y^2 - C_i \leq (x - a)^2 + y^2 - C_j,$$

or

$$2ax \leq a^2 + C_i - C_j.$$

Therefore, each district is defined by a number of linear inequalities. Hence, we have shown that our measure has the nice property that *the optimal districts with fixed p_i are convex*, so any local minimum of our objective function should consist of a partition into convex regions.

Computational Complexity

It would be nice to compute the global optimum, but we probably cannot do so in general. Adapting the linear program above, we wish to minimize

$$\sum_i \operatorname{Var} X_i$$

where X_i is a randomly chosen person in district i. This is equal to

$$\sum_i \left(\frac{k}{N} \sum_x m_{x,i} |\vec{x}|^2 - \frac{k^2}{N^2} \left| \sum_x m_{x,i} \vec{x} \right|^2 \right).$$

Notice that the term

$$\sum_i \sum_x m_{x,i} |\vec{x}|^2 = \sum_x |\vec{x}|^2 \sum_i m_{x,i} = \sum_x |\vec{x}|^2$$

is a constant. Hence, we wish to maximize the sum of the squares of the magnitudes of the centers of mass of the districts. This is an instance of quadratic programming where we try to maximize a positive semidefinite objective function. Since general quadratic programming is NP-hard, it seems likely that it is not easy to find a global maximum for our problem. On the other hand, we have shown that even local maxima have many properties that we want, e.g., convexity. Furthermore, these local maxima are significantly easier to find—e.g., from the quadratic programming formulation using the simplex method.

Unfortunately, the quadratic programming approach leads to an optimization involving kN variables, which can be quite large. Instead, we consider the formulation where to have a local maximum we need to pick p_i and C_i (thus defining our districts by "x goes in the district i for which $d(x, p_i)^2 - C_i$ is minimal") in such a way that the districts have the correct size and so that p_i is the center of mass of the ith district. This will imply that we have a local maximum of the quadratic program, since near our solution (up to first order) our objective function is

$$C - \sum_{x,i} m_{x,i} d(x, p_i)^2, \tag{5}$$

for some constant C. Since we have a global maximum of **(5)**, moving a small amount in any direction within our constraint does not decrease our objective, up to first order. Furthermore, since our objective is positive semidefinite, we are at a local maximum. This formulation is much better, since we are now left with only $3k$ degrees of freedom for k districts.

Comparison to Hess et al.

This procedure is very similar to that of Hess et al. [1965]. They too were attempting the minimize the summed moments of inertia of districts. They also converged on their solution via an iterative technique that alternates between finding the best districts for given centers and finding the best centers for given districts. Our approach differs from theirs in two main points: the method of finding new districts for given centers, and the general philosophy toward achieving exact population equality. Both of these differences stem from our having finer data (Hess et al. used 650 enumeration districts for dividing Delaware into 35 state House and 17 state Senate seats, whereas we have 10 times as many census tracts) and more computational power. We cannot determine exactly what algorithm Hess et al. used to determine optimal districts with given centers other than a "transportation algorithm," possibly the linear programming formulation from earlier (possibly using a min-cost-matching formulation). We have many more census tracts to work with and use an algorithm better adjusted to this problem. We also have different perspectives about what to do to even out population. Our fundamental units are sufficiently small that we can just run our algorithm, adjusting district sizes in a natural way until all districts are within 2% of the desired population. Hess et al. used a solution method that divided fundamental units of population between districts and later had to perform post-iteration checks and alterations so that units were no longer split and population equality still worked out. This readjustment has the potential to increase moments of inertia and could theoretically lead to a failure to converge.

Diminishing Halves Method

As an alternative against which to compare our moment of inertia algorithm, we use the Method of Diminishing Halves proposed by Forrest [1964].

Definition

The Diminishing Halves Method splits the state into two nearly-equal-sized districts and recurses on each of the two halves. The idea is to split into relatively compact halves. Forrest does not specify exactly how the state must be split into two halves at each step, but rather argues that the method for splitting the state in two could be adjusted based on preferences for keeping counties intact or other goals.

Suppose that we run a least-squares regression on the latitude and longitude coordinates of the state's census tracts. We would expect that dividing along this best-fit line would be a bad idea, since it would probably cut major cities in half or cover a long distance across the state. If we take a line perpendicular to the best-fit line, then hopefully we get the opposite properties. Therefore, we divide the state at each stage with a line whose slope is perpendicular to the best-fit line of the census tracts. We are not aware of this specific criterion being used in previous literature.

Mathematical Interpretation

The best-fit line is an approximation of the shape of the state is of the form

$$(X - \overline{X}) \sin\theta + (Y - \overline{Y}) \cos\theta = 0.$$

The left-hand side is the distance of a point (X, Y) from the line. We minimize

$$\mathrm{E}\left((X - \overline{X}) \sin\theta + (Y - \overline{Y}) \cos\theta\right)^2 = \\ \sin^2\theta \operatorname{Var} X + 2 \sin\theta \cos\theta \operatorname{Cov} XY + \cos^2\theta \operatorname{Var} Y.$$

This value is minimized when

$$\sin\theta \cos\theta(\operatorname{Var} X - \operatorname{Var} Y) + (\cos^2\theta - \sin^2\theta) \operatorname{Cov} XY = 0,$$

or when

$$\tan(2\theta) = \frac{-2 \operatorname{Cov} XY}{\operatorname{Var} X - \operatorname{Var} Y}.$$

To divide the population into k districts, we divide the state by a line perpendicular to the best fit that splits the population in ratio $\left\lfloor \frac{k}{2} \right\rfloor : \left\lceil \frac{k}{2} \right\rceil$. When we need to divide into an odd half and an even half, the ceiling half goes to the southern side.

Experimental Setup

Extraction of U.S. Census Data

We used a Perl script to extract data at the census tract level from the 2000 U.S. Census [U.S. Census Bureau 2001]. For New York, there are 6,661 tracts in the database, 6,398 of which have nonzero population. We extract the population along with the latitude and longitude of a point from each district. The districts have populations from 0 to 24,523 with a median of 2,518. We model the population density by assuming that the entire population of a tract is located at the coordinates given. We adjust for the fact that one degree of latitude and one degree of longitude represent different lengths on the Earth's surface by having our program internally multiply all longitudes by the cosine of the average latitude. We also extracted data for Arizona (small—8 congressional representatives), Illinois (medium—19 representatives), and Texas (large—32 representatives).

Implementation in C++

We use a C++ program to compute an approximate local minimum of our Moment of Inertia objective function. We do so without splitting census tracts between districts, and this discretization requires us to allow a little lenience about the exact sizes of our districts (we allow them to vary from the mean by as much as 2%).

We attempt to converge to a local optimum via two steps. First we pick guesses for the points p_i. We then numerically solve for the C_i that make the district sizes correct, giving us some potential districts. We allow a variation of 2% from the mean, beginning with a 20% allowable deviation in the first few iterations and tightening the constraint on subsequent iterations. We then pick the center of mass of the new districts as new values of p_i, and repeat for as long as necessary. Each step of this procedure decreases the quantity in **()**, because our two steps consist of finding the optimal districts for given p_i and finding the optimal p_i for given districts. We find the correct values of C_i by alternately increasing the smallest district and decreasing the largest one. When this adjustment overshoots the necessary value, we halve the step size for that district, and when it overshoots by too much, we reverse the change. For New York, convergence to the final districts took a couple of minutes.

After determining our districts, we output them to a PostScript file that displays the census tracts color-coded by district, so that one can visually determine compactness. Finally, we computed some of the compactness measures discussed in [Young 1988].

We also created a C++ program to implement the Diminishing Halves Method.

Measures of Compactness

We need an objective method for determining how successful our program is at creating compact districts. Young [1988] gives several measures for the compactness of a region. We use some of these to compare our districts with those produced by other methods. Since our algorithms generate convex districts except where the state border is nonconvex, we perform all of these results on the convex hull of our districts, so that the test results are not unfairly affected by awkwardly-shaped state borders.

Definitions

Inverse Roeck test. Let C be the smallest circle containing the region R. We measure Area(C)/Area(R), a number larger than 1, with smaller numbers corresponding to more compact regions. This is the reciprocal of the Roeck test as phrased in Young. We have altered it so that all of our measures in this section have smaller numbers corresponding to more compact regions.

Length-Width test. Inscribe the region in the rectangle with largest length-to-width ratio. This ratio is greater than 1, with numbers closer to 1 corresponding to more compact regions.

Schwartzberg test. We compute the perimeter of the region divided by the square root of 4π times its area. By the isoperimetric inequality, this is at least 1 with a value of 1 if and only if the region is a disk. This test considers a region compact if the value is close to 1.

Calculation in Mathematica

We used the tests above to check compactness of the proposed districts. We implemented the tests in Mathematica with aid of the Convex Hull and Polygon Area notebooks [Weisstein 2004; 2006].

For the Roeck test, we compute the area of the polygon by triangulating it. We find the circumradius by noting that if every triple of vertices can be inscribed in a disk of radius R, then the entire polygon fits into the disk. This is because a set of points all fit in a disk of radius R centered at p if and only if the disks of radius R about these points intersect at p. Let D_i be the disk of radius R centered around the ith point. If every triple of points can be covered by the same disk, then any three of the D_is intersect. Therefore, by Helly's theorem [Weisstein 1999] all the disks intersect at some point, and hence the disk of radius R at this point covers the entire polygon. Hence, we need for any three points the radius of the disk needed to contain them all. This is either half the length of the longest side if the triangle formed is obtuse, or the circumradius otherwise.

For the Length-Width test, we pick potential orientations for our rectangle in increments of $\pi/100$ radians. At each increment, we project our points parallel and perpendicular to a line with that orientation. The extremal projections determine the bounding sides of our rectangle. We choose the value from the orientation that yields the largest length-to-width ratio.

Calculating the Schwartzberg test is straightforward.

Results for New York

Figure 1 presents maps of the Moment of Inertia Method districts, the districts from the Diminishing Halves Method, and the actual current congressional districts of New York.

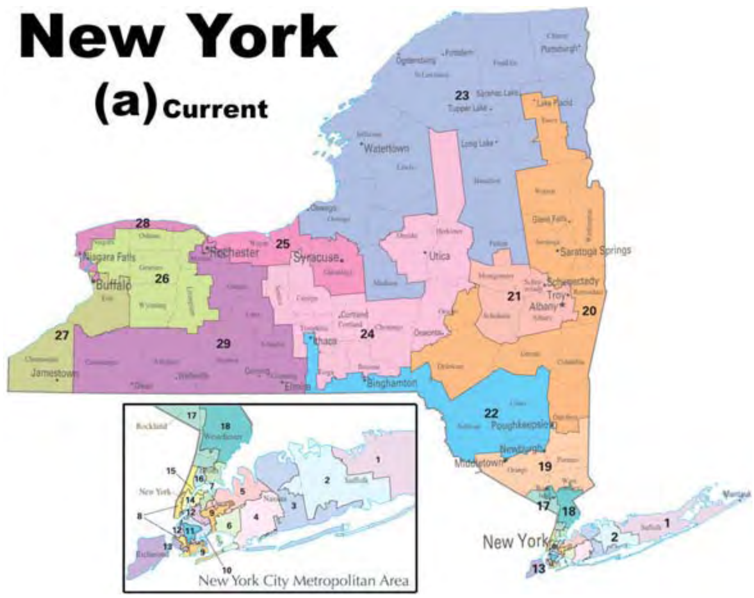
Our program's raw output plots the latitude and longitude coordinates of each census tract using a different color and symbol for each district. The state border and black division lines are added separately. There appears to be a slight color bleed across the borderlines near crowded cities, but this is due to the plotting symbols having nonzero width. Zooming in on our plot while the data are still in vector form (before rasterization) shows that our districts are indeed convex.

Discussion of Districts

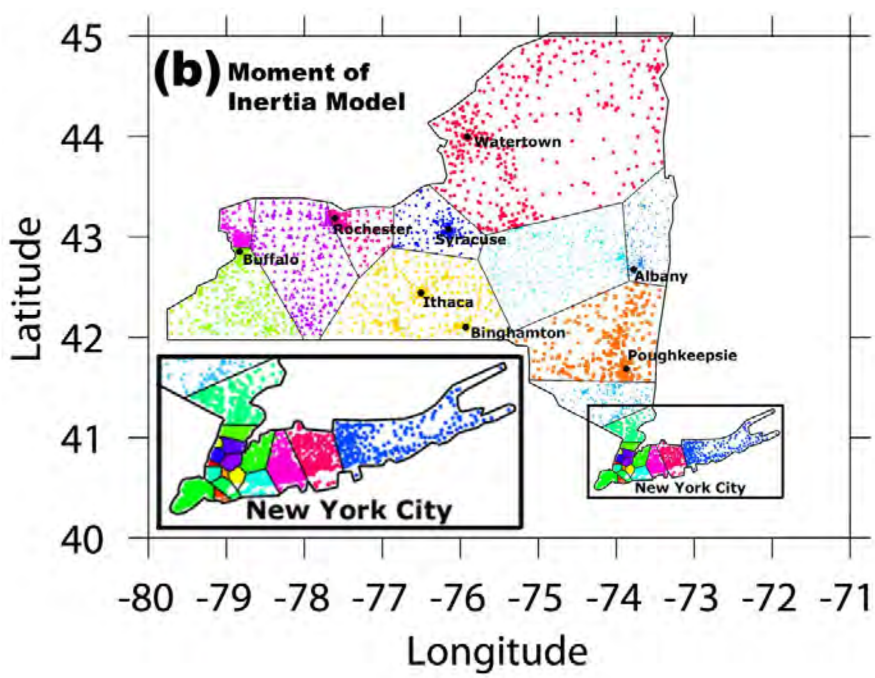
Both methods produce more compact-looking results than the current districts. Some current New York districts legitimately try to respect county lines, but there are a few egregious offenders, such as Districts 2, 22, and 28, where the boundaries conform to neither county lines nor good compactness. The current District 22 has a long arm that connects Binghamton and Ithaca, and the current District 28 hugs the border of Lake Ontario to connect Rochester with Niagara Falls and the northern part of Buffalo. Both of our methods allow Ithaca and Binghamton to be in the same district, but without stretching the district to the land west of Poughkeepsie. Buffalo and Rochester are kept separate in both of our models.

Our data do not contain information about county lines. However, both of our methods do a good job at keeping the major cities of New York intact. Buffalo and Rochester are divided into at most two districts in our methods, instead of three under the current districting. The Diminishing Halves Method has a cleaner division for Rochester, but the Moment of Inertia Method handles Syracuse much better.

Both methods produce districts with linear boundaries. The Diminishing Halves Method has a tendency to create more sharp corners and elongated districts, whereas the Moment of Inertia Method produces rounder districts. The Diminishing Halves Method tends to regions that are almost all triangles and quadrilaterals. Where three districts meet with the Diminishing Halves Method, the odds are that one of the angles is a 180° angle. The Moment of Inertia Method does a better job of spreading out the angles of three intersecting regions more evenly and thus results in more pleasant district shapes.

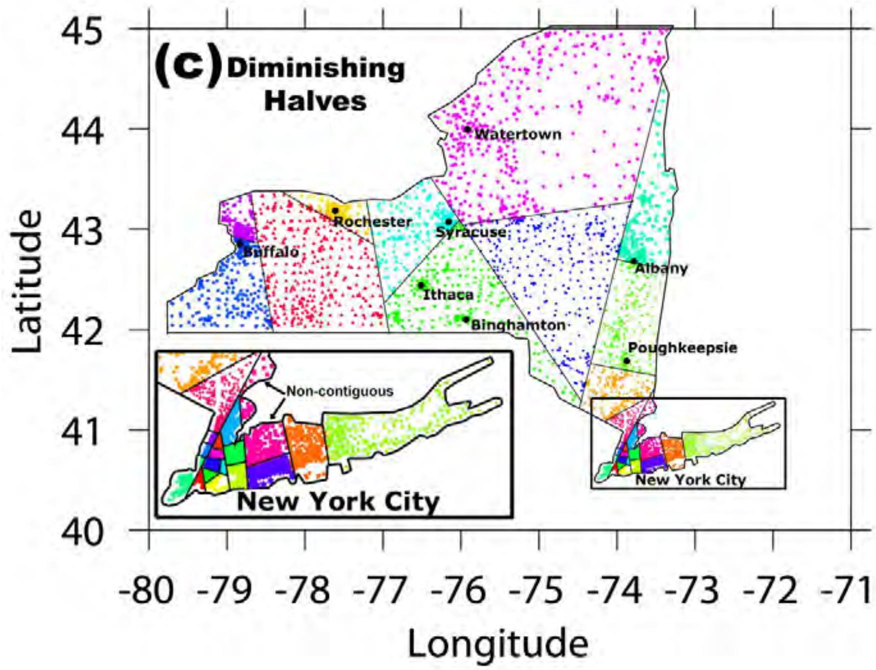


a. Current (adapted from U.S. Department of the Interior [2007]).



b. Moment of Inertia Method.

Figure 1. New York districts.



c. Diminishing Halves Method.

Figure 1 (continued). New York districts.

That Greater New York City contains roughly one-half of the state's population is convenient for the Diminishing Halves Method. However, that algorithm does not deal very well with bodies of water. This feature leads to the creation of one noncontiguous district (marked as "Non-contiguous" in Long Island Sound in **Figure 1c**. Overall, the shapes given by the Moment of Inertia Method look rounder and more appealing.

Compactness Measures

Table 1 lists results of the compactness tests. Smaller numbers correspond to more compact regions.

Table 1.

Mean and standard deviation for compactness measures of districts; smaller is better.

Districts	Inverse Roeck Test	Schwartzberg Test	Length-Width Test
NY (Moment of Inertia)	2.29 ± 0.66	1.64 ± 0.62	1.91 ± 0.61
NY (Diminishing Halves)	2.50 ± 0.87	1.74 ± 0.69	1.91 ± 0.77

According to these measures, the Moment of Inertia Method does marginally better than the Diminishing Halves Method. The diminishing halves numbers appear to be larger by about one-seventh of a standard deviation. This probably is caused by a few of the more misshapen districts.

All three measures are calibrated so that the circle gives the perfect measurement of 1. Roughly speaking, the Roeck test measures area density, the Length-Width test measures skew in the most egregious direction, and the Schwartzberg test measures overall skewness. Each measure tells us approximately the same thing: the Moment of Inertia Method performs a little bit better than the Diminishing Halves Method.

It would be desirable to compare the numbers in **Table 1** to the current districts, but there are two reasons why we cannot do this:

- The data that we used do not offer congressional district identification at the census-tract level. To compute compactness, we would need to choose a finer population unit, hence the numbers would not be directly comparable to those in **Table 1**.
- All of our districts in both methods are convex except for where the state border is nonconvex. This is not true for the current districts, and it is unclear how useful the compactness numbers are at comparing convex districts to nonconvex districts.

Results for Other States

To test how well our algorithms perform on states with different sizes, we also computed districts for Arizona (small—8 districts), Illinois (medium—19 districts), and Texas (large—32 districts). [EDITOR'S NOTE: We omit the corresponding figures and specific analysis.]

Compactness Measures

The Diminishing Halves Method produces consistently worse results by all three measures. This fact suggests (and the maps seem to confirm) that this fault is due largely to producing a small number of very elongated districts.

Given the evidence, we recommend the Moment of Inertia Method over the Diminishing Halves Method.

Sensitivity to Parameters

To test for robustness, we tweak some of the parameters to the Moment of Inertia Model and test variants of the Diminishing Halves Method.

Initial Condition

We ran the Moment of Inertia Model on each of the states with three different random seeds. The results were almost identical each time.

Population Equality Criterion

We ran the New York case of the Moment of Inertia Model using a 5% allowable deviation from the mean in district population instead of a 2% allowable deviation. We observed no significant change in the results.

Variants of the Diminishing Halves Method

We modified our criterion for determining the dividing line in the Diminishing Halves Method to use a mass-weighted best-fit line, weighted to account for different census tracts containing different numbers of people. We ran this modified method on New York, Arizona, and Illinois. We also tried a modification of the Diminishing Halves Method on the New York case that draws vertical and horizontal (longitude and latitude) lines.

In all these modified cases, results were visibly much worse. The modified methods tended to split cities into more districts than the original method.

Strengths and Weaknesses

Strengths:

- **Emergent behavior from simple criteria**. We specify criteria only for population equality and compactness. We satisfy contiguity and city integrity without explicitly trying to do so.
- **Simple, intuitive measure of complexity of districts**. In the Moment of Inertia Method, our measure of the noncompactness of a district gives a model that is easy to understand and does not use arbitrary constants that could be tuned to gerrymander districts.
- **Results in convex districts**. Both models produce districts guaranteed to be convex, aside from where the state border is nonconvex. This provides a fairly strong argument for the compactness of the resulting districts.
- **Easily computable**. Our final districting can be computed in a few minutes.
- **Nice-looking final districts**. The districts that we get appear very nice.

Weaknesses:

- **No theoretical bounds on convergence time**. We could not prove that our algorithm converges in reasonable time, although it has done so in practice.

- **Potential for elongated smaller districts**. Some of the smaller districts produced by the Moment of Inertia Method may be stretched to accommodate larger districts. The Diminishing Halves Method may not correctly divide regions such as discs or squares that are not described well by a best-fit line.

- **Does not respect natural or cultural boundaries**. Our algorithms do not take natural or cultural boundaries into account. Doing so would have the advantage of not having district boundaries cross rivers but could place pressure on making districts noncompact and allow for loopholes that could be exploited by malicious politicians.

- **Does not necessarily find the global optimum**. Our Moment of Inertia algorithm finds only a local minimum. This leads potentially to some nondeterminism in the resulting districts, which could allow gerrymandering; but the amount is small.

- **Can only draw new districts, not determine if existing districts are gerrymandered**. Cirincione et al. [2000] give a pseudoconfidence interval analysis to assess whether South Carolina's 1990 redistricting had been gerrymandered. We do not perform such analysis here.

Conclusion

We formulated and tested two methods for assigning congressional districts with a computer.

The Moment of Inertia Method searches for the answer that satisfies the intuitive criterion that people within the same district should live as close to each other as possible. We implemented this method and obtain results that would not have been computationally feasible in the 1960s and 1970s.

The Diminishing Halves Method recursively divides the population in half, which is very simple to explain to voters. To avoid elongated districts and to cut along sparsely populated areas rather than densely populated regions, our implementation chooses a dividing line perpendicular to the statistical best-fit line through the latitude and longitude coordinates of the census tracts.

We have some concrete recommendations for state officials:

- **Processing data at the census tract level or finer is computationally feasible**. It would not be unreasonable to process at the block group level if the extra resolution would be beneficial.

- **Districts should be convex**. Most models in the literature check only for contiguity. However, even severely gerrymandered districts such as Arizona District 2 satisfy contiguity. Requiring all districts to be convex greatly reduces the potential for political abuse.

- **City boundaries and contiguity of districts should be emergent properties, not explicit considerations**. Neither of our methods explicitly requires districts to be contiguous, yet the districts they generate are not only contiguous but convex. Neither of our methods attempts to preserve city or county boundaries, yet the Moment of Inertia Method does a good job of keeping cities together whenever reasonable. It is probably sensible for smaller states with a high ratio of counties to congressional representatives to be concerned with county boundaries; but for New York, where there are comparatively few counties, looking at city integrity instead of county integrity is more reasonable.

- **A good algorithm can handle states of different sizes**. Algorithms that perform well on large states might not yield good results for a small state with only one or two large cities. We tested our algorithms on states of different sizes; the Moment of Inertia Method behaves well in all cases.

- **We recommend a moment of inertia compactness criterion**. The Moment of Inertia Method, compared to the Diminishing Halves Method,

  – consistently produces more visually-appealing districts,

  – has better results on the compactness tests, and

  – does a better job of respecting city boundaries.

References

Altman, M. 1997. The computational complexity of automated redistricting: Is automation the answer? *Rutgers Computer and Technology Law Journal* 23: 81–136.

Bacao, F., V. Lobo, and M. Painho. 2005. Applying genetic algorithms to zone design. *Soft Computing* 9: 341–348.

Chou, C.I., and S.P. Li. 2006. Taming the gerrymander—Statistical physics approach to political districting problem. *Physica A* 369: 799–808.

Cirincione, C., T.A. Darling, and T.G. O'Rourke. 2000. Assessing South Carolina's 1990s Congressional redistricting. *Political Geography* 19: 189–211.

Forrest, E. 1964. Apportionment by computer. *American Behavioral Scientist* 8 (4) (December 1964): 23, 35.

Galvao, L.C., A.G.N. Novaes, J.E.S. de Cursi, and J.C. Souza. 2006. A multiplicatively-weighted Voronoi diagram approach to logistics districting. *Computers and Operations Research* 33: 93–114.

Garfinkel, R.S., and G.L. Nemhauser. 1970. Optimal political districting by implicit enumeration techniques. *Management Science* 16 (4): B495–B508.

Hess, S.W., J.B. Weaver, H.J. Siegfeldt, J.N. Whelan, and P.A. Zitlau. 1965. Nonpartisan political redistricting by computer. *Operations Research* 13 (6): 998–1006.

Mehrotra, Anuj, Ellis L. Johnson, and George L. Nemhauser. 1998. An optimization based heuristic for political districting. *Management Science* 44 (8): 1100–1114.

Nagel, S.S. 1965. Simplified bipartisan computer redistricting. *Stanford Law Review* 17: 863–899.

U.S. Census Bureau. 2001. Census 2000 Summary File 1 [New York]. `http://www2.census.gov/census_2000/datasets/Summary_File_1/New_York/` .

U.S. Department of the Interior. 2007. National Atlas of the United States. Printable Maps. Congressional Districts—110th Congress. `http://nationalatlas.gov/printable/congress.html` .

Weisstein, Eric W. 1999. Helly's theorem. From *MathWorld*—A Wolfram Web Resource. `http://mathworld.wolfram.com/PolygonArea.html` .

________. 2004. Polygon area. From *MathWorld*—A Wolfram Web Resource. `http://mathworld.wolfram.com/PolygonArea.html` .

________. 2006. Convex hull. From *MathWorld*—A Wolfram Web Resource. `http://mathworld.wolfram.com/ConvexHull.html` .

Williams, J.C. 1995. Political redistricting—A review. *Papers in Regional Science* 74: 13–39.

Young, H.P. 1988. Measuring the compactness of legislative districts. *Legislative Studies Quarterly* 13: 105–115.