
Electoral Redistricting with Moment of Inertia and Diminishing Halves Models

We propose and evaluate two methods for determining congressional districts. The models are defined so that they only explicitly contain criteria for population equality and compactness, but we show through a detailed analysis that other fairness criteria such as contiguity and city integrity are present as emergent properties.

In the Moment of Inertia Method, districts are created such that populations are within 2% of the mean district size and the sum of the squares of distances between each census tract weighted by population size and the district's centroid is minimized. We present a mathematical argument that this model will result in districts that are convex.

In the Diminishing Halves Method, the state is recursively divided in half by a line that is perpendicular to the statistical best-fit line describing the region's census tracts.

With the help of a Perl script we are able to parse US Census 2000 data, extracting the latitude, longitude, and population count of each census tract. By parsing data at the census tract level instead of the county level, we are able to run our model with high precision. We run our algorithms on census data from the states of New York as well as Arizona (small), Illinois (medium), and Texas (large).

We compare the results of our methods to each other and to the current districts in the respective states. Both 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 the district maps to check for community integrity. We evaluate our proposed districts with the Inverse Roeck Test, the Length-Width Test, and the Schwartzberg Test to obtain quantitative measures of compactness.

The initial conditions do not greatly affect the Moment of Inertia Method. We run additional variants of the Diminishing Halves Method and find that they do not improve over our normal method.

Based on our results, we would like to recommend to states that

- District shapes should be convex.
- City boundaries and contiguity can be emergent properties, not explicit considerations.
- A good algorithm can handle states of different sizes.
- We recommend our Moment of Inertia Method, as it consistently performed the best.

Contents

1	Problem Restatement	1
2	Assumptions and Assumption Justifications	1
3	Literature Review	2
4	Criteria for Fair Districting	3
5	Moment of Inertia Method	6
5.1	Description	6
5.2	Response to Prior Literature Commentary	7
5.3	Mathematical Interpretation	8
5.4	Computational Complexity	11
5.5	Comparison to Hess, Weaver, et al.	12
6	Diminishing Halves Method	13
6.1	Definition	13
6.2	Mathematical Interpretation	14
7	Experimental Setup	15
7.1	Extraction of US Census Data	15
7.2	Implementation in C++	15
7.3	Further Analysis in Mathematica	16
8	Measures of Compactness	17
8.1	Definitions	17
8.2	Calculation in Mathematica	17
9	Results for New York	19
9.1	Discussion of districts	19

9.2 Compactness Measures	21
10 Results for Other States	22
10.1 Discussion of Districts	26
10.2 Compactness Measures	27
11 Sensitivity to Parameters	28
11.1 Initial Condition	28
11.2 Population Equality Criterion	29
11.3 Variants of the Diminishing Halves Method	29
12 Strengths and Weaknesses	29
13 Conclusion	31

Appendix A: Full-Page Plots

Appendix B: Computer Code

1 Problem Restatement

Concerns over fair district apportionment have existed for over two hundred years. Given a map and a population distribution, we wish to divide the region into a number of districts with equivalent population size and characteristics such as geographic compactness. The specific factors that should be taken into account in enacting a fair apportionment are themselves unclear and before proposing an algorithm we must first determine what constitutes a fair division of congressional districts. Once the objective function is determined, the geography problem now turns into an optimization problem and we must determine an efficient and robust way to compute district divisions.

2 Assumptions and Assumption Justifications

About States

- **The Earth's Geometry is Euclidean.** We assume that 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 choose instead to use the census tract as our base unit of population.
- **Deviation from the current district division is not a major factor.** The criteria for what constitutes a fair division will be discussed in further detail in Section 4. We assume that there are no inherent transitional problems with switching to a completely new division if the new division can be shown to have a higher degree of fairness.

- **District populations are allowed to vary by as much as 2% from the average value.** This is fairly standard assumption to simplify modeling. In particular, 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 is no other reasonable data set upon which to base district apportionment, so this is a fair assumption.
- **Census Tracts individually satisfy fair apportionment criteria.** We assume that no US Census Tract is gerrymandered. There is no political benefit to gerrymandering a census tract.
- **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 locations for each census tract. We assume that the entire population of a census tract is located at this point. Since we have 6398 Census Tracts for New York, 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.

3 Literature Review

The problem of gerrymandering has attracted scholarly attention for decades. Before introducing our own model, we first provide a very brief overview of documented attempts to apportion congressional districts with computers.

Attempts to assign districts with computers began in the 1960's and 1970's with models such as Hess, Weaver, et al. [8], Nagel [10], and Garfinkel and Nemhauser [7]. These methods

typically represent population as a series of weighted (x, y) coordinates and attempt to draw equal-population districts based on rules of compactness and contiguity. The methods used for determining compactness vary slightly, and a collection of compactness metrics is reviewed in Young [16]. Resources were very limited in this era, and the Garfinkel and Nemhauser paper even reports being unable to compute a 55-county state. For a more detailed review of early papers see Williams [15].

It has been shown that the many versions of the redistricting problem are NP-hard [1]. Modern papers have attempted techniques such as graph theory [9], genetic algorithms [2], statistical physics [3], and Voronoi diagrams [6]. There are also papers such as Cirincione et al. [4] which are intellectual and stylistic descendants of the old papers running with modern computer resources that enable finer population blocks and tighter convergence criteria.

We developed the most intuitive criteria we could think of and came up with a moment of inertia model that winds up being similar in formulation to the method of Hess, Weaver, et al. [8]. However, there are a few differences in the optimization phase of the model between our implementation and Hess's implementation, discussed in further detail in Section 5.5. Our model also has access to much faster computational resources than the early models did, so we believe that our analysis will still be very informative.

4 Criteria for Fair Districting

Before we begin discussing an algorithm for dividing districts, we must first determine our objective function. Determining a fair objective function is of utmost importance, for otherwise computers could be used by malicious politicians to disguise gerrymandering rather than preventing it.

In this section, we list several factors that can be considered in choosing how to divide districts and we explain which ones we choose. Sections 5.3 and 6 describe the specific expressions of these criteria in our two models and their mathematical consequences.

First we enumerate criteria used by previous researchers then we explain which criteria we choose for our research. A quick survey of papers published in scholarly journals [3], [7], [8], [9] reveals that the most essential characteristics in dividing districts are

- *Equality of population*, which states that 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 (discussed in Section 5.1), but all agree that small wandering branches are bad.

A criterion which we will talk about but which does not appear to be emphasized in the literature is that of convexity, a stronger form of contiguity. Convexity is a very strong condition on a region. It requires that any two points in the region can be connected by a straight line segment contained within the region. This disallows a number of things like 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 more debatable criteria, examples of which are discussed further in Nagel [10] or Williams [15], vary in their phrasing but all serve one of two purposes:

- *Targeted homogeneity or heterogeneity*. Nagel's paper 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.* It seems intuitive that whenever possible people in the same city should have the same representative. Likewise, it can be viewed as unfair to representatives if the set of people they are representing changes 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.

In deciding which criterion we wish to incorporate in our model, we use the guiding principle that **the set of explicitly specified criteria for the model should be as minimalist as possible** so that more complicated measures of good districting come as **emergent properties** rather than objectives that we arbitrarily forced the computer to accomplish.

In the field of computer security, it is frequently stated that security and functionality are competing goals. In an effort to make software programs have more functions, opportunities arise for malicious users to exploit vulnerabilities. Analogously, the core set of criteria and the optional set of criteria listed above are in conflict with each other. Additionally, if we start to consider criteria that involve complicated objectives, politicians could be able to gerrymander their state by tweaking the parameters of the objective function. Although one could argue that the mathematical districting problem is made more complicated and thus more interesting by the addition of extra objectives, the end result is a more easily exploitable system that is not guaranteed to actually produce better results. We will explain why we choose not to include the two optional criteria listed above.

First, we do not consider targeted homogeneity or heterogeneity criteria because we consider it highly unethical to directly order 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.

Secondly, 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. Making the compromise that we should follow county borders whenever possible but split counties where reasonable has the problems that it involves much more work in pre-processing data to incorporate county information and that it still places pressure on creating non-compact districts.

We now formulate a methodology that involves only two criteria: we will divide the state into census tracts and assign each census tract to a congressional district while giving explicit consideration **based only upon equality of population and compactness**. We will examine and comment on the other measures retrospectively after first running a program that is blind to them.

5 Moment of Inertia Method

Our flagship model, which we believe is the most intuitive way to apportion districts, seeks to minimize the sum of the moments of inertia of the districts.

5.1 Description

As stated in Section 4, we need to consider district criteria based on equality of population and compactness.

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 [15], 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 block groups.

There are many differing opinions about how to define compactness. Young [16] lists

eight different measures of compactness that can be considered, none of which is perfect. The most intuitive definition of compactness, which occurred to our team before reading any of the literature, is to minimize the expected squared distance between all pairs of two people in any given district. This has the nice physical interpretation of being analogous to the moment of inertia (at least if the distance used is Euclidean). Some papers such as Galvao et al [6] use the variant of minimizing inertia based on the travel-time distance (adjusted for roads, lakes, etc) rather than absolute distance, but we choose to consider only absolute distance between points because 1) absolute distance data is easier to find and 2) if district borders are affected by travel-time, then it is technically possible to gerrymander by promoting the construction of strategic roads or bridges.

5.2 Response to Prior Literature Commentary

Our moment of inertia measure is one of the possible measures of compactness that Young [16] considers. He finds two problems with it: that it gives good ratings to “misshapen districts so long as they meander within a confined area” and that there is a significant bias based on the size of the district (the moment of inertia is proportional to the square of district size).

In response to the first of these complaints, we will show in 5.3 that as an emergent property of our optimization, we will get that all districts are not only contiguous, but also convex (except where they meet non-convex state lines). In other words, we will 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 non-convex.

The second of Young’s concerns about the moment of inertia measure being biased towards optimizing large districts is perhaps more serious. If the complaint is true, then the moment of inertia compactness criteria has to potential to lead to stretched or awkward

urban districts being formed in order to smooth out larger neighboring districts. In our experimental runs, this problem was not severe.

5.3 Mathematical Interpretation

In this subsection, 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 (excepting places where the state border is nonconvex).

We will use the average squared distance between two people in the same district as a measure of the misshapeness of that district. Although we could apply this measure using any distance function, things become significantly nicer when we assume a Euclidean metric. In the discussion below let $E[x]$ and $\text{Var}[x]$ represent the expectation and variance of a random variable x , respectively. If we 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

$$\begin{aligned} E[(x_1 - x_2)^2 + (y_1 - y_2)^2] &= E[x_1^2] - 2E[x_1x_2] + E[x_2^2] + E[y_1^2] - 2E[y_1y_2] + E[y_2^2] \\ &= 2E[x^2] - 2E[x]^2 + 2E[y^2] - 2E[y]^2 \\ &= 2\text{Var}[x] + 2\text{Var}[y]. \end{aligned}$$

Using the definition and standard properties about variance of a random variable this is equivalent to

$$2E[|(x, y) - (\bar{x}, \bar{y})|^2]$$

where (\bar{x}, \bar{y}) is the center of mass of people in the district. Furthermore, this quantity is increased if (\bar{x}, \bar{y}) is replaced by another point. Note that this quantity is twice the moment of inertia of the district.

We assume that there are N people in the state that need to be divided into k districts.

Our objective is equivalent to partitioning our people into k sets S_1, \dots, S_k of equal size, and picking k points in the plane p_1, \dots, p_k in order to minimize

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

where $d(a, b)$ denotes Euclidean distance between points. 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 our 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 that

$$m_{x,i} \geq 0 \tag{1}$$

since we need to have non-negative proportions. Since each person must be wholly divided we have that for any x ,

$$\sum_i m_{x,i} = 1. \tag{2}$$

Lastly, the restriction of district sizes says that for any i that

$$\sum_x m_{x,i} = \frac{N}{k} \tag{3}$$

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

$$\sum_{x,i} m_{x,i} d(x, p_i)^2. \tag{4}$$

By linear programming duality, at the point which minimizes our objective (a global minimum exists since $0 \leq m_{x,i} \leq 1$ implying that our domain is compact), our 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 Equation 3, D_x the coefficient of Equation 2, and $E_{x,i}$ the coefficient of Equation 1. We have that C_i and D_x are

arbitrary, but that $E_{x,i} \geq 0$ 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

$$d(x, p_i)^2 = C_i + D_x + E_{x,i}. \quad (5)$$

Now we note that if $m_{x,i} \geq 0$, that $E_{x,i} = 0$ and hence that $E_{x,i} \leq E_{x,j}$. In particular, person x can only be 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 i^{th} 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 i^{th} district is defined by the equations

$$d(x, p_i)^2 - C_i \leq d(x, p_j)^2 - C_j. \quad (6)$$

Rotating and translating the problem so that $p_i = (0, 0)$ and $p_j = (a, 0)$, and letting $x = (x, y)$, Equation 6 reduces to

$$x^2 + y^2 - C_i \leq (x - a)^2 + y^2 - C_j, \quad (7)$$

or

$$2ax \leq a^2 + C_i - C_j. \quad (8)$$

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

5.4 Computational Complexity

It would be nice to be able to compute the configuration with the global optimum of our Moment of Inertia objective function, but we probably cannot do so in general. Adopting the linear program from Section 5.3 above, we wish to minimize

$$\sum_i \text{Var} [X_i]$$

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 $(\frac{k}{N} \sum_x m_{x,i} \vec{x})$ of the districts i . This is an instance of quadratic programming where we try to maximize a positive semi-definite objective function. Since general quadratic programming is NP-hard, it seems quite 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 of the properties that we want, namely convexity. Furthermore these local maxima are significantly easier to find. At very least we can find them from the quadratic programming formulation using the simplex method.

Unfortunately, the quadratic programming approach leads to an optimization involving kN variables. This can be quite large. Instead we consider the formulation where to have a local maximum we need to pick p_i and C_i as in Section 4 (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 i^{th} 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

$$C - \sum_{x,i} m_{x,i} d(x, p_i)^2 \quad (9)$$

for some constant C . Since we have a global maximum of 9, moving a small amount in any direction within our constraint does not decrease our objective to first order. Furthermore, since our objective is positive semi-definite, this implies that we are at a local maximum. This formulation is much better since we are now left with only $3k$ degrees of freedom, where k is the number of districts.

5.5 Comparison to Hess, Weaver, et al.

We note that this procedure is very similar to that used by Hess, et. al. in 1965 in [8]. They too were attempting to minimize the summed moments of inertia of their districts. They also converged on their solution via an iterative technique that alternated 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 towards achieving exact population equality. Both of these differences seem to stem from the fact that we have both finer data (Hess used 650 enumeration districts for dividing Delaware into 35 state House and 17 state Senate seats, whereas we have on the order of 10 times more census tracts, specific numbers given in Table 1 in Section 7) and more computational power than Hess did.

We are unable to determine exactly what algorithm Hess used to determine optimal districts with given centers other than that he claimed to use a “transportation algorithm”. It is possible that he used the linear programming formulation from Section 4 (possibly using a min-cost-matching formulation). We had many more census tracts to work with and used an algorithm better adjusted to this problem (see Section 5.3).

We also had different perspectives about what to do to even out population. Our fundamental units were sufficiently small that we could just run our algorithm adjusting district sizes in a natural way until all districts were within 2% of the desired population. (In order to speed up computation, in the early iterations of our algorithm, we allowed errors of as much as 20% then gradually tighten the tolerance.) Hess used a solution method that divided his fundamental units of population between districts and later had perform post-iteration checks and alterations so that units were no longer split and population equality still worked out. As Hess points out this readjustment has the potential to increase moments of inertia and could theoretically lead to a failure to converge.

6 Diminishing Halves Method

In order to find an alternative solution against which to compare our moment of inertia algorithm, we use the Method of Diminishing Halves proposed by Forrest in [5].

6.1 Definition

The Diminishing Halves Method splits the state up into two nearly-equal sized districts and recurse on each of the two halves. The idea is to split up in such a way that the resulting halves are relatively compact. 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. We need to choose a methodology for finding a line that would split the state into two compact halves.

Suppose we run a least squares regression on the latitude and longitude coordinates of the state's census tracts. Then 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 will get the

opposite properties. Therefore, we will divide the state at each stage with a line whose slope is perpendicular to the best-fit line of the census tracts in the state. We are not aware of anyone in the literature who has used this specific criterion before.

6.2 Mathematical Interpretation

The best fit line is an approximation of the shape of the state. We compute best fit by attempting to minimize the mean squared Euclidean distance from the line. It is not difficult to see that for any given slope the best possible line with that slope contains the center of mass. Therefore the best fit line is of the form

$$(X - \bar{X}) \sin \theta + (Y - \bar{Y}) \cos \theta = 0. \quad (10)$$

Notice that the left hand side of 10 is the distance of a point (X, Y) from the line. Hence we wish to minimize

$$E \left[((X - \bar{X}) \sin \theta + (Y - \bar{Y}) \cos \theta)^2 \right] = \sin^2 \theta \text{Var} [X] + 2 \sin \theta \cos \theta \text{Cov} (X, Y) + \cos^2 \theta \text{Var} [Y]. \quad (11)$$

This value is minimized when

$$\sin \theta \cos \theta (\text{Var} [X] - \text{Var} [Y]) + (\cos^2 \theta - \sin^2 \theta) \text{Cov} (X, Y) = 0 \quad (12)$$

or when

$$\tan(2\theta) = \frac{-2\text{Cov} (X, Y)}{\text{Var} [X] - \text{Var} [Y]}. \quad (13)$$

Now that we have computed the proper slope of line, in order 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 $\lfloor \frac{k}{2} \rfloor : \lceil \frac{k}{2} \rceil$. When we need to divide into an odd and an even half, ceiling half goes to the southernmost side.

7 Experimental Setup

7.1 Extraction of US Census Data

We used a Perl script (see Appendix B) to extract the census data from US Census 2000 Summary File 1 [12] at the census tract level. For the state of New York, we detected 6661 tracts in the database, 6398 of which had non-zero population. We extracted the population along with the latitude and longitude of a point from each district. The districts had populations varying from 0 to 24523 with a median of 2518.

We used this data to model the population density of New York by assuming that the entire population of each tract was located at the coordinates given. We adjusted 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.

After modeling New York, we also used our Perl script to extract data for the states of Arizona (small – 8 congressional representatives), Illinois (medium – 19 representatives), and Texas (large – 32 representatives). Table 1 lists the states we tested, their populations, number of congressional districts, and number of non-empty census tracts.

State	Population	Number of 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

Table 1: Summary of Census Data

7.2 Implementation in C++

Using our data, we used 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. It should be noted that each step of this procedure decreases the quantity in Equation 4. This is because our two steps consist of finding the optimal districts for given p_i and finding the optimal p_i for given districts. We found the correct values of C_i by alternatively increasing the smallest district and decreasing the largest one. When this adjustment overshoot the necessary value we halved the step size for that district, and when it overshoot by too much, we reversed the change. We found that for New York we were able to converge to our final districts in a couple of minutes on a laptop with a 1.42GHz G4 processor.

After determining our districts, we were able to output them into a postscript file where we displayed the census tracts color-coded by district so that one could visually determine compactness. Finally, for these districts we computed some of the compactness measures discussed in [16] in order to get objective measures of their compactness.

We also created a C++ program to implement the Diminishing Halves method (see Section 6).

7.3 Further Analysis in Mathematica

The Inverse Roeck Test, Length-Width Test, and Schwartzberg Test were used on the regions generated by the C++ program to verify compactness of the proposed districts. These three tests were implemented in Mathematica with aid of the Convex Hull and Polygon Area notebooks from the Wolfram Mathworld website [13] [14].

8 Measures of Compactness

We need an objective method for determining how successful our program is at creating compact districts. In [16], Young gave several measures for the compactness of a region. We will use some of these to help compare our districts with those produced by other methods. Since our algorithms generate convex districts except in places 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.

8.1 Definitions

The Inverse Roeck Test. Let C be the smallest circle containing the region, R . We measure $\frac{\text{Area}(C)}{\text{Area}(R)}$. This is a number bigger 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. We compute the ratio of length to width of this rectangle. This will be a number more than 1, with numbers closer to 1 corresponding to more compact regions.

The Schwartzberg Test. We compute the perimeter of the region divided by the square root of 4π times its area (we use different wording from Young, but this test is mathematically equivalent). 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.

8.2 Calculation in Mathematica

The Inverse Roeck Test, Length-Width Test, and Schwartzberg Test were used on the regions generated by the C++ program to verify compactness of the proposed districts. These three

tests were implemented in Mathematica with aid of the Convex Hull and Polygon Area notebooks from the Wolfram Mathworld website [13] [14].

First we assumed that the regions were defined by the convex hull of the census tracts that they contain. This is reasonable assuming the districts are convex, which they are except where they meet non-convex state boundaries. We determined the bounding polygon by taking the convex hull of the census tracts contained within each district.

For the Roeck test, we computed the area of the polygon by triangulating it. We found the circumradius by noting that if every triple of vertices can be inscribed in a disk of radius R , then the entire polygon can be fit into such a 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 i^{th} point. If every triple of points can be covered by the same disk, then any three of the D_i 's intersect. Therefore, by Helly's theorem 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.

The Length-Width test is computed as follows. 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 will determine the bounding sides of our rectangle. We chose the value from the orientation that yields the largest length-to-width ratio.

Calculating the Schwartzberg test is straightforward. We compute the area of the polygon and its perimeter, and the resulting answer is $\frac{\text{Perimeter}}{\sqrt{4\pi\text{Area}}}$.

9 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. Full page versions of each map are printed in Appendix A.

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 have been 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 is still in vector form (before rasterization) shows that our districts are indeed convex.

9.1 Discussion of districts

One can see visually that both the Moment of Inertia and the Diminishing Halves Method produce more compact looking results than the districts currently in place. Some of the current New York districts legitimately try to respect county lines, but there are a few egregious offenders such as Congressional 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 in order 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 model does not contain information about county lines, so we cannot evaluate its ability to keep communities intact based on counties. However, by looking at regions where the census tracts are clustered, we can see the location of cities on the map. Both of our methods do a good job at keeping the major cities of New York intact (excepting the fact

that it is difficult to evaluate the New York City area, which contains about half the state’s population). The cities of Buffalo and Rochester are divided into at most two districts in our methods, whereas they are divided into 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 the Moment of Inertia and the Diminishing Halves 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 produced rounder districts. The Diminishing Halves Method tends to regions that are almost all triangles and quadrilaterals. Whenever three districts meet with the Diminishing Halves Method, 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.

The fact that the greater New York City area contains roughly one half of New York’s population is convenient for the Diminishing Halves algorithm. However, the Diminishing Halves algorithm does not deal very well with bodies of water. This led to the creation of one noncontiguous district (given by pink squares in Figure 1c). Overall, the shapes given by the Moment of Inertia Method look rounder and more appealing.

9.2 Compactness Measures

Table 2 lists the results of the compactness tests. All of the tests we use produce numbers larger than 1 where smaller numbers correspond to more compact regions.

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

Table 2: Mean and standard deviation districts for compactness measures of districts. Smaller numbers correspond to more compact districts.

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 a 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 of these measures tells us approximately the same thing: the Moment of Inertia Method performs a little bit better than the Diminishing Halves Test.

It would be desirable to compare the numbers in Table 2 to the current US districts, but there are two reasons why we cannot do this. First, the US Census 2000 census file that we read from does not offer congressional district identification at the census tract level. In order to compute compactness, we would need to choose a more fine population unit, so the numbers would not be directly comparable to those in Table 2. Second, 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.

10 Results for Other States

In order to test how well our algorithms performed on states with different sizes, we also computed districts for the states of Arizona (small – 8 districts), Illinois (medium – 19 districts), and Texas (large – 32 districts). Figures 2–4 give the maps of the Moment of Inertia Method districts, those generated by the Diminishing Halves Method, and the current congressional districts. Full page versions of each map are printed in Appendix A.

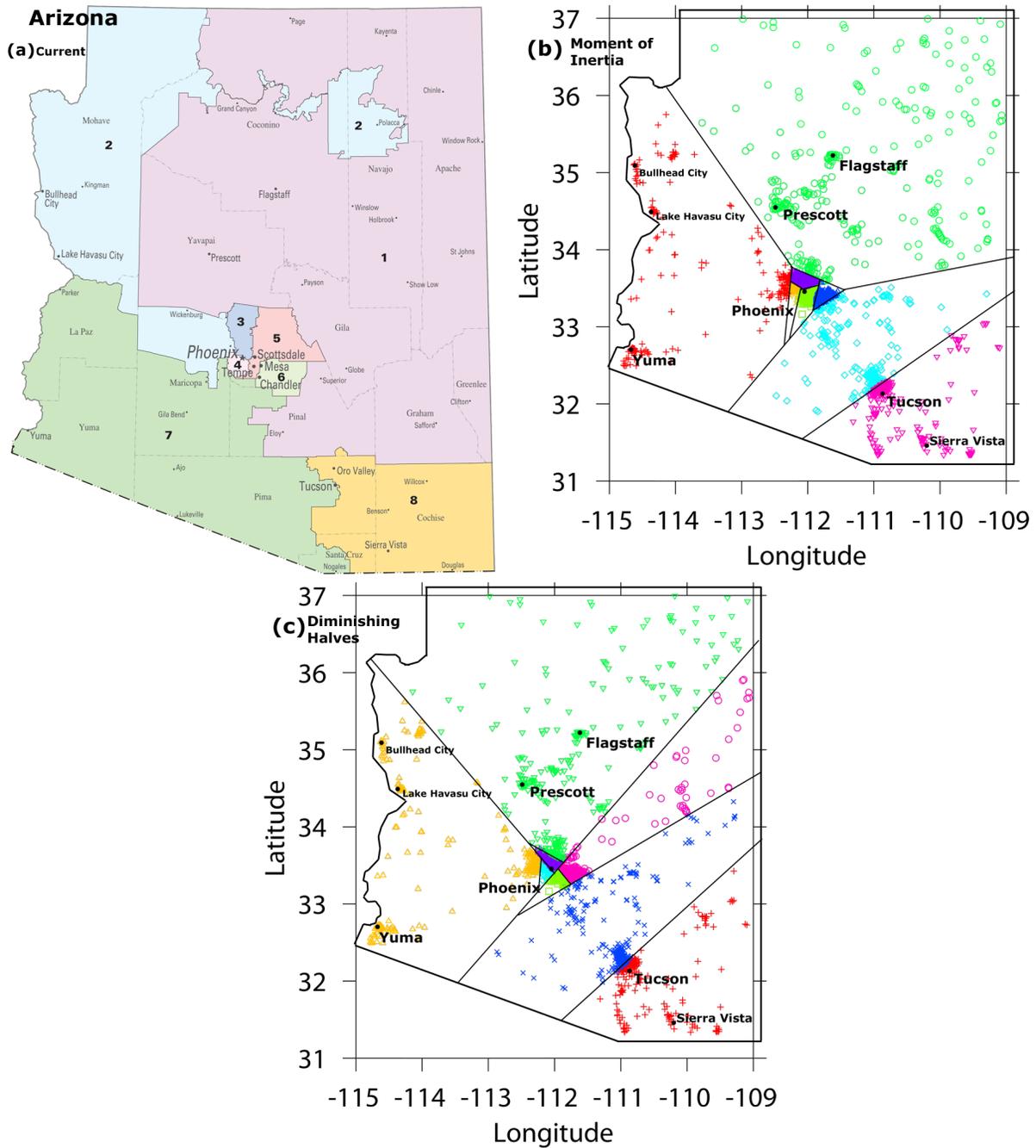


Figure 2: Arizona district assignments. (a) Current (adapted from [11]). (b) Moment of Inertia Method. (c) Diminishing Halves Method. Larger copies printed in Appendix A.

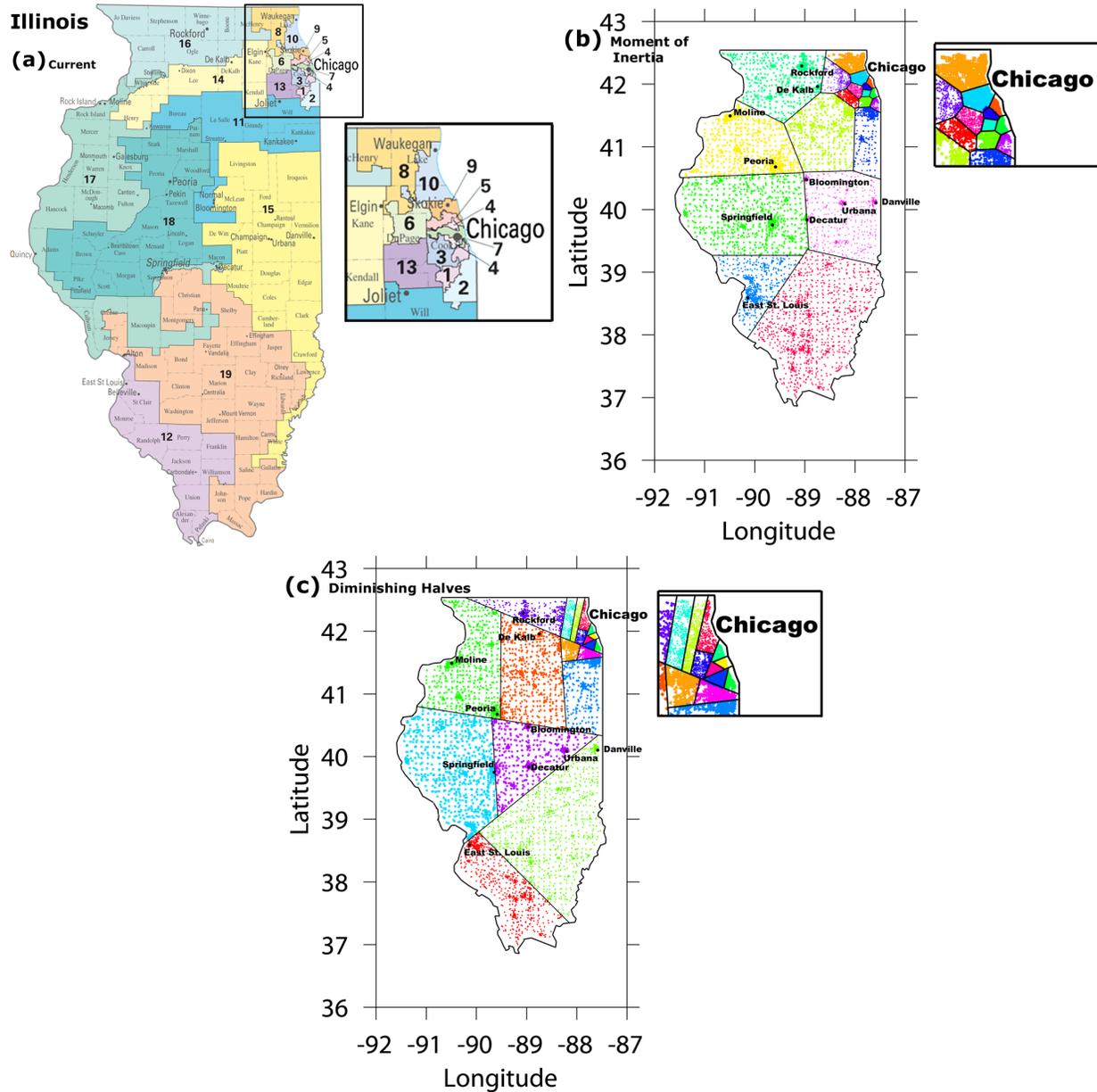


Figure 3: Illinois district assignments. (a) Current (adapted from [11]). (b) Moment of Inertia Method. (c) Diminishing Halves Method. Larger copies printed in Appendix A.

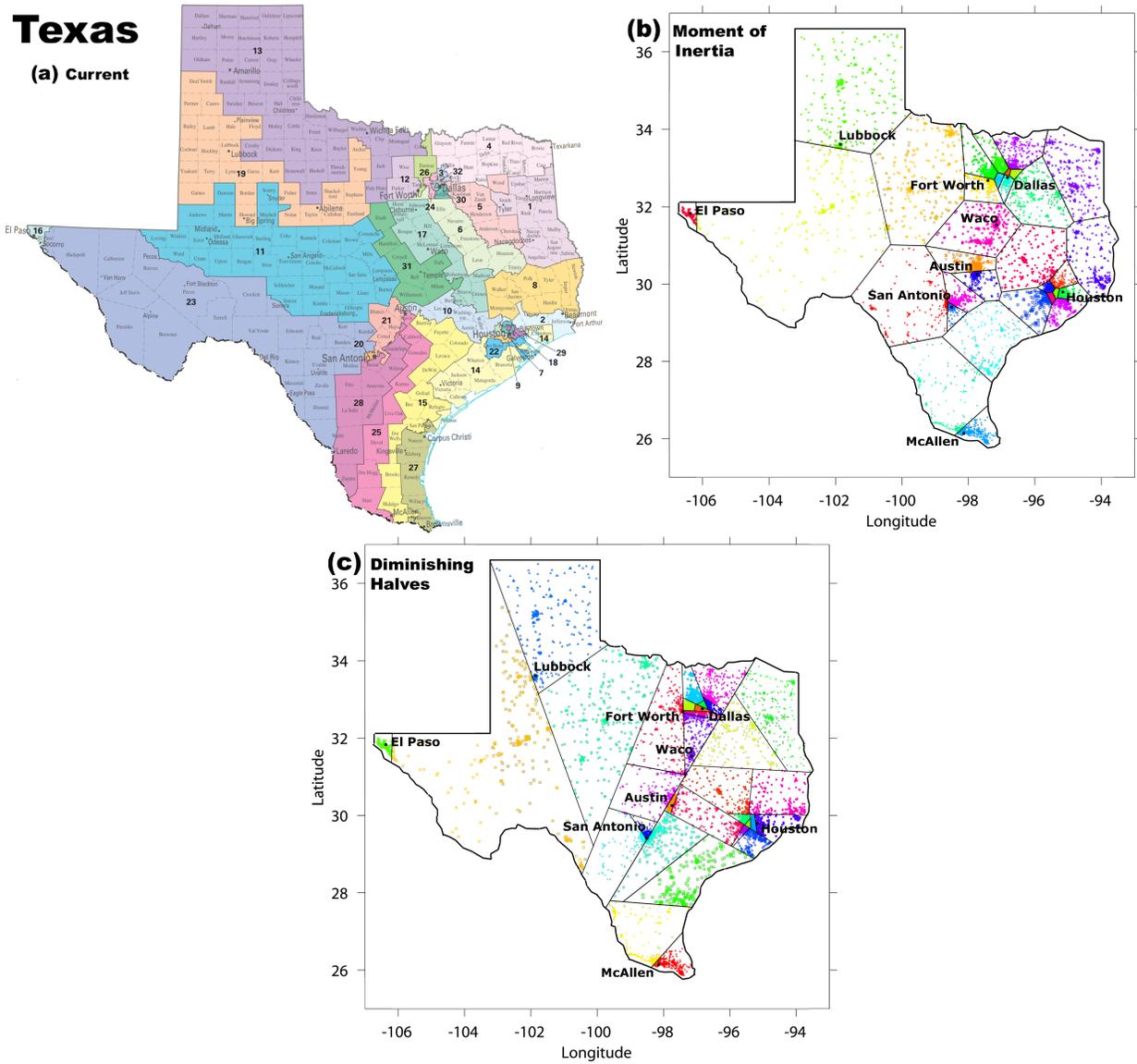


Figure 4: Texas district assignments. (a) Current (adapted from [11]). (b) Moment of Inertia Method. (c) Diminishing Halves Method. Larger copies printed in Appendix A.

10.1 Discussion of Districts

1. *Arizona.* Under the current division, Arizona District 2 is a very blatant case of gerrymandering. Neither of our models produces a district this bad.

Unlike New York, which has half of its population in the corner of the state surrounding New York City and the other half spread out somewhat evenly, Arizona contains two cities with a large population (Tucson and the Phoenix area) and is sparsely populated elsewhere. Such a high concentration of population causes the Diminishing Halves Method to make unappealing triangular cuts that dangle into the far corners of the state. Aside from District 2, even the current districts appear to look better on average than the Diminishing Halves Method's districts. The Moment of Inertia Method does a much nicer job.

The case of Arizona suggests that the Moment of Inertia Method will have an easier time adjusting to smaller states where there are fewer lines to draw and population density can be concentrated in only one or two small regions.

2. *Illinois.* The current Illinois District 17 is strongly noncompact. Districts 11 and 15 also have suspicious tails that do not follow any county line. Both of our models provide a fairer redistricting. The Moment of Inertia Method does unfortunately split the cities of Bloomington and Decatur, but not any worse than they are split by the current US district assignment. The Diminishing Halves Method does a better job of keeping Bloomington and Decatur intact, but splits Springfield and the fringes of Peoria, so there is a tradeoff. In the region surrounding the Chicago area, the Moment of Inertia Method's districts look much more organic than the stripes painted by the Diminishing Halves Method, so we recommend the Moment of Inertia Method overall.
3. *Texas.* Texas is a large state comparable in population to New York. However, unlike New York, Texas has its population distributed across several very large cities instead

of primarily one. Houston, San Antonio, Dallas, and Austin are all large cities. The Diminishing Halves Method has a difficult time dealing with all the large cities, and draws too many thin awkward triangles. The Moment of Inertia Method cleanly divides all the major cities into as few components as is reasonable and avoids filling the southern part of Texas with thin districts the way that Texas Districts 15, 25, and 28 appear in the current plan.

Even though the size of Texas is comparable to New York, the Moment of Inertia Method takes substantially longer to compute an answer. The computation runs in a little under half an hour, presumably due to the fact that Texas has more distinct population centers than New York. The computation time is still very reasonable compared to the timescale of calculations in fields such as computational fluid mechanics.

10.2 Compactness Measures

Table 3 lists the results of the compactness tests. All of the tests we use produce numbers larger than 1 where smaller numbers correspond to more compact regions.

Districts	Inverse Roeck Test	Schwartzberg Test	Length-Width Test
TX (Moment of Inertia)	2.04 ± 0.64	1.14 ± 0.09	1.72 ± 0.57
TX (Diminishing Halves)	2.76 ± 1.66	1.27 ± 0.20	2.30 ± 1.73
IL (Moment of Inertia)	1.90 ± 0.36	1.28 ± 0.26	1.55 ± 0.39
IL (Diminishing Halves)	2.49 ± 0.99	1.35 ± 0.24	2.01 ± 0.96
AZ (Moment of Inertia)	2.18 ± 0.56	1.17 ± 0.08	1.77 ± 0.51
AZ (Diminishing Halves)	2.69 ± 0.91	1.29 ± 0.15	2.07 ± 0.79

Table 3: Mean and standard deviation for compactness measures of districts. Smaller numbers correspond to more compact districts.

The Diminishing Halves algorithm produces consistently worse results by all three measures, most of the individual data sets are still within the borderline of statistically

significant plausibility. Additionally, for these states the Diminishing Halves Method produces results with extraordinarily high standard deviations, especially for the smaller states. This fact suggests (and the maps seem to confirm) that the discrepancy is due largely to a small number of very elongated districts produced by the Diminishing Halves Method.

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

11 Sensitivity to Parameters

As a test for robustness, we tweaked some of the parameters to the Moment of Inertia Model to see if the results would be substantially impacted. Neither of these parameters is relevant to the Diminishing Halves Method, which is deterministic aside from having to choose which half of the state receives the $\lfloor \frac{k}{2} \rfloor$ share and which half receives the $\lceil \frac{k}{2} \rceil$ share of the population when dividing a region whose assigned number of representatives k is odd (the Diminishing Halves Method currently assigns the larger division to the southernmost half). We additionally tested variants of the Diminishing Halves Method to see if we could improve the performance to be comparable to the Moment of Inertia Method.

11.1 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. A very close inspection could reveal a small number of borderline census tracts that were not always identically grouped, but the centers of mass of the districts were in essentially the same location regardless of the initial seed.

11.2 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.

11.3 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 instead of a best-fit line that was not 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 that alternately draws vertical and horizontal (longitude and latitude) lines on the New York case. In all these modified cases, results were visibly much worse than those generated with the method as described in Section 6. The modified methods tended to split cities into more districts than our normal method.

12 Strengths and Weaknesses

Strengths:

- **Emergent behavior from simple criteria.** We only specify criteria 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 non-compactness of a district is equivalent to its moment of inertia. This gives us a model that is easy to understand and does not use any arbitrary constants that could be fine-tuned to gerrymander districts.

- **Results in convex districts.** In both models we produce districts that are 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 could be computed in a few minutes with modest computational resources.
- **Nice looking final districts.** The districts resulting from our algorithm appear very nice.

Weaknesses:

- **No theoretical bounds on convergence time.** We were unable to prove that our algorithm converges in reasonable time, although it has done so in practice.
- **Potential for elongated smaller districts.** It is possible that some of the smaller districts produced by the Moment of Inertia Method are stretched in order 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 crossing rivers, but could place pressure on making districts noncompact and allows for loopholes that could be exploited by malicious politicians.
- **Does not necessarily find the global optimum.** Our Moment of Inertia algorithm only finds a local minimum of our objective function. As far as we know finding the global minimum is not computationally tractable. This leads potentially to some non-determinism in the resulting districts, which could in turn allow gerrymandering, but the amount is small.

- **Can only draw new districts, not determine if existing districts are gerrymandered.** The paper by Cirincione et al. [4] contained a pseudoconfidence interval analysis where they used their model to predict whether South Carolina's 1990 redistricting had been gerrymandered. We do not perform this analysis here. The problems of analyzing existing districts for signs of gerrymandering and drawing new districts are separate, but share many of the same concepts.

13 Conclusion

We have formulated and tested two methods for assigning congressional districts with a computer. We have written a Perl script that allows us to easily obtain US Census data from all of the census tracts in any state.

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. The fact that variants of this method were among the first algorithms to be seen in the literature testifies to the intuitiveness of this algorithm. With modern computer technology, we are able to implement this method and obtain results that would not have been computationally feasible in the 1960s and 1970s.

The Diminishing Halves Method uses the concept of recursively dividing the state in half, which is very simple to explain to voters. In an attempt to avoid elongated districts and to cut along sparsely populated areas rather than densely populated regions, our implementation of the diminishing halves method chooses a dividing line perpendicular to the statistical best-fit line of the latitude and longitude coordinates of the census tracts. We are not aware of any location in the literature that uses this specific variant of the Diminishing Halves Method.

In reading the literature concerning computer districting schemes, one finds that a good number of these papers contain warnings against putting too much trust in the supposed unbiasedness of computer automation. Although no computer system is perfect, we do have

some concrete recommendations that we would advise state officials to implement when considering how to assign congressional districts in a fair manner:

- **Processing data at the census tract level or finer is computationally feasible.**

Our methods process data at the census tract level and run in a couple minutes, (slightly longer for Texas). It would not be unreasonable to attempt data processing at the block group level if it was felt that 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 require districts

to be contiguous, yet the districts they generate are not only contiguous but convex. Neither of our methods attempt to preserve city or county boundaries, yet the Moment of Inertia Method does a good job at 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 the state of New York where there are comparatively few counties, looking at city integrity instead of county integrity is a more reasonable idea.

- **A good algorithm can handle states of different sizes.** Algorithms that perform

well on large states might not necessarily yield good results when applied to a smaller state with only one or two large cities. We have tested our algorithms on states of different sizes and found that the Moment of Inertia Method behaves well in all cases.

- **We recommend a moment of inertia compactness criterion.** The Moment

of Inertia Method consistently produced more visually appealing districts than the

Diminishing Halves Method. All of the additional compactness tests that we performed came down in favor of the Moment of Inertia Method. The Moment of Inertia Method does a better job at respecting city boundaries. It is possible that there is a variant of the Diminishing Halves method that performs better, but in all of our tests the Moment of Inertia Method was the strongest performer.

References

- [1] Altman, M. “The Computational Complexity of Automated Redistricting: Is Automation the Answer?” *Rutgers Computer and Technology Law Journal* **23** (1997) 81–136.
- [2] Bacao, F; Lobo, V; Painho, M. “Applying genetic algorithms to zone design.” *Soft Computing* **9** (2005) 341–348.
- [3] Chou, CI; Li, SP. “Taming the Gerrymander—Statistical physics approach to Political Districting Problem.” *Physica A* **369** (2006) 799–808.
- [4] Cirincione, C; Darling, TA; O’Rourke, TG. “Assessing South Carolina’s 1990s congressional districting.” *Political Geography* **19** (2000) 189–211.
- [5] Forrest, E. “Apportionment by Computer.” *The American Behavioral Scientist*. Thousand Oaks: Dec 1964. Vol. 8, Iss. 4; p. 23, 35.
- [6] Galvao, LC; Novaes, AGN; de Cursi, JES; Souza, JC. “A multiplicatively-weighted Voronoi diagram approach to logistics districting.” *Computers & Operations Research* **33** (2006) 93–114.
- [7] Garfinkel, RS; Nemhauser, GL. “Optimal Political Districting by Implicit Enumeration Techniques.” *Management Science* **16** (1970) B495–B508.

-
- [8] Hess, SW; Weaver, JB; Siegfeldt, HJ, Whelan, JN, Zitlau, PA. “Nonpartisan Political Redistricting by Computer.” *Operations Research* **13** (1965) 998–1006.
- [9] Mehrotra, A; Johnson, EL; Nemhauser, GL. “An Optimization Based Heuristic for Political Districting.” *Management Science* **44** (1998) 1100–1114.
- [10] Nagel, SS. “Simplified Bipartisan Computer Redistricting.” *Stanford Law Review*. **17** (1965) 863–899.
- [11] National Atlas of the United States. “Printable Maps – Congressional Districts.” <http://nationalatlas.gov/printable/congress.html>
- [12] US Census Bureau. “Census 2000 Summary File 1 Delivered via FTP.” http://www2.census.gov/census_2000/datasets/Summary_File_1/New_York/
- [13] Weisstein, EW. “Convex Hull.” From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/ConvexHull.html>
- [14] —. “Polygon Area.” From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/PolygonArea.html>
- [15] Williams, JC. “Political Redistricting — A Review.” *Papers in Regional Science* **74** (1995) 13–39.
- [16] Young, HP. “Measuring the Compactness of Legislative Districts.” *Legislative Studies Quarterly* **13** (1988) 105–115.