Class Project:

Mathematical Measures to Estimate Partisan Gerrymandering

CSCI 8715 Spring 2022
Group 4

Fafa Hoshyargar, Yiming Yao

University of Minnesota
Driven to Discover℠

Problem Statement

- **Input:**
    - General election dataset
    - Redistricting maps
    - Geospatial data (polygon shapes)
- **Output:**
    - Partisan bias in different maps by various mathematical methods
- **Objectives:**
    - Assess proposed plans in terms of partisan gerrymandering
    - Assess mathematical measures in estimating partisan gerrymandering
- **Constraints:**
    - Data only available from the last US census
    - Vote counts only available from past elections

Motivation

- Developing scientific tools to quantify gerrymandering bias, using outlier detection approaches
- To help government, politicians, and people who devote themselves to politicking determine the most appropriate redistricting plans
- Using modern scientific methods to close/reduce manipulating district maps as loopholes in elections
- To ensure that people who have the right to vote participate in an unbiased election

Contribution

Contribution Claim:

1. Study the basic concepts of the gerrymandering problem in Wikipedia and define the spatial bias
2. Formally define several measurements to estimate partisan gerrymandering
3. Using different mathematical approaches to perform similar calculations for various maps in Minnesota to quantify partisan gerrymandering
4. Validate the effectiveness and accuracy of our measurements in quantifying spatial bias

Key Concepts

Gerrymandering

- The process of manipulating district lines for partisan purposes
- In 1812, Republican Governor Elbridge Gerry signed a controversial redistricting plan that favored the Republicans
- Federalists responded with a cartoon comparing one of the districts to a salamander, dubbing it a "gerrymander"
- The term stayed in the American political lexicon

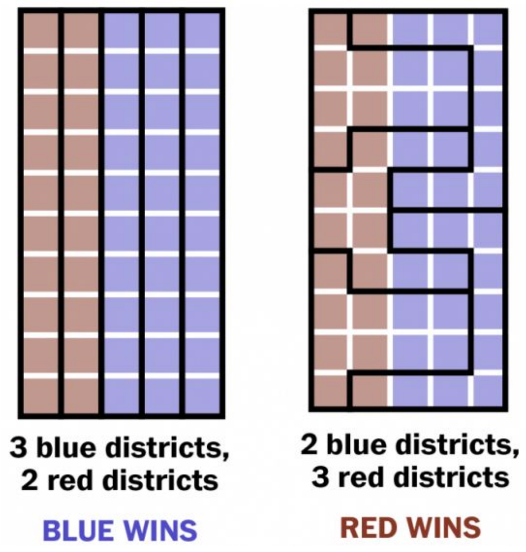


Fig. 1 An example of gerrymandering [1]

Key Concepts

Efficiency gap

- Partisan gerrymandering is carried out in one of two ways:
  - <u>Cracking</u>: preferred candidate *loses* by narrow margins across *many* districts
  - <u>Packing</u>: preferred candidate *wins* by overwhelming margins in a *few* districts
- Cracking and packing produce "wasted votes"
  - Votes cast for the losing candidate are wasted
  - Votes cast for the winning candidate above 50% are wasted
- The efficiency gap is one party’s total wasted votes, minus the other party’s total wasted votes, divided by the total number of votes

Fig. 2 An example of packing and cracking [2]

Limitations of Existing Work and Proposed Approach

- Limitations
  - None compared across many measures (at least for MN)
  - Hard to provide an absolute definition of spatial bias

- Proposed Approach
  - Analyze well-known metrics for the existing map and six proposed maps
  - Compare various measures to find the least gerrymandered plan(s) in MN
  - Figure out the most robust estimator(s) of spatial bias

Definition of Measures

- **Perimeter to area ratio**
  - Use the perimeter to area ratio, $\frac{P_i}{A_j}$, where P_i is the district's total perimeter, and A_i is the district's area, where *n* denotes the number of congressional districts within a map:

$$\frac{P}{A} = \sum_{i=1}^{n} \frac{P_i}{A_i}$$

- **Ranked-marginal deviation**
  - For a given redistricting plan $\vec{d}_i$, we take Party A's votes for each district when the districts are ordered from the most to the least, $\vec{m}$, and consider the differences between the mean and the observed Party A's percentage. This provides a set of numbers to compute the two-norm, a distance metric:

$$\left\| \vec{m} - \vec{d}_i \right\|_2 = \sqrt{\sum_{j=1}^{n} (m_j - d_{ij})^2}$$

Definition of Measures

- **Distance to mean**
  - Use average number of seats won in all districts $m_{Seats} = \frac{\sum_{i=1}^{r} S_i}{r}$, where *r* denotes the number of redistricting plans, to calculate the absolute distance between the share of winning votes in a given redistricting plan, S_i, from this average value:

$$d_i = \left|S_i - m_{Seats}\right|$$

- **Seats to votes ratio**
  - $\frac{S_i}{V_i}$, is calculated for each district and summed, where S_i denoted seats per district and is ***1*** if won in that district (gained more than fifty percent of casted voted) and ***0*** otherwise; V_i s the vote percentage in that particular district. Therefore:

$$\frac{S}{V} = \sum_{i=1}^{n} \frac{S_i}{V_i}$$

Definition of Measures

- **Area Ratio Convexity Measure**
  - The ratio of the area of the largest convex set, A_i, n a polygon to that of the convex hull, C_i, of the polygon, *i.e.*, $\frac{A_i}{C_i}$, is calculated and summed:

$$ARCM = \sum_{i=1}^{n} \frac{A_i}{C_i}$$

- **Efficiency gap**
  - The efficiency gap is the absolute difference in the two parties' wasted votes, divided by the total number of votes:

$$Efficiency\ gap = \frac{\left| Wasted\ votes_{Party\ A} - Wasted\ votes_{Party\ B} \right|}{Total\ votes}$$

Experimental Methodology

- Shapefiles of the <u>existing map</u> joined with election data (precincts assigned to districts)
- Shapefiles for <u>six alternative maps</u> from the Minnesota Legislature website
- Calculated geometries based on shapefiles using "GeoPandas"

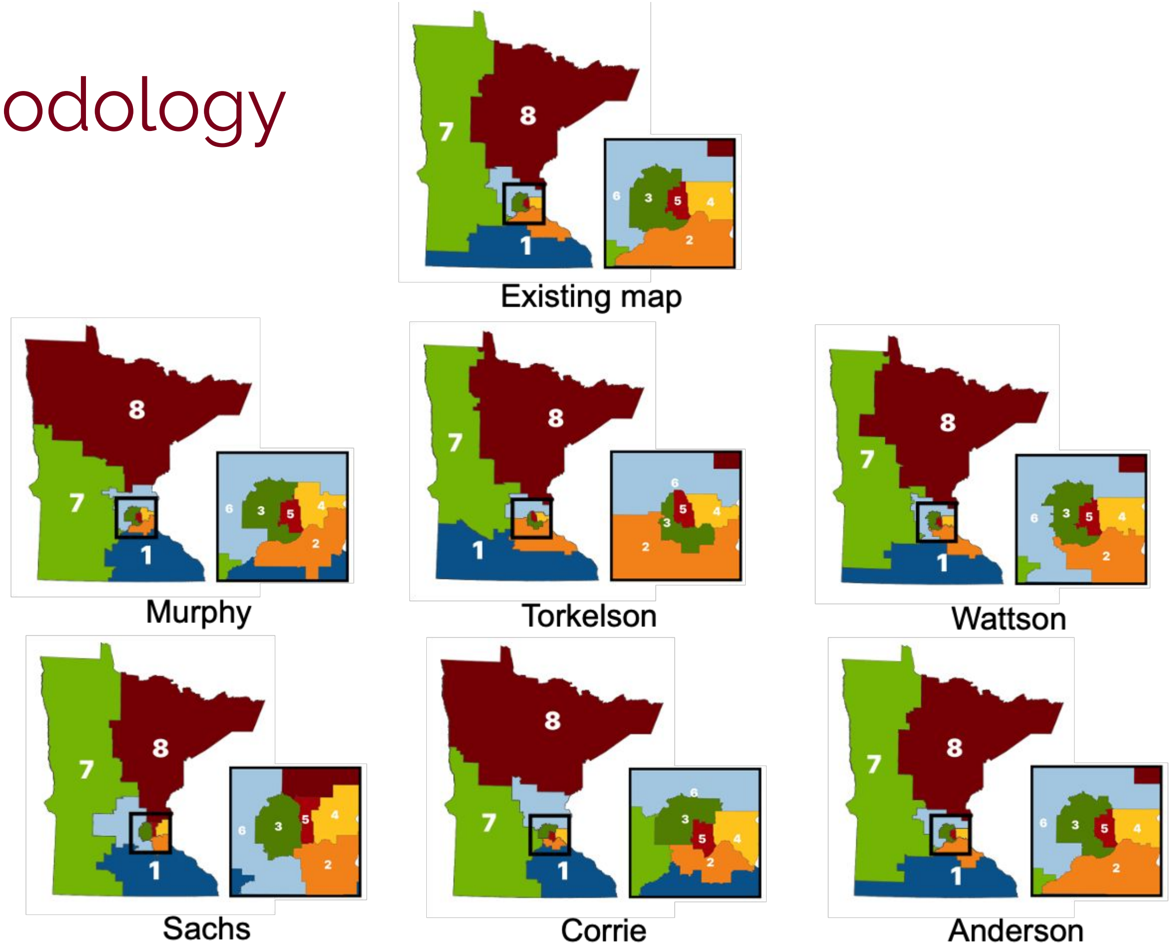


Fig. 3 Minnesota redistricting plans [3]

Summarized Results

Calculation of 2020 Minnesota presidential election gerrymandering bias in the proposed approaches

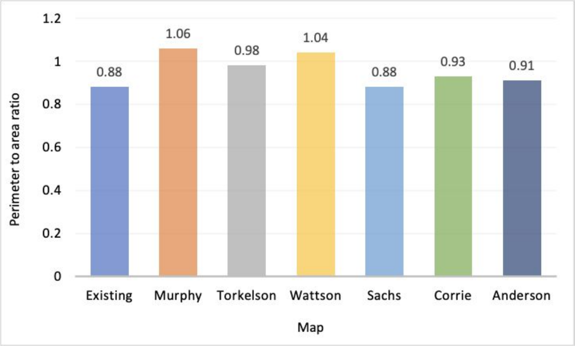
	Perimeter to area ratio	Ranked-marginal deviation	Distance to mean	Seats to votes ratio	Area Ratio Convexity Measure	Efficiency gap
Map1 Existing	**0.88**	**2.08**	**0.14**	**6.21**	**6.15**	**6.6%**
Map 2 Murphy	1.06	2.81	0.14	6.11	6.10	7.4%
Map 3 Torkelson	0.98	**8.95**	**0.86**	**4.19**	6.15	**20.2%**
Map 4 Wattson	1.04	1.82	0.14	6.15	5.87	7.4%
Map 5 Sachs	0.88	6.81	0.14	6.04	6.42	7.2%
Map 6 Corrie	0.93	5.32	0.14	6.09	6.58	7.2%
Map 7 Anderson	0.91	2.17	0.14	6.16	6.11	7.3%

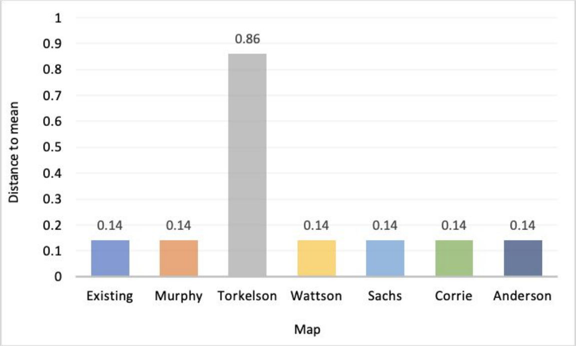
Table 1: MN results

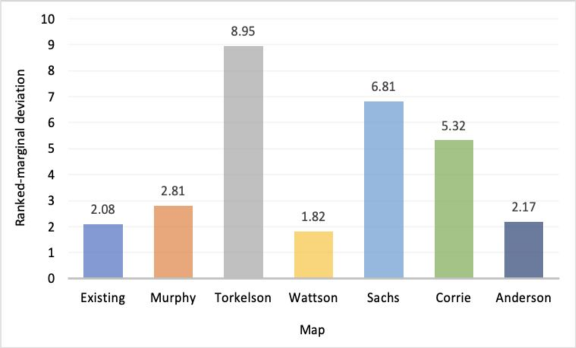
Green: existing plan for the 2020 Minnesota presidential election

Blue: extreme values, i.e., anomalies

Summarized Results



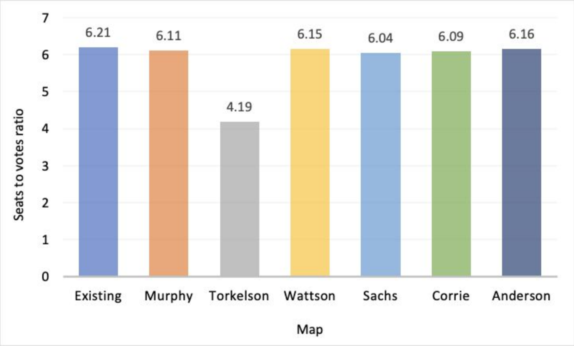


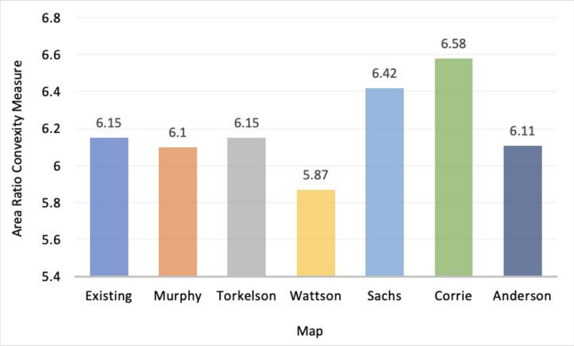


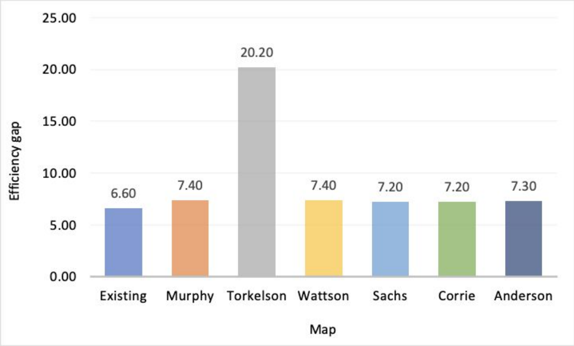
- X-axis:
  - Minnesota redistricting plan
- Y-axis:
  - Different for each measure

- Anomaly detected in the Torkelson map (**grey bar**):
  - This map cuts out suburban area in favor of more rural territory thus leaning towards Republican (Republican party won 5 out of 8 districts in the election)

Summarized Results (Cont'd)

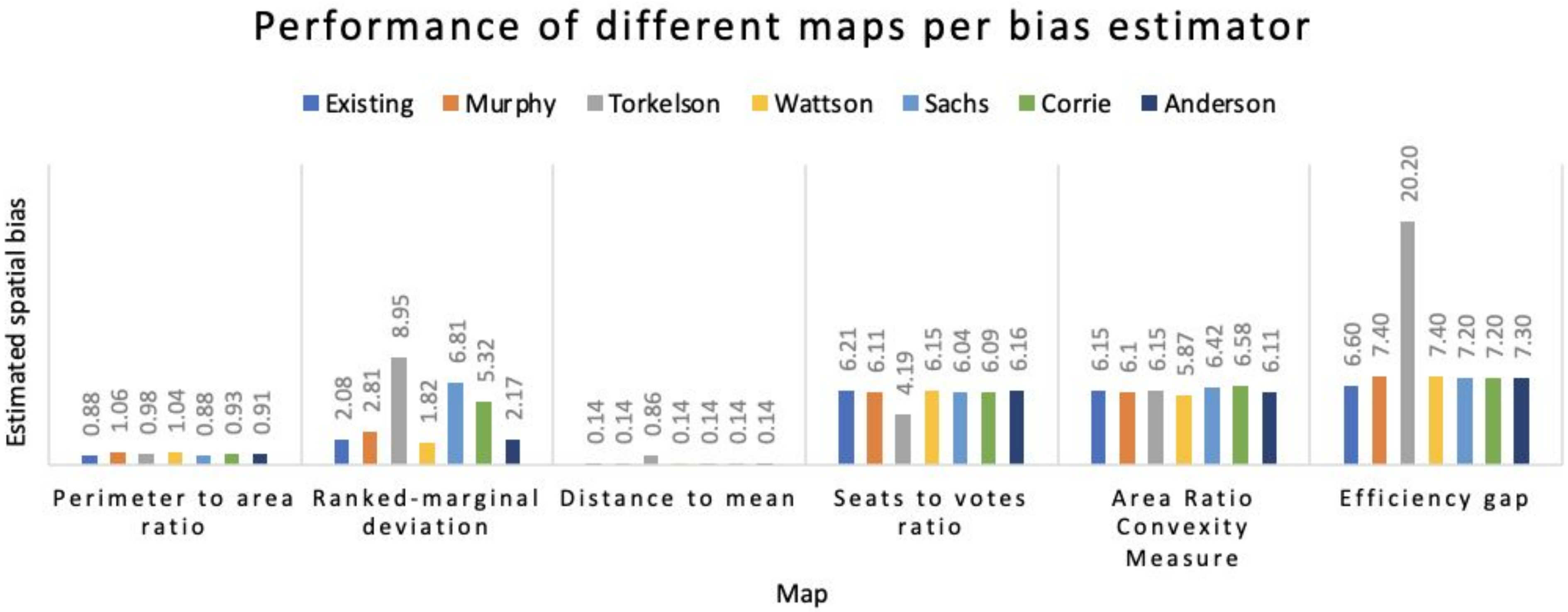






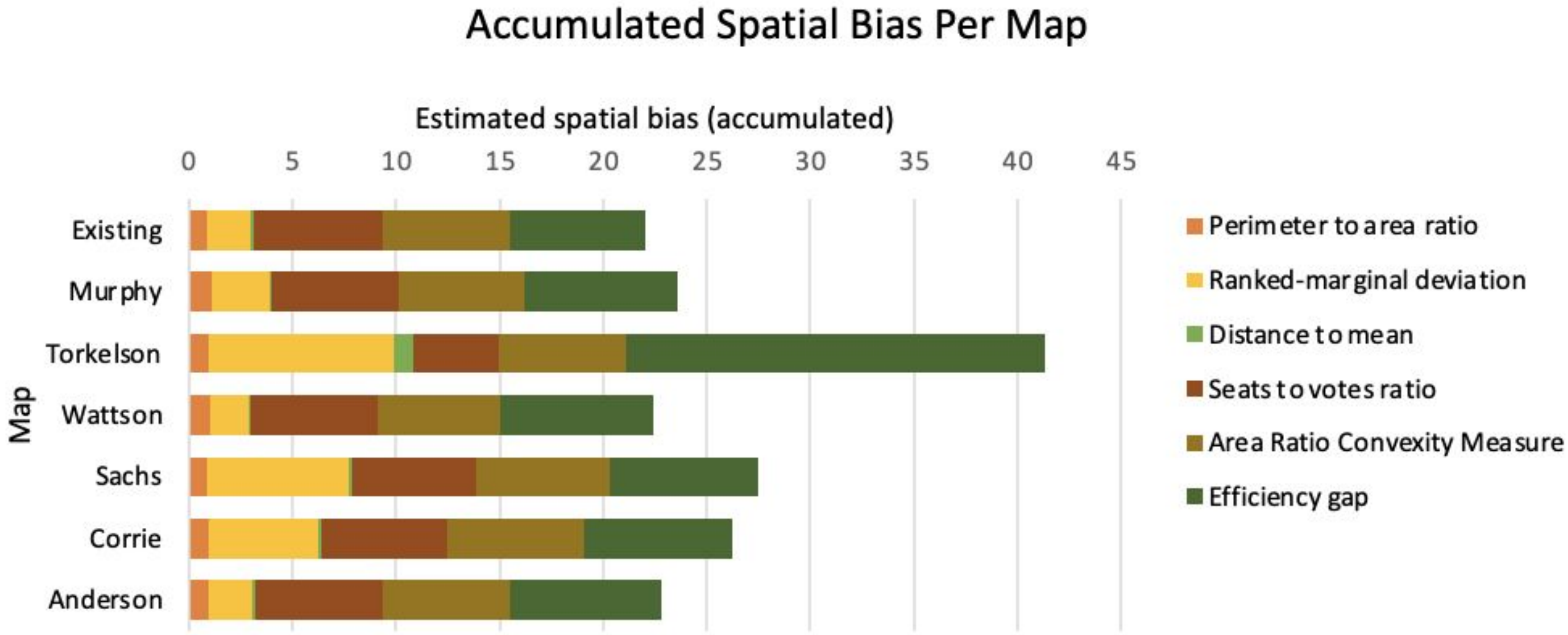
- Torkelson map would create five Republican-leaning districts by **packing** more Democratic votes into the three most urban districts

Results



Performance of different maps per bias estimator
Existing
Murphy
Torkelson
Wattson
Sachs
Corrie
Anderson
Estimated spatial bias
0.88
1.06
0.98
1.04
0.88
0.93
0.91
2.08
2.81
8.95
1.82
6.81
5.32
2.17
0.14
0.14
0.86
0.14
0.14
0.14
0.14
6.21
6.11
4.19
6.15
6.04
6.09
6.16
6.15
6.1
6.15
5.87
6.42
6.58
6.11
6.60
7.40
20.20
7.40
7.20
7.20
7.30
Perimeter to area ratio
Ranked-marginal deviation
Distance to mean
Seats to votes ratio
Area Ratio Convexity Measure
Efficiency gap
Map

Results



- Torkelson map would create five Republican-leaning districts by **packing** more Democratic votes into the three most urban districts (anomaly)

Discussion

- **How do mathematical measures behave in estimating partisan gerrymandering?**
    - *Efficiency gap*: the smaller the value of the efficiency gap, the smaller the gerrymandering in a given map (generally less than 7-8%)
    - *Seats to votes ratio*: smaller ratios (fewer votes won more seats) means more considerable partisan bias
    - *Distance to mean*: smaller distance to average seats won means less partisan bias
    - *Ranked-marginal deviation*: minor deviation means less partisan bias
    - *Perimeter to area ratio & Area Ratio Convexity Measure*:
        - Least robust in estimating partisan bias

Discussion

Torkelson map (Map 3) has the largest partisan bias (anomaly)

The Existing map (Map 1) demonstrates the best performance of these four measurements:

1) Efficiency gap = 6.6% (1st lowest)
2) Seats to votes ratio = 6.21 (1st highest)
3) Ranked-marginal deviation = 2.08 (2nd lowest)
4) Distance to mean = 0.14 (lowest in tie)

	Distance to mean	Efficiency gap	Seats to votes ratio	Ranked-marginal deviation
Map1 Existing	0.14	6.6%	6.21	2.08
Map 2 Murphy	0.14	7.4%	6.11	2.81
Map 3 Torkelson	**0.86**	**20.2%**	**4.19**	**8.95**
Map 4 Wattson	0.14	7.4%	6.15	1.82
Map 5 Sachs	0.14	7.2%	6.04	6.81
Map 6 Corrie	0.14	7.2%	6.09	5.32
Map 7 Anderson	0.14	7.3%	6.16	2.17

Table 1: partial results of our experiment

Discussion

- **What are suitable mathematical measures to estimate partisan gerrymandering?**
  - Efficiency gap
  - Ranked-marginal deviation
  - Seats to votes ratio
  - Distance to mean
  - P/A and ARCM need further mathematical modeling to measure urbanization level, then used to analyze partisan gerrymandering

- **The rank for performance in proposed maps:**

1. Existing map
2. Wattson
3. Anderson
4. Murphy
5. Corrie
6. Sachs
7. Torkelson

Conclusion and Future Work

- Conclusion
  - The perimeter to area ratio & Area Ratio Convexity Measure have the least significant power in estimating partisan bias
  - Torkelson map (Map 3) shows the most significant bias
  - The existing map (Map 1) shows the best performance
  - The efficiency gap, seats to votes ratio, distance to mean, and ranked-marginal deviation are suitable mathematical measures to estimate partisan gerrymandering

- Future Work
  - Incorporate more political aspects into the computation
  - Explore further methods to improve spatial bias detection accuracy
  - Add weights to each measure defining bias to create an integrated approach to estimating partisan gerrymandering

References

[1] Ingram, C. (2015). This is the best explanation of gerrymandering you will ever see. https://www.washingtonpost.com/news/wonk/wp/2015/03/01/this-is-thebest-explanation-of-gerrymandering-you-will-ever-see/?utm˙term=.5cf029656cab/.

[2] Klarreich, E. (2017). How to Quantify (and Fight) Gerrymandering. https://www.quantamagazine.org/the-mathematics-behind-gerrymandering-20170404/.

[3] Montgomery, D. H. (2022). What you need to know about redistricting in Minnesota. https://www.mprnews.org/story/2022/02/11/what-you-need-to-know-about-minnesota-redistricting/.

[4] MGGG Staff (2020). Minnesota Election Shapefile Dataset. GitHub. https://github.com/mggg-states/MN-shapefiles.

[5] Geographic Information Services: 2020 Redistricting Plans. Minnesota Legislature. https://www.gis.lcc.mn.gov/redist2020/plans.html.