Little House (Seat) on the Prairie: Compactness, Gerrymandering, and Population Distribution

Allan Borodin
University of Toronto
Toronto, Canada
bor@cs.toronto.edu

Omer Lev
Ben-Gurion University of the Negev
Beersheba, Israel
omerlev@bgu.ac.il

Nisarg Shah
University of Toronto
Toronto, Canada
nisarg@cs.toronto.edu

Tyrone Strangway
Ben-Gurion University of the Negev
Beersheba, Israel
tyrone@cs.toronto.edu

ABSTRACT

Gerrymandering is the process of creating electoral districts for partisan advantage, allowing a party to win more seats than what is reasonable for their vote. While research on gerrymandering has recently grown, many issues are still not fully understood such as what influences the degree to which a party can gerrymander and what techniques can be used to counter it. One commonly suggested (and, in some US states, mandated) requirement is that districts be "geographically compact". However, there are many competing compactness definitions and the impact of compactness on the gerrymandering abilities of the parties is not well understood. Also not well understood is how the growing urban-rural divide between supporters of different parties impacts redistricting.

We develop a modular, scalable, and efficient algorithm that can design districts for various criteria. We confirm its effectiveness on several US states by pitting it against maps "hand-drawn" by political experts. Using real data from US political elections we use our algorithm to study the interaction between population distribution, partisanship, and geographic compactness. We find that compactness can lead to more fair plans (compared to implemented plans) and limit gerrymandering potential, but there is a consistent asymmetry where the party with rural supporters has an advantage. We also show there are plans which are fair from a partisan perspective, but they are far from optimally compact.

KEYWORDS

Voting; Gerrymandering; Redistricting; Compactness; Social Choice; Fairness

ACM Reference Format:
Allan Borodin, Omer Lev, Nisarg Shah, and Tyrone Strangway. 2022. Little House (Seat) on the Prairie: Compactness, Gerrymandering, and Population Distribution. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022,* IFAAMAS, 9 pages.

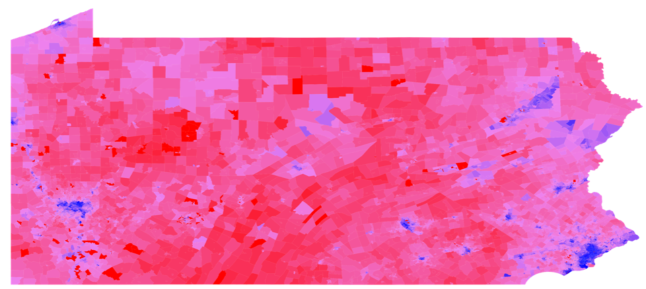

Figure 1: Voter distribution for the 2016 presidential election in Pennsylvania. The more blue a region the more it leaned left (Democrats), more red for the right (Republicans).

1 INTRODUCTION

In many democracies, politicians are elected to represent the people of particular geographic areas, called districts. There is country wide aggregation of votes. Instead voters within a district pick a winner from the alternatives vying to represent their district. Political power is based on the number of districts won.

How voters are partitioned into these districts directly affects the makeup of the legislative body. The partitioning is often governed by hard constraints. For example, most jurisdictions require that the districts be geographically connected (with certain exceptions) and have roughly equal populations. In addition, there are many competing goals when designing a districting plan [42]. One could prioritize not breaking up communities of interest, such as those with a shared culture and history.[1] It may also be desirable to be as compatible as possible with established city and county boundaries, a consideration studied by Wheeler and Klein [43]. Another reasonable goal would be to obtain geographically compact regions (a goal enshrined in some US states' laws and regulations[2]). A less defensible goal is *gerrymandering*: designing districts for partisan gain, i.e., creating districts which help a particular party gain a number of seats beyond its popular support.

In the US, following every 10-year census, state legislatures decide their new federal congressional and state legislative districts, and partisan concerns are often part of the consideration [41]. For example, in the 2020 federal election in North Carolina, a state accused of gerrymandering (partially overturned by courts [8]), the Democratic party received 49.96% of the vote and won five districts;

[1]Regarding ethnic minorities, this is required by the US' Voting Rights Act of 1965.
[2]For example, California's constitution states, in article XXI, "districts shall be drawn to encourage geographical compactness".

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

the Republican party received 49.41% of the vote and won eight districts [30]. As noted above, due to the many competing goals, even with clearly stated goals it is not clear what is fair or optimal when it comes to non-partisan redistricting (see Wasserman [42] for further discussion and a comparison of objectives).

Parallel to the political partisan redistricting process there is a more complex, ongoing population-wide process. In the US [3] and Europe [31], voters are "reorganizing" themselves for various economic and social reasons. As Figure 1 shows, in Pennsylvania voters of left leaning parties tend to cluster in dense urban centres, while voters of right leaning parties spread out into surrounding rural regions. People in the large cities of Philadelphia in the east and Pittsburgh in the west are overwhelmingly Democrat voters, while the surrounding rural regions lean Republican. In short, in many democracies around the world the left-right split can be characterized as an urban-rural divide as well.

Our contribution. Our work explores aspects of both of these processes – the immediate partisan one and the process of population dynamics. In the first part of our work (Section 5), we introduce our automated redistricting procedure, which is flexible and can be used to design plans for various objectives, both partisan and nonpartisan. To prove the utility of our algorithm, we compare its performance against hand-drawn plans from election experts. We also explore (Section 5.4) the social contribution of our algorithm.

Once we show the power of our algorithm, we begin using it to understand the interplay between population distribution, geographic compactness constraints, and political power. Our algorithm allows us to examine possible requirements that have been suggested as a means to mitigate or eliminate gerrymandering. In particular, we study the impact of a compactness requirement. In Section 6, a few compactness measures are considered and we see that in the US, the more rural party (Republicans) still consistently outperforms the more urban party (Democrats). Moreover, we see that this advantage is robust *even in a non-gerrymandered, optimally-compact plan.* This advantage is not due to political gaming of the division process, but rather due to the geographic spread of each party's supporters. That being said, we show the existence of "fair" but not ideally compact plans.

In Section 7 we examine how compactness constraints affect gerrymandering possibilities. We show that demanding stringent compactness constraints reduces the ability of parties to reach extreme gerrymanders. However, in most cases, the compactness requirement allows for relatively greater rural-party gerrymandering. Indeed, under the most stringent compactness constraints, the urban party sometimes cannot even achieve its vote proportion.

Finally, note space limitations force us to omit some details throughout the paper. All missing details can be found in the full version.[3]

2 RELATED WORK

Political districting, and in particular gerrymandering, has long been the examined in the humanities and social sciences. There have been legal discussions [17, 23, 36], and recently the US Supreme Court ruled [35] partisan gerrymandering cannot be addressed by federal courts. Gerrymandering has long been studied by historians [5, 13] and political scientists [14, 15, 20, 38].

There is a long history of trying to detect gerrymandering by proposing and comparing various scores in potential plans [16, 19, 33, 41]. Recently, the AI community has become interested in redistricting. Pegden et al. [32] proposed treating the redistricting process as a repeated "cut-and-choose" game, while Bachrach et al. [1] measure how much districts can distort voter representation.

Most related to our work is automated redistricting, which originally was proposed as a solution to gerrymandering by Vickrey [40]. A popular automated redistricting technique uses Markov Chains to generate a sequence of related districts (see DeFord et al. [10], Fifield et al. [16] for two types of chains). These chains allow us to see where values from a plan, such as various measures of partisanship or compactness, fall within the distribution of a random ensemble of them. Outlier values may be a sign of poor district design, or gerrymandering. One such analysis was used to help with the 2018 Pennsylvania redistricting [11]. Another line of automated redistricting work focuses not on generating ensembles of plans, but instead designing one optimized plan. Within the AI community many published automated methods violate basic legal requirements such as population balances [6, 26]; or are unable to scale to real-life sized problems [4]. Some other methods that produce legal plans for real data are only able to design plans for objectives based on linear combination of individual district scores [21]. These shortcomings are not surprising since finding an optimal gerrymandering, or certain graph partitions, is known to be NP-hard in various settings [12, 26, 39]. For a recent survey on various automated redistricting methodologies we refer the reader to Becker and Solomon [2].

As discussed, the growing urban-rural divide and its impact on redistricting is an inspiration for our work. Commentators [24, 29] have argued this divide is amplifying the effects of partisan gerrymandering. Using simulated data, Borodin et al. [4] argue the rural party can stretch its vote share more effectively when gerrymandering than the urban party – although at extreme levels of divide the opportunity to gerrymander is limited. Rodden [34] goes further, arguing that the urban-rural divide is a fundamental disadvantage to the urban party in almost every scenario, though his argument boils down to the advantage of rural areas having more urban party supporters than vice versa.

3 MODEL

We examine gerrymandering with a graph-theoretic formulation. We shall use a US-oriented terminology (states, precincts, etc.), but the formulation represents most geographic districting settings. A state is an undirected graph $G(V, E)$, and each node $v \in V$ represents a precinct, a small geographic region where votes are tallied. An edge $(u, v) \in E$ represents that precincts u and v share a physical boundary. For $v \in V$ let n_v be the number of people who live in precinct v, and $n_v^{p,e}$ be the number of people who live in v and vote for party p in election e. We will omit e when the context is obvious. Let $N = \sum_{v \in V} n_v$ be the total number of people in the state. We limit our focus to two parties: the rural party (in the US, Republicans (R)), and the urban party (in the US, Democrats (D)).

[3] https://www.cs.toronto.edu/~nisarg/papers/gerrymandering_compactness.pdf

Creating a districting plan requires partitioning G into K vertex-disjoint subgraphs $G_1, \cdots, G_K$ (the districts). The number of districts K is extrinsically determined (in the US, by a census every 10 years). There are two widely accepted requirements for legal districts in the US and elsewhere:

Contiguity For each $k \in [K]$, G_k must form a connected subgraph of G. In the real world, this translates to being able to walk from any point in the district to any other point in the district without crossing into another district.

Population balance-δ Given $\delta > 0$, for each $k \in [K]$,

$$1 - \delta \leq \frac{\sum_{v \in V(G_k)} n_v}{N/K} \leq 1 + \delta.$$

The exact value of δ required in the U.S. changes between states (and judicial decisions). Informally, districts should be as near equal-sized in population as possible [25]. We take $\delta = 0.005$, so that the maximum population deviation between any two districts is at most 1% of the state's population. This is the legal requirements in some states, and a far tighter constraint than what is respected by many previously proposed automated redistricting methods.

Parties are interested in winning as many districts as they can. The party with the most voters in a district is typically said to win that district. For example, if $\sum_{v \in V(G_k)} n_v^{D,2012} > \sum_{v \in V(G_k)} n_v^{R,2012}$, we say the Democrats win district k according to the 2012 presidential vote totals in a given state. If the inequality is reversed, we say the Republicans win the district in that election. Note that this is only one way to define "winning a district". In Section 5.1 we calculate the probability of winning a hypothetical district, based on historic vote totals and outcomes, and find plans with several high probability wins.

4 ELECTION SETTINGS

In this paper, we use election data from three US states — Pennsylvania, North Carolina, and Wisconsin — from the 2012 and 2016 presidential elections.[4] These are states and elections for which granular, precinct-level, data is available.[5] Each of these three states has a particular election of interest. We also include the number of nodes (precincts) and edges in the graphs of each state.

Pennsylvania (PA) 2012 *Sizeable Democrat advantage.* The Democrat (Obama) won 51.97% of the vote, vs. 46.59% to the Republican (Romney). PA has 9,255 nodes and 25,721 edges.

North Carolina (NC) 2016 *Sizeable Republican advantage.* The Republican (Trump) won 49.83% of the vote, vs. 46.17% to the Democrat (Clinton). NC has 2,692 nodes and 7,593 edges.

Wisconsin (WI) 2016 *Near tie.* The Republican (Trump) won 47.22% of the vote, vs. 46.45% to the Democrat (Clinton). WI has 6,634 nodes and 18,126 edges.

In addition to these elections, they provide a good mix of geographic features. WI, for example, has its north-east corner carved up by lake Michigan, forming a jagged bay. PA and NC, on the other hand, have a much more convex shape. Furthermore, the population distribution is varied: PA's large urban centres are in its east and west edges, whereas in NC, the urban centres are concentrated in the middle of the state.

4.1 Voting and the Urban-Rural Divide

State	2012 correlation	2016 correlation
NC	0.79	0.80
PA	0.47	0.58
WI	0.38	0.57

Table 1: Spearman correlation between a precinct's fraction of D party votes and its density (total population divided by area) in three states and two elections.

As noted above, a key geographic feature of US political parties is the growing divide between a more rural Republican party and a more urban Democratic party. We examined different states, with differing ethnic makeup, education patterns, and history, but this feature was common across all our data: densely populated urban centres favour the Democrats, while sparsely populated rural regions favour the Republicans (see Table 1).

5 THE GREAT ALGORITHM

To study the role of compactness and population distribution in gerrymandering we need an algorithm that can optimize for various compactness and partisan fairness metrics (or handle them as constraints) on real-world data. To that end, we introduce our Goal-based Redistricting for Elections Automatically using Technology (GREAT) algorithm, that can create plans from graph representations. As we will demonstrate, our algorithm, with minimal engineering effort, can be used to optimize various measures of partisan fairness (e.g., to minimize the robust partisan bias metric introduced in Section 6), partisan gain (e.g., the number of districts won by a given party either by achieving a plurality of votes, or at least a threshold fraction of votes, or with at least a certain probability), and compactness (e.g., the Polsby-Popper and Convex Hull scores defined in Section 5.2). Furthermore, the algorithm can optimize towards one of these goals while satisfying strict constraints on other metrics (e.g., optimize compactness while ensuring that a given party wins at least a fixed number of districts).

To show its capabilities, we will now demonstrate our algorithm is capable of matching the performance of human experts when creating partisan plans (Section 5.1), and compact plans (Section 5.2). In Section 5.3 we show our algorithm is able to compete with human experts in a prestigious redistricting competition.

First, we give a brief overview of our algorithm (see full version for details). Our method is based on simulated annealing, a local-search like method which can make non-improvement steps, allowing it to escape local optima. After some fixed number of iterations or elapsed time, the process ends and the best of all iterated solutions is returned. Starting from an (often random) initial plan, a step is considered by using a modification of the tree-recombination procedure proposed by the Metric Geometry and Gerrymandering Group [28]. Briefly, the method takes a set of m adjacent districts from the current solution, and recombines and redivides the nodes in them to form m new districts. This is done by drawing random

[4]In Sections 5.1 and 5.2 we also use data from Maryland and Massachusetts for a proof of concept. But, because of missing votes and an overwhelming Democrat lean respectively, we omit them from subsequent experiments.
[5]Data from MGGG (https://github.com/mggg-states).

State	Total seats	Our D	538 D	Our R	538 R
MD	8	7	5 (8)	4	4 (4)
MA	9	9	9 (9)	0	0 (0)
NC	13	7	8 (8)	11	10 (10)
PA	18	8	8 (9)	13	13 (13)
WI	8	5	5 (5)	6	6 (6)

Table 2: First column is the number of seats in the state. Second and third columns are the number of districts D take with over 82% probability with our algorithm and the 538 optimally-gerrymandered plans, respectively. Fourth and fifth columns are the same for the R party. The 538 numbers show the number of districts won according to their districting based on our election data. In parentheses are 538's results using absentee data (which we did not have access to).

spanning trees over the precincts of the m districts and cutting random edges in the trees to separate the nodes into the desired number of districts. For efficiency reasons, we generally use $m = 2$. Using larger m values did not noticeably improve the results. Any objective that can be expressed numerically and calculated from an arbitrary plan may be used. Additionally, any binary constraint (for which it can be checked whether a given plan satisfies) can be incorporated by ensuring that the algorithm only considers steps which satisfy the given constraint.

Like the work before us, we are unable to provide guarantees (with respect to optimal solutions) on our method's performance. Instead, we compare against the best plans human experts create. As far as we are aware, we are the first to publish work that compares against, let alone matches, state of the art hand-drawn plans.

5.1 Proof of Concept: 538 Gerrymandering

Nate Silver and the election experts at 538's gerrymandering project [37] drew thousands of hand-crafted districts for various objectives. While there is no guarantee their plans are optimal, they do serve as an excellent, and publicly available, benchmark.

As noted above, winning a plurality of votes is just one of the measures of what it means to win a district. At 538, they took a probabilistic view, designing partisan plans that maximized the number of districts that were won with a sufficient probability. This non-trivial measure of victory also serves as an ideal goal to show the modularity of our algorithm. Unfortunately, they released few details regarding their method. However, we believe we were able to reconstruct it using released results (see full version for a detailed reconstruction). Briefly, 538 uses the Cook Partisan Voting Index (CPVI) [9], which measures a district's D party bias according to the 2012 and 2016 elections, and transforms it into the probability that the D party wins that district. The R party wins it with the remaining probability. When gerrymandering for party P, 538's objective was to maximize the number of districts for which P's probability of winning was at least 82%. To guide our method, we used a combination of the expected number of districts won by P and the total number of districts won with at least 82% probability (see full version for exact optimization details).

The availability of presidential election data at the precinct level is inconsistent, so we are unable to compare against 538 in all states.

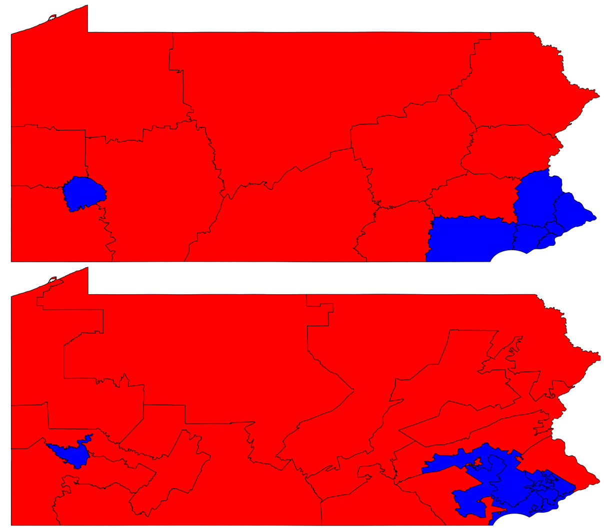

Figure 2: PA districts (R wins in red); (D wins in blue) based on the 2016 PA election data. Top, our plan, optimizing the convex hull score; bottom, PA's actual 2011 districts.

There are five states for which we have publicly available data, and for each of them we optimized for the 538 objective for each party. For each state and party we ran our algorithm on 60 cores with 2.10 GHz computing power for 24 hours, though our algorithm often stopped advancing well before this deadline. Out of all solutions iterated, we took the one with the most districts above 82% for the indicated party. Our results are shown in Table 2.

Overall, there was only one case, NC for D, where we did not match 538. Even here, we only missed by one district out of the 13. We did outperform 538 in Maryland for the Ds, but we caution we were missing 25% of their vote for each party – the absentee data (mail-in ballots), for which we have no precinct level data (the other states had over 99.7% ballots accounted for). In NC for the R party, we also outperformed 538.

5.2 Proof of Concept: Compact Redistricting

As mentioned, compactness is often a legislated requirement, even if the mathematical definition and the required levels are not specified. Despite this ambiguity our algorithm is able to easily optimize for a variety of compactness scores. To measure compactness in a single district, we use the following scores:

Polsby-Popper (PP) A district's PP score is its area divided by the area of the circle with the same length perimeter.

Convex Hull (CH) A district's CH score is its area divided by the area of its minimum convex hull.

For these two metrics, we use our algorithm to find a plan optimized for the mean score across all districts. From a visual standpoint (see Figure 2), our plans pass an "eye test" for looking compact, especially compared to the plans enacted in practice.

As was the case for gerrymandering, 538 implemented a compact plan for each state. These plans were designed to minimize "the average distance between each constituent and his or her district's geographic centre". In addition, we have plans created by the public

	PP	CH	538	DRA
Our Polspy-Popper	**44**	86	0.31	**93**
Our Convex Hull	37	**88**	0.29	85
Our 2% Fair	26	76	0.34	49
538's Compact Plan	34	87	**0.27**	81
DRA's Compact Plan	40	82	0.29	70
2011 Plan	16	62	0.41	15
Updated Plan	32	78	0.30	64

Table 3: The compactness scores for various plans in Pennsylvania. The best score for each metric is bolded. For PP, CH, and DRA, higher is better, for 538 lower is better.

using Dave's Redistricting App (DRA). DRA is the most popular open source tool for redistricting, and 538 used it too to create all their plans. Amongst all plans ever published on DRA, the website features the most compact (according to their internal metric) for each state. In addition, we have the 2011 plan for all relevant states, and for NC and PA a court mandated updated plan as well.

Compared to *all* of the above mentioned plans, in *every* state our plans had the best mean compactness scores for their respective metrics. They are compact even according to metrics they were not optimized for (in PA and NC our PP plans have the best DRA score). We are not claiming the compactness measures we chose are superior to others. As Table 3 shows, in PA, according to any metric the four compact plans have similar scores. Far more compact than the 2011 and our 2% Fair plan (the later of which will be introduced in Section 6.2). Even the Updated Plan, partially designed to address the 2011 plan's compactness issues, is generally less compact than the compact ones. We only argue, for a variety of measures, our algorithm is capable of creating plans just as compact as those from human experts. For the WI, NC, MA, and MD compactness tables information about the 538 and DRA metrics, see Appendxi refsubsec:compactnes-tables.

5.3 Proof of Concept: Princeton Redistricting

Finally, we used our algorithm for a redistricting competition, *The Great American Mapoff*, hosted by Princeton University's Gerrymandering Project. Here, we were finalists from among almost 150 entrants. We competed in two categories. First, designing compact but gerrymandered plans ("stealth gerrymandering"), is the exact topic of Section 7. For the second, partisan fairness, we used our robust partisan bias score (explained in Section 6). Our plans, created in a day, were judged by human experts to be among the best, as good as the handcrafted plans submitted by other participants. Because the contest goals were open ended, we are unable to make exact quantitative comparison. For more details of the contest and our submissions, see the full version.

5.4 The Ethics of Automated Redistricting

Before discussing our main results, we wish to touch upon the ethics of automated redistricting and its implications. There is an understandable concern our tool could be used to advance partisan interests. This point is especially salient for our tool, which, in hours, can match what human experts take much longer to produce.

However, the actual redistricting process takes years, and is only done once every ten years in the United States (and in many other democracies). In these situations, partisan groups would have years – and near unlimited resources – to have experts craft plans by hand, limiting the utility of an automated tool for gerrymanderers. Furthermore, the actual redistricting process involves a certain human element. When crafting a plan there is bargaining and dealing between the various interested actors. To protect their position within their district, a representative of one party may wish to keep communities of similar ethnicity, income, or shared history together. Thus they may bargain with and make concessions to members of the other party. While this behaviour would be interesting to model, it is not something a one shot algorithm is capable of.

We see our tool as something researchers can use to study redistricting. Here, we use it to explore the impact and limitations of compactness requirements. Furthermore, it can be used to help combat gerrymandering: if a plan is as biased as the highly partisan plans produced by our tool, then there is strong evidence of gerrymandering. Because it is highly modular, it can be used to quickly propose alternative plans, optimizing a diverse set of desiderata. The plans our tool suggests may not be the final ones, but can give a sense of what is and is not achievable.

6 FAIRNESS IN DISTRICTING

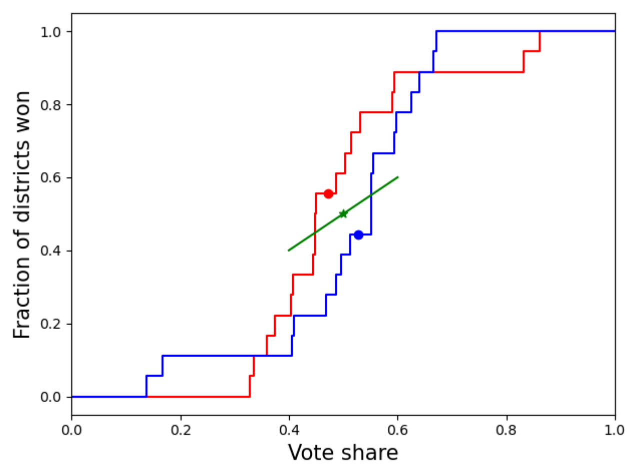

Figure 3: Uniform swing for $R(D)$ in red (blue), in the 2012 presidential election in PA under our Convex Hull compact plan. Vertical axis shows the fraction of districts won; horizontal axis the vote fraction. The dots on the party curves indicate the actual election outcome (0 swing). The green line is the range of proportional outcomes on the range $[0.4, 0.6]$. A green star marks the point $(1/2, 1/2)$.

In this section we examine the bias presented in various plans. In particular we measure the deviation from what is "fair". Formally, we want the fraction of districts won by each party to match – as closely as possible – its fraction of overall votes. However, measuring how proportional a plan is based on data from just one election is not very robust. To address this, the *uniform swing* model [14] is widely used. In this model, hypothetical elections are generated starting from a baseline election by shifting the vote shares of the parties. Specifically, the vote share of a given party is increased or decreased by an equal amount in every district. The fraction of districts won by each party is then measured in these hypothetical elections in order to measure the amount by which proportionality would likely be violated if vote shares change over time. See Figure 3 for the uniform swing model applied to the 2012 PA presidential election under the plan produced by our algorithm for optimizing the Convex Hull score, where the fraction of districts won by a party is plotted against the fraction of votes received by the party at different uniform swing levels.

There are several metrics that use the uniform swing model to measure the partisan bias in a given plan. We are interested in the *partisan bias* score [18]. This value measures the vertical displacement of the swing curve from the point $(1/2, 1/2)$. Intuitively, the partisan bias measures the divergence from the idea that "half the votes should translate to half the seats". More generally, we can measure the vertical displacement from any point (a, a) for $a \in [0, 1]$. We introduce a robust version of this metric. Fixing a line segment $[l, r]$ $(l, r \in [0, 1], l < r)$, we measure the average vertical distance from the swing curve to the line $y = x$ over this line segment. We use $[0.4, 0.6]$ as the reasonable range (as is the case with most presidential elections, the vote shares of both parties are between 40% and 60%). The 45° line in this range is shown in green in Figure 3. There are two ways to measure the distance between a party's swing curve $(s(x), x \in [0, 1])$ and a line segment $[l, r]$. The first is a signed version,

$$\frac{\int_l^r (s(x) - x)dx}{r - l} \tag{1}$$

measuring on average how much higher or lower the swing curve is over the proportional line. A positive (negative) value indicated this party, over the range of reasonable vote shares, can expect more (fewer) seats than what is proportionally fair. Alternatively we could take the unsigned version,

$$\frac{\int_l^r |s(x) - x|dx}{r - l} \tag{2}$$

which measures the average deviation from proportionally fair. That is, how much does the plan deviate from "an α fraction of the vote share should translate into an α fraction of the seats". While these measures are similar, and often correlated, they can differ. For example, consider a plan where a vote share of $50\% + \epsilon$ ($50\% - \epsilon$) results in winning (loosing) each district. For any symmetric range about 50% vote share, this plan has the best possible score for Equation 1 and the worst possible score for Equation 2. Both of these measures provide important information. For a fixed party p and its swing curve we quantify its partisan advantage (disadvantage), over $[l, r]$, by how positive (negative) Equation 1 is. On the other hand, Equation 2 tells us, over $[l, r]$, on average how disproportionate p's outcomes are, its total advantage plus total disadvantage.

For any $t < 0.5$, the fraction of districts won by one party with $0.5 - t$ vote share is exactly one minus the fraction of districts won by the other party with $0.5 + t$ vote share. Hence, for a symmetric range around the 0.5 vote share point, their swing lines are mirrors of each other about the point $(0.5, 0.5)$. Thus for both parties, the value of Equation 1 is identical in magnitude (but opposite in sign), and the value of Equation 2 will be identical.

We note that Equation 2 is non-linear, it can not be calculated by combining scores calculated from individual districts. This is unlike the 538 partisan measure, which is just the sum of how much each district leans towards a target party. Even the compactness measures we examine are the simple means (and sums) of individual district compactness scores. Thus this measure is fairly complex to optimize for. Even the algorithm presented recently by Gurnee and Shmoys [21], which was designed with the goal of finding fair plans, would be unable to optimize our second fairness objective.

6.1 Compactness Can Improve Fairness

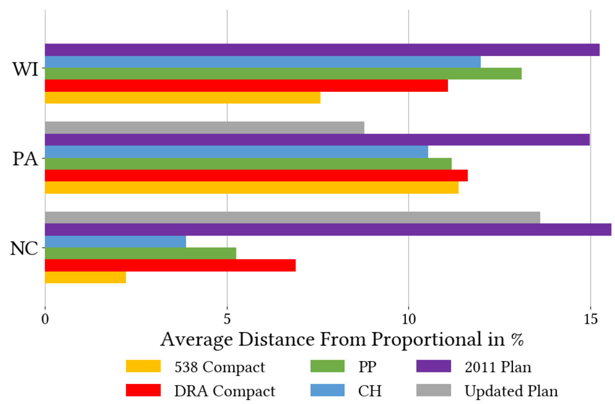

Figure 4: Average signed distance from the R swing curve to the $y = x$ line over the range $[0.4, 0.6]$ (Equation 1). In each state there are four compact plans, DRA's, 538's, our Convex Hull, and our Polsby-Popper. And in most states two implemented ones (2011 and Updated). Each score in PA uses 2012 presidential data, while each plan in WI and NC uses 2016 presidential data. The WI 2011 plan was not struck down so there is no WI updated plan. The full version has scores for other elections and comparison ranges.

As mentioned, compactness is often a primary goal when redistricting. It has even been suggested it is a path to partisan neutrality [22]. A priori, it is not clear if compact plans are more free of partisan bias than less compact ones. Thus, in this section we study plans designed to optimize various notions of compactness, contrasting them with the currently used plans. We find that optimizing for various definitions of compactness can reduce the signed partisan bias (Equation 1), relative to the plans used in real life. Despite this, we find a persistent bias to these compact plans, i.e., they favour the R party despite being optimized for a non-partisan goal.

We find that optimizing for any form of compactness yields plans that have improved partisan fairness relative to the plans enacted in 2011, according to our signed partisan bias score (over the range $[0.4, 0.6]$). This improvement is consistent across every states and independent of the compactness measure optimized. Figure 4 shows our signed partisan bias score (closer to zero is fairer) for various plans in three states using the presidential elections of interest.

This improvement is sometimes extreme: in NC, the 2011 districting (with a 17% robust partisan bias towards the R party) is more than two times as biased as any of the compact plans. It is worth noting that both NC and PA 2011 plans were struck down by the courts for being overly biased. The NC 2011 plan was found to disenfranchise minority voters [8], while in PA the plan was found to disenfranchise Ds [27]. The updated plan from 2016 (with an 11% robust partisan bias towards the R party) was still almost twice as biased as any compact plan.

Interestingly, the updated plans from 2016 in NC and 2018 in PA seem dissimilar. While the updated NC plan is still significantly more R-biased than any of the compact ones, the opposite holds for the new PA plan. Its R bias is lower than the compact plans, although it is, of course, less compact according to almost any metric (Table 3). It has been suggested the new PA plan was designed with partisan proportionality in mind [7]. That said, it should be noted that neither of the compact plans are designed to optimize the Equation 1. In each state, when we use our algorithm to optimize for this metric specifically we find plans that have near-zero bias according to Equation 1. These are not shown in Figure 4 because the bars would be virtually invisible.

For all plans, in all states, both elections, and both ranges of comparison there is one consistent pattern: The R party always has a positive score in our metric, and from symmetric considerations noted above, this means a negative D score. That is, the more rural party can expect to gain more seats than its proportional voter share. This includes every single plan designed to optimize some notion of compactness (supposedly the 2018 PA was also designed to consider proportional fairness). The advantage was significant. On average in PA this was a 10% (and often higher) advantage.

6.2 An $\alpha\%$ of the Vote can be an $\alpha\%$ of the Seats

As we saw with the updated PA plan and the compact plans, optimizing purely for compactness may not be the be the most effective way to eliminate partisan bias. For each state we use our algorithm to show there is a plan that effectively has no rural bias, a "fair" plan. We use use our algorithm to optimize for the unsigned partisan bias, Equation 2 over the range $[40, 60]$ (optimizing for Equation 1 can lead to plans with huge jumps in the swing curve). In each state the resulting plans had an unsigned partisan bias score from 3.5 to 10 times lower than any of the the existing and compact plans. Furthermore, in these plans there is near zero advantage at for either party, when measured by the signed partisan bias.

This gain in fairness comes at a cost of compactness. While we can make these fair plans more compact by optimizing for the Polsby-Popper score, with the constraint that the score from Equation 2 should remain near the best found, we are still nowhere near the best compactness scores. As Table 3 shows, in PA, if we keep the unsigned bias under 2% (the best score was 1.5%), the only plan with worse compactness scores is the 2011 plan.

Figure 5 shows this sacrifice in compactness ensures partisan fairness, over the $[40\%, 60\%]$ vote range. For either party, an $\alpha\%$ of the vote means a majority in an $\alpha\%$ of the seats. This is unlike other plans (such as our CH plan in Figure 3) where both parties are far from proportional. An exact description of our optimization,

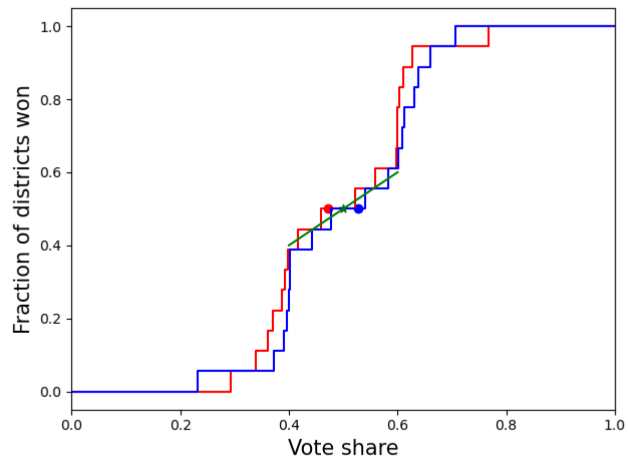

Figure 5: Uniform swing for $R(D)$ in red (blue), in the 2012 presidential election in PA under our fair but compact plan. Vertical axis shows the fraction of districts won; horizontal axis the vote fraction. The dots on the party curves indicate the actual election outcome (0 swing). The green line is the range of proportional outcomes on the range $[0.4, 0.6]$. A green star marks the point $(1/2, 1/2)$.

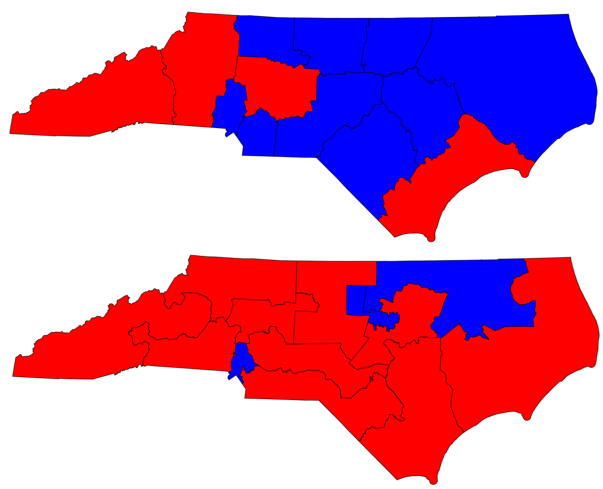

Figure 6: NC districts (R wins in red); (D wins in blue) based on the 2016 NC election data. Top, our plan, optimizing the Polsby-Popper score while maintaining 9 D wins. Bottom, NC's 2016 court mandated plan.

and similar results showing a near identical tradeoff in fairness and compactness for WI and PA can be found in the full version online.

7 DESIGNING PARTISAN PLANS

In this section we examine what limits compactness thresholds impose on partisan gerrymandering. For compactness, we use the average Polsby-Popper score of a plan. To measure gerrymandering ability, we use *gerrymandering power* (introduced in Borodin et al.

[4]). For a particular election, the gerrymandering power of party p is defined as the difference between the share of seats it can optimally gerrymander to win and the seat share it would have received in a purely proportional election. A high gerrymandering power indicates there is a plan that stretches p's vote share into a disproportionately large number of districts. A low (or negative) gerrymandering power indicates p is unable to stretch its vote into many extra wins (or even win a proportional number of seats).

To gerrymander for party p while staying compact, we run our algorithm with the objective of generating plans which are as compact as possible while maintaining k wins (at least 50% of the vote) for party p. To ensure a diversity of election outcomes we use the elections specified in Section 4. As we saw, our algorithm is capable of generating highly compact districts and highly partisan districts. Unsurprisingly, we find it performs quite well when combining these goals. As Figure 6 shows, our compact gerrymander easily passes the eye test, especially when compared to the implemented plan. In NC our algorithm can stretch the number of districts Rs win to all 13 using the 2016 election data, all while creating a plan more compact than the existing one (in the existing plan, Rs won 10 of 13). We can also create a map, more compact than the existing, and where Ds where they win 8 more seats.

We vary k from the most partisan outcome (maximal number of districts won with no compactness constraint), to the most compact outcome (number of seats won by p in the Polsby-Popper compact plan from Section 6).

7.1 Effect of Increasing Compactness

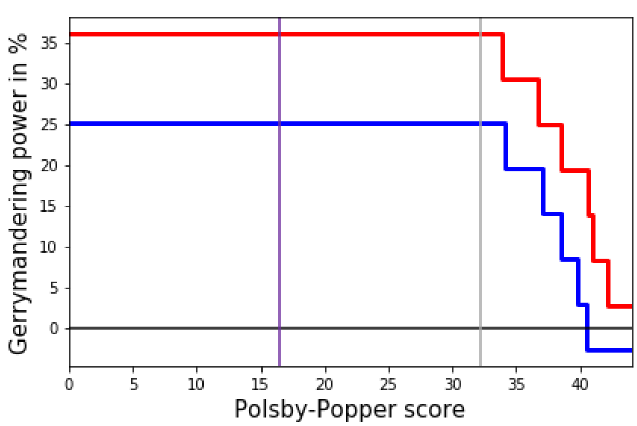

Figure 7: Gerrymandering power in PA when faced with a minimum required Polsby-Popper score using data from the 2012 PA presidential election. R **in red;** D **in blue. The vertical purple (grey) line is the Polsby-Popper score of the 2011 congressional plan (2018 court mandated plan). Average distance between the two curves is a** 10.8% R **advantage.**

Figure 7 shows in PA, increasing the required Polsby-Popper score lowers the gerrymandering power of both parties. But, to have an impact, we require a steep increase beyond current compactness levels. Similar results for WI and NC can be found in the full version.

Additionally, compactness requirements are unable to entirely remove the urban disadvantage. For almost any Polsby-Popper score, the R gerrymandering power is well above the D one. In PA there is no requirement level where the Ds have an advantage. In NC and WI there is a brief period of near maximum compactness requirements where the Ds have a small, temporary, advantage. In WI, when the compactness requirement is lower than that of the current plan the democrats can have a minuscule advantage, but more stringent requirements give a large R advantage. The average distance between the two curves shows a 10% R advantage in gerrymandering power in PA and NC, and 4% in WI.

In every single state, even with the most extreme compactness requirements, Rs are able to stretch their vote share beyond proportional. Conversely, Ds, have a negative gerrymandering power when compactness requirements are high – no legal plan reaches their proportional allocation.

8 DISCUSSION

In this work we introduced a modular and powerful automated redistricting technique. Our technique can generate plans comparable to ones from human experts for both partisan and non-partisan goals. Our method is able to generate compact districts, far more compact (according to various metrics) than the plans used in practice or the ones produced by electoral experts. While these plans, which were optimizing compactness, reduce partisan bias, we find they do not eliminate it and still always favor the rural party. Furthermore, we use our algorithm to explore the effects of an often proposed solution to gerrymandering, compactness restrictions. We find that while this can reduce the ability of either party to gerrymander, the potential for some degree of gerrymandering remains, and the rural party can still gerrymander more than its urban counterpart. These results contribute to growing evidence that the urban-rural divide leads to imbalanced outcomes that disadvantage the urban party. Despite this lopsidedness we show there are plans which are near totally proportionally fair, but to achieve this we must sacrifice a significant amount of compactness.

We see many applications of our algorithm for future work, including further examination of the rural party advantage. We have preliminary results on extending the gerrymandering power metric used in Section 7. We find that even if we require districts to have large margins of victory, compactness constraints still more negatively impact the urban party. We also intend to use our algorithm to explore the tradeoff between partisan fairness and non-partisan goals, such as not splitting counties. We began exploring the tradeoff between fairness and compactness in Section 6.2, the next step would be mapping out its Pareto frontier.

Possibly more important than partisan considerations is ensuring that minority voices are heard in the political process. Beyond the basic requirements of majority-minority districts that satisfy the 1965 Voting Rights Act, the function for measuring minority representation could be quite intricate and difficult for human experts to analyze and optimize for. We believe that such non-trivial objective functions, along with restrictions such as compactness, make this problem an ideal application of our algorithm.

Acknowledgments. This work was supported by Israel Science Foundation grants #1965/20 and #3152/20.

REFERENCES

[1] Yoram Bachrach, Omer Lev, Yoad Lewenberg, and Yair Zick. 2016. Misrepresentation in District Voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. New York, 81–87.
[2] Amariah Becker and Justin Solomon. 2021. Redistricting Algorithms. *Political Geometry* (2021). arXiv:2011.09504 [cs.DS]
[3] Bill Bishop. 2009. *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart.* Mariner Books.
[4] Allan Borodin, Omer Lev, Nisarg Shah, and Tyrone Strangway. 2018. Big City vs. the Great Outdoors: Voter Distribution and How it Affects Gerrymandering. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. Stockholm, Sweden, 98–104.
[5] David Butler. 1992. The electoral process The redrawing of parliamentary boundaries in Britain. *British Elections and Parties Yearbook* 2, 1 (1992).
[6] Amittai Cohen-Zemach, Yoad Lewenberg, and Jeffrey S. Rosenschein:. 2018. Gerrymandering Over Graphs. In *Proceedings of AAMAS*. Stockholm, Sweden, 274–282.
[7] Nate Cohn, Matthew Bloch, and Kevin Quealy. 2018. The New Pennsylvania Congressional Map, District by District. New York Times.
[8] 2019. *Common Cause v. Lewis*. Wake County Superior Court.
[9] Charlie Cook. 2018. The Cook Political Report. https://cookpolitical.com/. Accessed: 2021-01-01.
[10] Daryl DeFord, Moon Duchin, and Justin Solomon. 2021. Recombination: A Family of Markov Chains for Redistricting. *Harvard Data Science Review* (31 3 2021). https://doi.org/10.1162/99608f92.eb30390f https://hdsr.mitpress.mit.edu/pub/1ds8ptxu.
[11] Moon Duchin. 2018. Outlier analysis for Pennsylvania congressional redistricting. (2018). https://mggg.org/publications/md-report.pdf
[12] M.E. Dyer and A.M. Frieze. 1985. On the Complexity of Partitioning Graphs into Connected Subgraphs. *Discrete Applied Mathematics* 10 (1985), 139–153.
[13] E. J. Engstrom. 2006. Stacking the States, Stacking the House: The Partisan Consequences of Congressional Redistricting in 19th Century. *APSR* 100 (2006), 419–427.
[14] Robert S. Erikson. 1972. Malapportionment, Gerrymandering, and Party Fortunes in Congressional Elections. *The American Political Science Review* 66, 4 (1972), 1234–1245. http://www.jstor.org/stable/1957176
[15] Dan S. Felsenthal and Nicholas R. Miller. 2015. What To Do about Election Inversions under Proportional Representation? *Representation* 51, 2 (2015), 173–186.
[16] Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. 2018. A New Automated Redistricting Simulator Using Markov Chain Monte Carlo.
[17] John N. Friedman and Richard T. Holden. 2009. The Rising Incumbent Reelection Rate: What's Gerrymandering Got to Do With It? *Journal of Politics* 71 (2009), 593–611. Issue 02. https://doi.org/10.1017/S0022381609090483
[18] Bernard Grofman. 1983. Measures of bias and proportionality in seats-votes relationships. *Political Methodology* (1983), 295–327.
[19] Bernard Grofman and Gary King. 2007. The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering after LULAC v. Perry. *Election Law Journal* 6, 1 (2007), 2–35.
[20] B. Grofman, W. Koetzle, and T. Brunell. 1997. An Integrated Perspective on the Three Potential Sources of Partisan Bias: Malapportionment, Turnout Differences, & the Geographic Distribution of Party Vote Shares. *Electoral Studies* 16, 4 (1997), 457–470.
[21] Wes Gurnee and David B. Shmoys. 2021. Fairmandering: A column generation heuristic for fairness-optimized political districting. arXiv:2103.11469 [cs.CY]
[22] Christopher Ingraham. 2016. This is actually what America would look like without gerrymandering. https://www.washingtonpost.com/news/wonk/wp/2016/01/13/this-is-actually-what-america-would-look-like-without-gerrymandering/. Accessed: 2022-01-07.
[23] Samuel Issacharoff. 2002. Gerrymandering & Political Cartels. *Harvard Law Review* 116, 2 (2002), 593–648.
[24] Brian Klaas. 2017. Gerrymandering is the biggest obstacle to genuine democracy in the United States. So why is no one protesting? *Washington Post* (10 February 2017).
[25] Justin Levitt. 2020. Where are the lines drawn? https://redistricting.lls.edu/where.php/
[26] Yoad Lewenberg, Omer Lev, and Jeffrey S. Rosenschein. 2017. Divide and Conquer: Using Geographic Manipulation to Win District-Based Elections. In *Proceedings of the 16th International Coference on Autonomous Agents and Multiagent Systems (AAMAS)*. São-Paulo, Brazil, 624–632.
[27] 2018. *League of Women Voters of Pennsylvania v. Commonwealth of Pennsylvania*. Number 159 in MM 2017. Supreme Court of Pennsylvania.
[28] MGGG. 2020. Metric Geometry and Gerrymandering Group; https://mggg.org. Accessed: 2021-01-01.
[29] New York Times Editorial Board. 2017. A Shot at Fixing American Politics. *New York Times* (30 September 2017), A22.
[30] NBC News. 2020. North Carolina House Election Results 2020. https://www.nbcnews.com/politics/2020-elections/north-carolina-house-results. Accessed: 2022-01-07.
[31] Rick Noack. 2016. The urban-rural divide that bolstered Trump isn't just an American thing; it's prevalent in Europe, too. *Washington Post* (November 27 2016).
[32] Wesley Pegden, Ariel D. Procaccia, and Dingli Yu. 2017. A partisan districting protocol with provably nonpartisan outcomes. ArXiv:1710.08781.
[33] Clemens Puppe and Attila Tasnádi. 2008. A computational approach to unbiased districting. *Mathematical and Computer Modelling* 48, 9-10 (2008), 1455–1460. https://doi.org/10.1016/j.mcm.2008.05.024
[34] Jonathan Rodden. 2019. *Why Cities Lose: The Deep Roots of the Urban-Rural Political Divide.* Basic Books.
[35] 2019. *Rucho v. Common Cause*. Number 18-422 in 588 U.S. Supreme Court of the United States.
[36] P.H. Schuck. 1987. The Thickest Thicket: Partisan Gerrymandering and Judicial Regulation of Politics. *Columbia Law Review* 87, 7 (1987), 1325–1384.
[37] Nate Silver. 2020. 538: The Gerrymandering Project. https://fivethirtyeight.com/tag/the-gerrymandering-project/.
[38] Andranik Tangian. 2010. Computational application of the mathematical theory of democracy to Arrow's Impossibility Theorem (how dictatorial are Arrow's dictators?). *Social Choice and Welfare* 35, 1 (June 2010), 129–161.
[39] René van Bevern, Robert Bredereck, Jiehua Chen, Vincent Froese, Rolf Niedermeier, and Gerhard J. Woeginger. 2015. Network-Based Vertex Dissolution. *SIAM Journal on Discrete Mathematics* 29, 2 (2015), 888–914.
[40] William Vickrey. 1961. On the Prevention of Gerrymandering. *Political Science Quarterly* 76, 1 (1961), 105–110. http://www.jstor.org/stable/2145973
[41] Samuel S.-H. Wang. 2016. Three Tests for Practical Evaluation of Partisan Gerrymandering. *Stanford Law Review* 68 (June 2016), 1263–1321.
[42] David Wasserman. 2018. Hating Gerrymandering Is Easy. Fixing It Is Harder. FiveThirtyEight. https://fivethirtyeight.com/features/hating-gerrymandering-is-easy-fixing-it-is-harder/
[43] Archer Wheeler and Philip N. Klein. 2020. The impact of highly compact algorithmic redistricting on the rural-versus-urban balance. In *SIGSPATIAL '20: 28th International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, November 3-6, 2020.* 397–400.