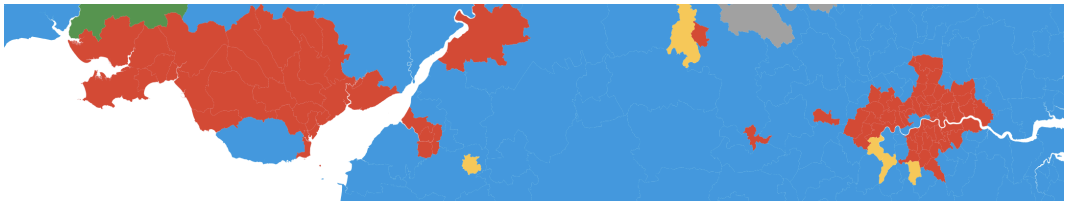


A Reinforcement Learning Algorithm for Fair Electoral Redistricting in Parliamentary Systems

Callum Evans
2nd May 2022

University of Exeter

Abstract

The primary goal of a strong democracy should be to most accurately represent its electorate, and the way they are divided into electoral districts can drastically affect this. As a result, many methods have been proposed to algorithmically generate fairer boundaries, the majority of which focus on eliminating bias through qualitative measures, however, these often fail to produce truly fair results. This paper, therefore, aims to demonstrate how fairness can and should become a higher priority within our electoral systems through the development, implementation and application of a new reinforcement learning-based method for algorithmic redistricting that directly optimises for fairness. Specifically, the model has been applied to the parliamentary system of the UK, representing a significant absence within the literature, meaning the paper also outlines a new metric for measuring fairness in parliamentary systems that directly rewards proportionality, the seats-votes difference. The algorithm has then been evaluated on the current parliamentary constituency boundaries in the UK, as well as the initial proposals for the 2023 boundary review, and was ultimately found to improve the map's fairness in all experiments. The paper subsequently concludes with some of the limitations of the model and the seats-votes difference and ways the redistricting algorithm could be further expanded in the future.

I certify that all material in this dissertation that is not my own work has been identified.

Source code: `https://github.com/calluume/UK_redistricter`

Contents

1 Introduction **1**

2 Background Work **2**
2.1 Literature Review . . . 2
2.1.1 Evaluating Electoral Boundaries . . . 2
2.1.2 Related Works . . . 3
2.2 Project Specification . . . 3

3 Problem Definition **4**

4 UK Ward-level Dataset **4**
4.1 Data Sources . . . 4
4.2 Dataset Challenges . . . 5
4.3 Linear Regression Model . . . 5

5 Original Methodology **6**

6 Redistricter Methodology **8**
6.1 Selection Methods . . . 8
6.1.1 Hybrid Selection . . . 8
6.1.2 Voronoi Selection . . . 9
6.1.3 Acceptance Selection . . . 9
6.2 Descent-based Local Search . . . 10
6.3 Reinforcement Learning . . . 11
6.4 Multi-stage Optimisation . . . 11

7 Fitness Metrics **11**
7.1 Compactness Measures . . . 12
7.2 Fairness Measures . . . 12
7.2.1 Fairness in Parliamentary Systems . . . 12
7.2.2 The Efficiency Gap & the Wasted-vote Score . . . 13
7.2.3 Seats-Votes Difference . . . 14

8 Evaluation & Results **14**
8.1 Overview . . . 14
8.2 Model Evaluation . . . 14
8.2.1 Current Boundaries . . . 14
8.2.2 2023 Boundaries . . . 17
8.2.3 Generation of New Boundaries . . . 18
8.3 SV Difference Evaluation . . . 18

9 Conclusions **20**
9.1 Future Work . . . 20
9.2 Final Conclusions . . . 20

References **21**

Appendices **24**

1 Introduction

The way that a country decides to separate itself into electoral districts, or constituencies, can have large and unforeseen consequences long into the future, creating meaningful impacts on how its people are represented in government. As a result, this process of drawing or updating electoral district boundaries, also called electoral redistricting, is often a long, laborious and sometimes controversial one that involves many iterative stages of planning and public consultations [1]. Typically, this is due to the many factors that have to be considered during electoral redistricting, as not only do boundaries have to follow the rules and constraints defined by law, they should also be an accurate reflection of the country's electorate and cannot show a bias towards any political party or demographic. However, there are still many examples of unfair boundaries being proposed in recent years, especially in the US, where many electoral district maps are still drawn by the governing party [2]. Even in the UK, where independent commissions decide boundaries, the sixth periodic review of the Westminster parliamentary constituencies was abandoned after two attempts in 2013 and 2018 [3] after criticism surrounding a potential bias towards the Conservative party and a reduction in Scotland and Wales' influence in government, particularly due to attempts to reduce the number of MPs from 650 to 600.

Therefore, whilst many methods have been proposed to detect or quantify the effect of biased boundaries, other research has focused on creating algorithmic methods for electoral redistricting in order to completely remove a bias inherent in human-made maps. A majority of these methods aim to do this by optimising boundaries solely for quantitative factors, such as population distribution or district geometries [4, 5, 6], however, others reject this idea and instead propose that only maps that consider potential election outcomes can be fair [7, 8]. Following this concept, boundaries could subsequently be generated in a way that rewards proportionality between parties' vote shares and seat shares, maintaining the local representation advantages of first-past-the-post electoral systems whilst aiming toward the ability of proportional vote systems to fairly represent the electorate [9].

Therefore, for this project, I will be proposing a new reinforcement learning algorithm for electoral redistricting that is able to directly optimise boundaries for fairness and their ability to represent the electorate. As a result, this paper will outline the methodology, implementation and analysis of this new algorithm and will describe its application to a parliamentary system, specifically the UK. As the majority of the literature focuses on the US and, to the author's knowledge, no research has been done into how redistricting methods apply to parliamentary systems, the extra necessary considerations made are also outlined. This includes the development of a fairness measure, the seats-votes difference, that can evaluate boundaries for any number of parties, which many of the existing fairness measures struggle to do, whilst still making direct comparisons to the proportional vote. Fairness calculations will be done using the results from the 2017 general election, estimated at the ward level to evaluate boundary changes, whilst some geographic considerations, mainly the compactness of constituencies, are used in order to ensure boundaries remain somewhat logically shaped. After the implementation of the model, parameter tuning and evaluation of the algorithm and fairness metric were carried out to evaluate how well it is able to improve the current UK boundaries. In light of the current 2023 boundary review in the UK, which began in January 2021 [3], and criticisms of their changes, the initial boundary proposals have also been evaluated in comparison to the current constituency map.

This paper is organised as follows. First, section 2 contains an overview of the background research done for the project, with a focus on the evaluation of electoral boundaries and existing redistricting algorithms. Section 3 then explains the context of the model, specifically the rules that must be followed by any of the UK boundary commissions' proposals and subsequently the final model's proposals. Section 4 then outlines the creation of the ward-level dataset used in the model, which was fundamental for the fairness calculations used. Section 5 describes the methodology of the original algorithm that was used as the basis for the proposed algorithm, with section 6 how the current proposed algorithm has been adapted to the problem. Section 7 then describes the development and implementation of the algorithm's fitness function and challenges of measuring fairness in a parliamentary system. Section 8 then outlines the parameter tuning and model experiments done to evaluate the model's performance, as well as an evaluation of the initial boundary proposals for the 2023 review currently happening in the UK. Lastly, section 9 concludes the paper with a critique of the model and potential ways it could be expanded and improved in the future.

2 Background Work

2.1 Literature Review

2.1.1 Evaluating Electoral Boundaries

Gerrymandering can be defined as the process of manipulating the boundaries of electoral districts in a way that favours certain political parties or groups over another [10]. Therefore, this means that those who decide electoral boundaries are can favour themselves in the results of future elections, called partisan gerrymandering, or it can be used to reduce the influence of a particular demographic, ethnic, religious or class group within government. For example, this was the case in Northern Ireland in the 1920s, where parliamentary constituency boundaries were drawn to help protestant unionist parties in areas where they did not have the majority [11]. The practice itself is done mainly through two methods, packing and cracking [12]. Here, packing aims to concentrate a voting group into a small number of districts, allowing them to win a small number of seats, but usually stopping them from winning a majority. Cracking, on the other hand, splits a group into many different districts, usually spreading them thinly enough that they are not able to win any seat.

Countries that use first-past-the-post (FPTP) electoral systems, which is a system where voters choose a single candidate to vote for and the candidate with the most vote wins, are usually the most vulnerable to gerrymandering. This is because candidates are still able to win even if they have received less than 50% of the total votes cast, especially in elections with more than 2 candidates, which highlights how prone the system is to gerrymandering due to how easily votes can be wasted [13]. This idea of a "wasted vote" is therefore important to one of the most commonly used metrics for measuring the fairness of electoral boundaries, the efficiency gap [14]. The efficiency gap was proposed by Stephanopoulos and McGhee in 2015 and relies on the fact that districts with a single representative produce a lot of wasted votes, and so, it assumes that the goal of a partisan gerrymander is to most efficiently win seats. This means the efficiency gap can be calculated as the difference between each candidate's wasted votes, where wasted votes account for votes wasted through cracking (i.e those cast for a losing candidate), and packed votes (i.e surplus votes cast for the winning candidate above what is needed to win, 50% in 2 party systems). This can also be very easy to interpret, which Stephanopoulos and McGhee describe could be compared to thresholds defined by lawmakers to denote potential gerrymandering. However, they also note the efficiency gap can display some instability over time and can produce unexpected results with significant differences in vote share.

Another method for measuring fairness is the plotting of the seats-votes (SV) curve [15], first proposed by Theil in 1969, which looks directly at the relationship between votes gained and seats won. This is done simply by plotting the percent of seats that a party would win for a given vote share, with an S-shaped curve representing an unbiased election, and a large area between the parties representing a biased one. One of the main differences between the two methods is that the SV curve and resulting metrics are calculated using hypothetical election results and although this has the advantage of being able to forecast elections under a wide range of situations, results rely on the model used and can be more computationally expensive to calculate and more difficult to interpret[16].

Alongside metrics for measuring fairness in elections, district compactness is often used for detecting gerrymandered districts, which ultimately follows the idea that gerrymandered boundaries tend to have unusual or illogical shapes. This subsequently gave the name to the practice after, in 1810, the governor of Massachusetts, Elbridge Gerry, manipulated the senate state boundaries to benefit his party, resulting in a shape that was later described as looking like a salamander (hence the "gerrymander") [17]. Therefore, many compactness measures will compare the shape of a district's geometry to that of some base geometric object, which include; the Reock score [18] which finds the ratio of a district's area to that of its minimum bounding circle, the Polsby-Popper score [19] that instead compares it to the area of the circle whose circumference equals the district's perimeter, or simply using the ratio of a district's length to width. Ultimately, however, as noted by Kaufman, King and Komisarchik, whilst many US state constitutions state that legislative boundaries should be compact, there usually is no formal definition, and they instead rely on the idea people recognise gerrymandering when they see it [20]. This also means there is no standard measure and any choice in which to use is generally subjective, although computability and the resolution of the maps used can be a factor [21].

2.1.2 Related Works

Many methods for algorithmically generating electoral boundaries have already been proposed within the literature, with the majority focusing on non-machine learning methods and qualitative factors that optimise boundaries solely for compactness and population distribution, and not fairness. Although these still typically see improvements in fairness compared to their real-world counterparts, most share the common theme that these methods are proposed as a solution to an inherent bias in maps drawn by humans, and so, they should not use any characteristics related to voter intention in order to remove any impact on elections from consideration. For example, the splitline method [4] proposed by Smith in 2011 functions deterministically to only produce one possible district map for a given state, meaning that it cannot be manipulated. The vast majority of algorithms have also been applied to the US, where admittedly gerrymandering is more of an issue than some other countries, with no research having been done into redistricting in parliamentary systems.

Other well-known methods include the optimal compactness algorithm [5], used to improve existing boundaries through randomly swapping census blocks between districts in a gradient descent approach, which was found to improve the efficiency gap of the current US congressional map by 4% [22]. The Markov Chain Monte Carlo (MCMC) method [6] frames the problem as a graph partitioning problem, and expands earlier Monte Carlo simulations by splitting the network into random subgroups and accepting districts according to an acceptance probability. In testing their algorithm, Fifield et al. found that they were able to eliminate the Democrat bias in the 2008 Pennsylvania map by swapping only 3% of precincts, however, the MCMC algorithm cannot currently include multiple constraints and performance is dependent on the mixing of the resulting Markov Chain. Meanwhile, some population-based approaches have also been proposed, such as the open-source "Antimander" project [23], which again frames the problem as a graph partitioning problem, but instead uses NSGA-II, an evolutionary multi-objective algorithm by Deb et al. [24]. Another example is the "Auto-Redistricter" project [7] which uses a genetic algorithm and was also one of the first to include a fairness heuristic in its optimisation alongside geometric considerations. The later "fairmandering" project [8] has similar goals and also explicitly rejects the idea that redistricting should not consider the effect on election outcomes. The project by Gurnee and Shmoys proposes a column generational heuristic approach and proposes stochastic hierarchical partitioning to represent the trillions of possible unique plans for a given number of districts in a tree structure. This also differs from the other outlined methods as this approach aims to generate a list of many possible maps that can then be given to decision-makers who can choose which map to use.

However, very few machine learning methods have been proposed for drawing electoral boundaries, such as the intelligent clustering approach by Hung and Chen [25]. Some other grouping algorithms have been proposed that could be applied to the problem of redistricting, such as the reinforcement learning algorithm for grouping problems by Zhou, Hao and Duval in 2016 [26], which is outlined in more detail in section 5. This method originally used graph colouring as its use case, however, the authors note how this could be applied as a general algorithm for grouping problems.

2.2 Project Specification

In order to further research absences in the literature regarding the lack of machine learning-based approaches and applications to parliamentary systems, the project would therefore consist of the implementation and evaluation of a reinforcement learning algorithm for redistricting, based on the algorithm proposed by Zhou, Hao and Duval. The proposed redistricting algorithm would also directly optimise for fairness using the efficiency gap or similar metrics, again representing an area that lacks research in the literature. The project would use the UK as its application, with electoral wards as the base geographic unit, meaning the initial stage of the project would involve the creation of a ward-level election results dataset for use in the fairness calculations. The final model could then be evaluated based on how well it can improve the fairness of the current UK boundaries, and could also be used to evaluate the boundaries from the initial proposal of their 2023 periodic review.

3 Problem Definition

The government of the United Kingdom uses a parliamentary monarchy system and can be classified as a bicameral Parliament, meaning that it has two chambers; the House of Lords and the House of Commons [27]. Whilst the House of Lords has 800 appointed members, the House of Commons consists of 650 directly elected members and holds the majority of the legislative power of the Houses of Parliament. Each member of Parliament (MP) within the Commons represents a parliamentary constituency with an electorate of around 70,000, which allows for greater local representation within government [28].

As outlined by 'the Parliamentary Constituencies Act 1986' [29], the boundaries of these constituencies are decided by four independent commissions, with one for each country within the UK. The act also defines a strict set of rules that any proposed boundaries must follow, which include; constituencies cannot be larger than 13,000km^2 and their electorates must be within 5% of the electoral quota, which is the average electorate size for mainland constituencies. Exceptions are made for smaller electorates in island constituencies or those with areas above 12,000km^2, and a set of 4 protected constituencies are also defined that cannot be changed (which are the Isle of Wight, Na h-Eileanan an Iar, Orkney and Shetland, and Ynys Môn). These constraints are important to note as the proposed redistricting algorithm should also obviously have to generate boundaries that fit these rules, whilst being able to best optimise for fairness.

4 UK Ward-level Dataset

The four boundary commissions within the UK typically use electoral wards as the base geographic unit for creating constituency boundaries, which is done as wards are generally well-defined, well-understood and are indicative of areas with a broad community of interest [30]. Therefore, although constituency boundaries do not always follow ward boundaries, due to a variety of factors including geographical considerations, local government boundaries or the commission's ability to produce constituencies following electorate or area constraints, electoral wards are generally accepted as the base for constituencies. As a result of this, and the fact that any country-scale redistricting model could not reasonably be expected to produce boundaries that dynamically diverge from already established boundaries or simulate the iterative consultations that the real-world boundaries undergo, it was decided that electoral wards would be used as the main base geographic model for use in the dataset.

However, from this, the problem arises that the UK does not collect election results on such a small scale, the data from which is fundamental for the fairness optimisation aspect of the model. Therefore, ward-level election results had to be created for use within the model, done using a linear regression model and demographic statistics, similar to the model used by Electoral Calculus [31], a political forecasting company in the UK.

4.1 Data Sources

Ward-level demographic data was mainly collected from datasets collected for the 2011 census, the most recent census with publicly available data, with datasets for each country obtained separately. Specifically, England and Wales release census data through Nomis, from the Office of National Statistics (ONS), whilst Scotland and Northern Ireland use Scotland's Census and the Northern Ireland Statistics and Research Agency (NISRA) [32, 33, 34] respectively. In total, 5 different demographic factors were chosen, which were age, education, economic activity, household type and ethnicity, each with their own subclasses that gave a total of 33 inputs for the model. Datasets containing boundary shapefiles, ward ONS codes and constituency lookups were then also obtained from ONS, Boundaries Scotland and Open Data NI releases.

Figure 1 therefore shows some mapped samples of the dataset, including the network of ward adjacencies generated from the shapefile data and examples of the ward-level demographic data collected for use in the model.

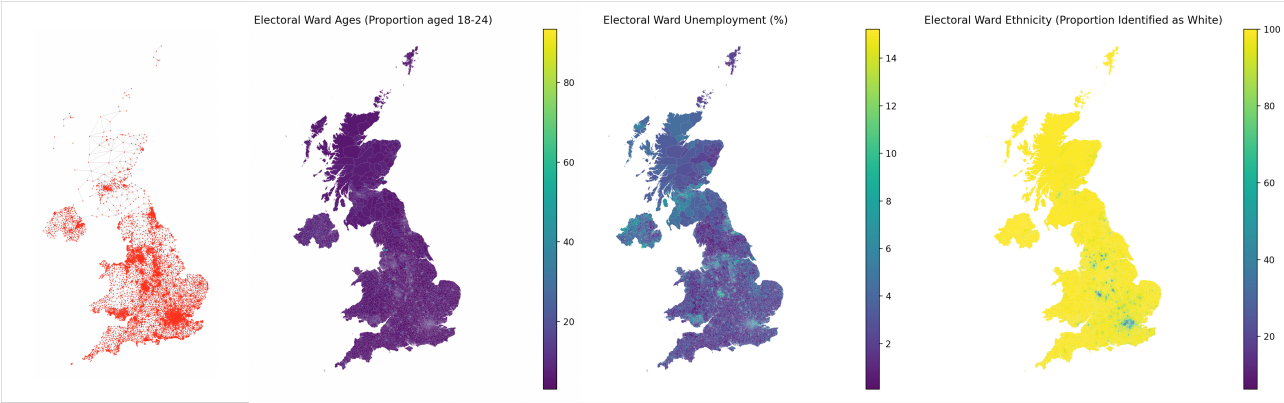


Figure 1: Examples of the ward-level demographic statistics and geographic information collected for use in the model. From left to right, this includes; the network of ward adjacents generated from the boundary shapefiles, proportion of each ward's population aged 18 to 24, ward unemployment and proportion of the ward's population identified as white.

4.2 Dataset Challenges

Some small modifications and decisions had to be made to the dataset to be suitable for use in the model, including:

- For privacy reasons, the ONS merges any small wards until they have a population above 1000 and over 400 households, at which point the ward is given a new "*E36*" code separate from the standard ONS electoral ward code beginning "*E05*". These merged wards are called census merged wards, and were created solely for use in the 2011 census. This mainly only affected two places within the UK, Bryher and St Agnes in the Isles of Scilly and the electoral wards within the City of London, and so, to ensure all data sources matched, Bryher was merged into the St Agnes ward and the City of London is considered as one ward, using its local authority boundaries.
- According to the available lookup datasets between constituencies and wards, constituency boundaries in Scotland appeared to deviate from ward boundaries a lot more than in the other countries in the model. This meant that each of the over 100 affected wards in Scotland that crossed constituency boundaries had to be manually placed into the most appropriate constituency. This was mostly done by placing wards in the constituency with the most overlap, however, some wards had to be placed in other constituencies in order to maintain contiguous constituencies. A similar decision had to be made in Northern Ireland with Derryaghy, which was split into a north and south ward, each placed into Lagan Valley and West Belfast respectively. Therefore, for this dataset, Derryaghy is considered as one ward and has been placed into Lagan Valley.

4.3 Linear Regression Model

Once the demographic data had been collected together, these were used as the input for a linear regression model to estimate ward-level election results. This was done by first merging all wards into their constituencies to generate constituency-level demographic data that was consistent with the ward data. Then, as constituencies are the smallest level that election results are collected at, a least squares linear regression model could then be used to fit the demographic data to the party vote shares for each constituency. The results were then normalised and changed to minimise the difference between the real election results and the results obtained from merging the wards back into their constituencies, ensuring the vote shares remain weighted according to the values generated from the linear regression model. Examples of the resulting ward-level support can be seen in Figure 2.

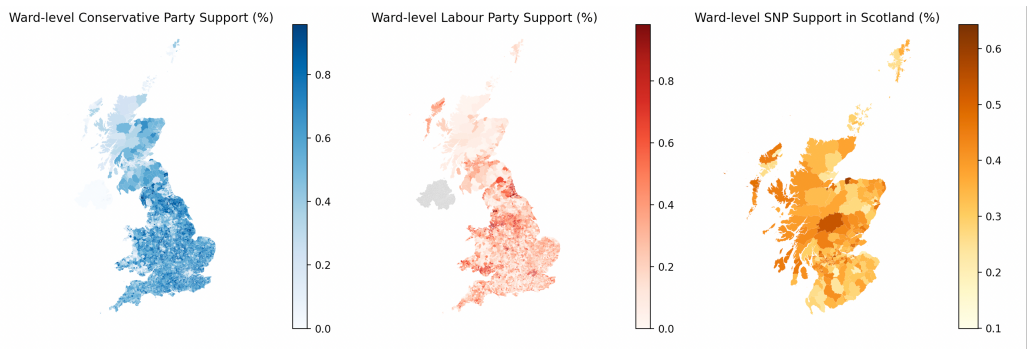


Figure 2: Examples of the ward-level party support dataset, displaying from left to right, support for the Conservative party, Labour party (who do not run in Northern Ireland elections) and the Scottish National Party (SNP).

A linear regression model was chosen due to the limited amount of data available to train the model on as, to be more accurate, the model was trained separately for each country and region. The model used 12 regions, including the 9 former English government regions and one region each for Scotland, Wales and Northern Ireland, and although this meant that the results for each region would be more tailored to that specific region, fewer constituencies could be used for training. For example, whilst there are 73 constituencies that can be used for training the model for Greater London, only 18 constituencies could be used for Northern Ireland. Regardless of this, using region-specific models was found to give much more accurate results compared to nationwide models, and allowed for easier accommodation of smaller, country-specific parties such as the Scottish National Party or Plaid Cymru.

Another element to note is that it was decided that election results from the 2017 general election would be used, as they represented a "cleaner" dataset. This is because, as may be evident, the political landscape within the UK can change very quickly and can vary dramatically between elections, particularly after the Brexit referendum in 2016. For example, a study by the British Election Study found that across the 2010, 2015 and 2017 elections, only 51% of voters voted for the same party every time [35]. Therefore, results from the 2017 election were chosen as it was thought there was less of an outsized impact from political events than in the 2015 or 2019 elections, whilst still remaining fairly recent. Another potentially interesting factor within the 2017 results is that the difference in vote shares of the two largest parties, the Conservatives and Labour, was much smaller in 2017 than in 2015 and 2019 (with a difference of 2.4%, compared to 6.5% and 11.4% respectively). Consequently, this may make 2017 more of an interesting use case when attempting to measure and optimise for fairness as the model will likely find it easier to change the election results. As a result, the model uses 13 parties as they were in 2017, including; Conservatives, Labour, Liberal Democrats, SNP, Plaid Cymru, Greens, UKIP, DUP, Sinn Fein, SDLP, UUP, Alliance and an "other parties" category (which includes independents and the speaker of the house).

Ultimately, however, the linear regression model was not able to replicate, with complete accuracy, the constituency-level results using its prediction of the ward-level results, as can be seen in appendix A. However, the estimated results and vote shares remained sufficiently close for use in this project and close enough to demonstrate the algorithm whilst still somewhat accurately reflecting the real world. Subsequently, developing a full prediction model for ward-level election data was deemed outside the scope of this project, although methods such as those used by Electoral Calculus, who use demographic data with multi-level regression, regularised prediction and post-stratification [31] could be used in order to improve the data used within the model.

5 Original Methodology

The redistricting model is based on a reinforcement learning-based local search algorithm for generalised grouping problems, proposed by Zhou, Hao and Duval in 2016, which originally used graph colouring as its case study. This is a problem within graph theory, first studied algorithmically as part

of Karp's 21 NP-complete problems from 1972 [36], which aims to colour vertices in a graph in such a way that no two vertices of the same colour share an edge. Their algorithm consists of four main components; a descent-based local search (DBLS), a group selection method, probability updating and smoothing stages, and the model was shown to give competitive results on popular graph colouring benchmarks, DIMACS and COLOR02, in comparison to other colouring methods. A diagram of the main schema, as outlined below, can be seen in Figure 3.

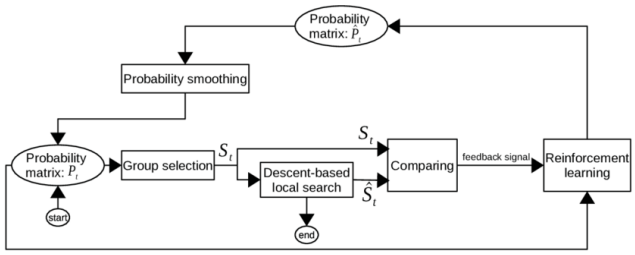


Figure 3: A schematic diagram of the original reinforcement learning based local search algorithm, as proposed by Zhou et al. (2016) [26]

The paper defines the partitioning problem as: given a set V of n distinct items, find an optimal set g, of k partitions. This is achieved through maintaining a probability matrix P, of size $(n \times k)$, where each element in the matrix p_{ij} $(0 < i \leq n$ and $0 < j \leq k)$, represents the probability that an item v_i belongs to the grouping g_j. All probabilities within the matrix are initialised as $1/k$, denoting that every item has an equal probability of belonging to any grouping. This probability matrix can then be used in the group selection method to generate the initial solution S_t, with the authors noting that a hybrid greedy/random selection performed best. This would choose to either assign items randomly or by taking the grouping with the highest probability, according to a defined probability, which allowed for the algorithm to better avoid local minima by stopping it from becoming too greedy.

Once the initial solution has been generated, a DBLS algorithm can then be applied to optimise the solution for any given metric. Here, items are iteratively and randomly moved between groupings, ensuring there is an increase in the fitness value of the solution with each move, essentially reflecting a hill climbing local search. The original authors do note how the performance could be improved through the use of a more powerful improvement algorithm, however, using hill climbing reflects the assumption that only good changes have been made to the solution, which can be useful in the application of a reinforcement learning model.

The last two stages represent the main reinforcement learning aspect of the algorithm, and first, involve a comparison between the initial solution S and the DBLS-optimised solution $\hat{S}$. Here, for each item v_i, we compare its group assignment in S and $\hat{S}$, and update the probability according to any changes. Specifically, if v_i has remained in its original grouping, we can assume that this is a good assignment as the DBLS did not affect it, and so the probability is increased by a reward factor. Conversely, if v_i has changed grouping during the DBLS, we can therefore assume that its new grouping produces a fitter solution, and so the probability for the original grouping is decreased by a penalisation factor and the new grouping is increased by a compensation factor. If no items have changed grouping throughout the DBLS, we can assume that S is the optimal solution and the algorithm is terminated. Otherwise, the probability smoothing stage can take place, which is where all probabilities above a certain threshold p_0 are reduced by multiplying by a smoothing factor $\rho < 1$. This stage is important as it is done to forget older decisions and avoid less helpful and potentially misleading assignments.

The paper then shows a set of comparative results of their method to other well-known graph colouring algorithms, specifically those based on advanced local search methods such as tabu search or simulated annealing. The reinforcement learning method showed equal performance in around half of their experiments on the DIMACS benchmark, with other instances performing worse, compared to the COLOR02 benchmark where the algorithm performed competitively, even against some more complex population-based algorithms. Ultimately, the authors conclude by saying how the simplicity of the local

search means that the reinforcement learning method will struggle to compete with other, typically more complex, graph colouring algorithms that combine several techniques like the population-based approaches. However, they then note how future work should be undertaken to explore how other more powerful improvement algorithms could be applied to the problem, as well as work to apply the algorithm to solve other grouping problems.

6 Redistricter Methodology

The redistricting model proposed here remains fairly similar to that proposed by Zhou et al., as outlined in section 5, however, a diagram showing the main schema for the adapted algorithm can be seen in Figure 4.

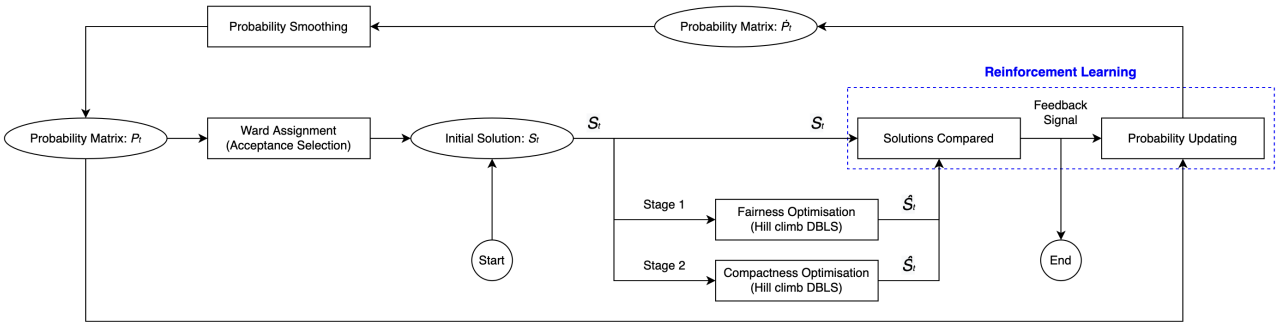


Figure 4: A schematic diagram of the new proposed redistricter algorithm.

6.1 Selection Methods

Three selection methods were developed throughout the project, as it was found that the selection method had a larger and more unpredictable impact on the algorithm's solutions than any other aspect of the model. As mentioned in section 5, the selection phase of the algorithm is when a solution is generated from the probability matrix, through the assignment of items to the chosen grouping. Therefore, even though 100 wards may be swapped during the local search, this only represents a small percent of the total number of wards, and many of these swaps may not even show an improvement in the fitness of the solution through a constituency seat changing hands. So, the selection method, where all of the 9,500 wards are assigned, can have a large effect on the solutions.

In the original paper, Zhou et al. use a hybrid selection method that either chose a grouping based on a random choice, or a greedy choice, where the grouping with the highest probability is chosen. This highlights how reinforcement learning can learn and improve through trial and error, meaning that any good selection method has to be able to fully exploit the knowledge learnt throughout the rest of the process. However, one constraint that does not exist within the original case study is that, in the context of redistricting, proposed constituencies have to be contiguous whilst still allowing for some reasonable geographical factors, unlike graph colouring where neighbouring nodes must be in different groupings. For example, island wards still have to be assigned to a constituency, usually the closest mainland constituency. Another change was made to initialise all values in the probability matrix as 0, except for the real ward assignments, which were set to 1 to maintain the real boundaries as the initial solution. Ultimately, the three methods developed were; hybrid selection, Voronoi selection, and acceptance selection. The acceptance selection method was ultimately found to perform best and has been used throughout the rest of the model's development and experimentation.

6.1.1 Hybrid Selection

The hybrid selection method was the first developed and remains fairly similar to that proposed in the original paper, except that instead of a random selection, a roulette wheel selection was chosen. This meant that, according to a defined threshold, wards would either be assigned to the constituency

with the highest probability, or the constituency would be chosen randomly, with choices weighted by their value in the probability matrix. A roulette wheel was added as this made contiguity constraints much easier to follow and ensured that the algorithm's performance was more stable. The method was then also changed to work iteratively, by assigning wards to a constituency only when it was adjacent to another ward in that constituency, and repeating this until all wards had been assigned. However, this was later modified further to create the acceptance method, as outlined below.

6.1.2 Voronoi Selection

The Voronoi selection method is the simplest of those developed and is based on Voronoi tessellations, however, it was found to not be suitable for redistricting. It works iteratively by first choosing a central ward for each constituency, done similarly to the hybrid selection, by either taking the ward with the highest probability of belonging to the constituency or by using a roulette wheel selection. Then, the model iterates through all unassigned wards and assigns them to the closest of these central wards, only where it would keep the constituency contiguous, i.e when there is a path to the central ward through the assigned wards. The exception to this is island wards, or wards without any adjacent wards, which are simply assigned to the constituency of the closest central ward. Due to the closeness and size of constituencies generated, the majority of wards can be assigned using this method, typically within around 5 iterations, however, this cannot account for wards surrounded by other constituencies, which can happen due to the shape of some wards, or islands that contain multiple wards. In these cases, contiguity is ignored and wards are assigned to the constituency of the closest assigned ward.

This was found to function well and be reasonably fast, however, the algorithm was not able to achieve the same performance using this approach. Specifically, although this method worked well for generating *compact* boundaries, it was not able to create *fair* boundaries. This likely is because the nature of the Voronoi approach, based on Voronoi tessellations, meant that constituencies would tend to spread out and grow larger. This meant that ward assignments became less precise, which is made worse by the fact that the method is not able to fully exploit knowledge stored in the probability matrix as it is only referred to once, during the assignment of central wards. Runs tested using the Voronoi method tended to have a much more even distribution of constituency sizes and regularly grouped fairly demographically different wards. Ultimately though, this resulted in most runs performing very poorly in terms of fairness, consistently having the Conservative party gaining around an extra 100 seats compared to the initial results.

However, in applications that aim to more quantitatively generate boundaries, such as those based solely on area and population constraints like many in the literature, the Voronoi selection method could perform much better, potentially with some modifications to analyse population distribution. It would then also be worth noting that in non-parliamentary systems or those that have a much smaller number of seats, having a more even area distribution between constituencies may be desirable, unlike the UK, where smaller, urban constituencies must exist alongside those that are much larger and more sparsely populated.

6.1.3 Acceptance Selection

The acceptance selection method was the last developed and was adapted from the original hybrid selection method, with the final method working using the method displayed in Figure 5. This approach was ultimately found to produce the most favourable results as it was able to most fully exploit the learnt knowledge stored in the probability matrix, whilst still allowing for some stochastic decisions, and was chosen as the main selection method used throughout the rest of the project.

First, each constituency is assigned one starting ward, chosen similarly to the central ward in the Voronoi selection method, by either a greedy selection or roulette wheel. This is done solely to ensure that every constituency is guaranteed to have at least one ward, a constraint not applied to the original hybrid selection method. This then works iteratively by looking at the neighbours of each ward, checking whether or not they have been assigned a constituency, and if they have, the probability that the ward will be assigned to that same constituency is stored in a list. The list is then sorted from the highest to the lowest probability constituency, at which point the ward can iterate through the

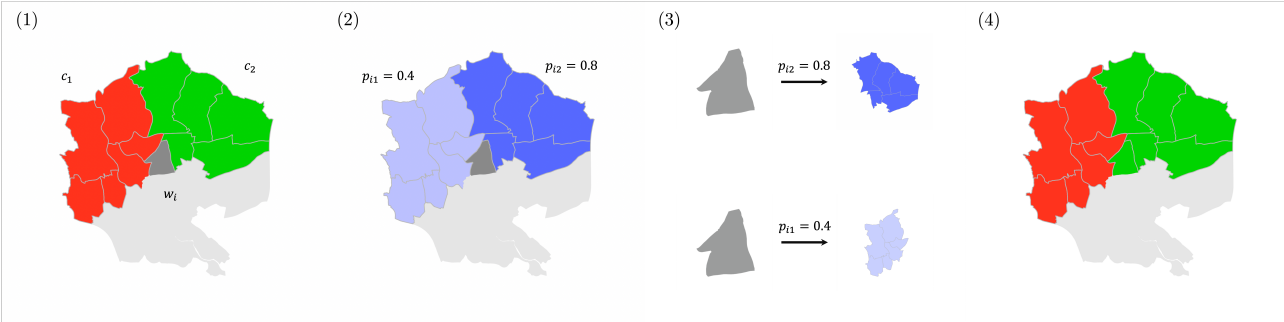


Figure 5: A diagram demonstrating ward assignment using the acceptance selection method. (1) For each ward, find the constituencies of neighbouring assigned wards. (2) Find the probability of each neighbouring constituency. (3) Iterate through the possible constituencies, and accept the assignment using the value from the probability matrix. (4) Assign the ward according to the first accepted constituency.

list, choosing whether to "*accept*" each one based on their probability. This is then repeated until the majority of wards have been assigned, leaving only those that cannot be due to contiguity constraints, and so, these can similarly be assigned to the constituency of the closest assigned ward.

6.2 Descent-based Local Search

The next stage within the model is the descent-based local search algorithm, which is where each solution is improved by randomly swapping wards between constituencies. Each swap made can then be evaluated by looking at changes in the fitness function and chosen to be accepted or rejected, with fitness calculations outlined in section 7. This is also where many of the constraints placed on the real boundaries are implemented, specifically, the population and area constraints, alongside the contiguity constraints.

Therefore, in order to make a swap, a ward is randomly chosen out of all wards, excluding those in protected constituencies, and its neighbours are examined. If the ward borders a constituency different to its own, both the original and the new potential constituency can be analysed to ensure that the swap would not produce an invalid result. This means that the areas of the two new constituencies are checked to ensure they remain below 13,000km^2, and the populations checked to ensure they do not go outside of the allowed 5% deviation from the electoral quota (unless they are above 12,000km^2 large). An adjacency matrix is then generated from all of the wards in the original constituency, which can then easily be used to check whether the constituency would remain contiguous without the swapped ward, ignoring any island wards without adjacencies.

Once a ward is swapped, this swap will only be accepted if there has been a positive improvement in the fitness of the solution, which is the same as the hill climbing method seen in the original algorithm proposed by Zhou et al. Initially, this had been changed to use simulated annealing, which is a probabilistic method first proposed by Kirkpatrick, Gelatt and Vecchi and later, Černý [37, 38], where worse solutions can sometimes be accepted according to a probability given by the Boltzmann distribution. However, this was changed to use a hill climb approach in order to help the algorithm converge on a better solution, particularly when overall fewer swaps were made within the local search. A check was also added that specifically ensured that there is an improvement in the fairness of any affected country, in order to try and solve some of the limitations of the fairness metrics as outlined in section 8. Alongside these checks, in order to better try and represent local communities and avoid constituencies becoming too spread out, wards will not swap into different regions, which are defined using the same 9 regions as were used in the linear regression model during the dataset creation, although this could be expanded into further sub-regions or counties.

6.3 Reinforcement Learning

Once the solution has been improved, the new solution is compared to the initial solution, making note of any changes and rewarding or penalising constituencies accordingly. This section of the algorithm remains almost identical to the probability updating and smoothing stages seen in the original method. Therefore, at each iteration t, the constituency assignment for a given ward, w_i, is compared in the original solution S_t and the improved solution $\hat{S}_t$. If it has stayed in its original constituency, c_u, we reward c_u and update the probability vector p_i according to the reward factor α $(0 < \alpha < 1)$, using equation (1).

$$p_{ij}(t+1) = \begin{cases} \alpha + (1-\alpha)p_{ij}(t) & j = u \\ (1-\alpha)p_{ij}(t) & \text{otherwise.} \end{cases} \tag{1}$$

Otherwise, when ward w_i has moved from its original constituency c_u to a new constituency c_v, where $v \neq u$, we penalise the original constituency c_u and compensate the new constituency c_v according to the penalisation and compensation factors β $(0 < \beta < 1)$ and γ $(0 < \gamma < 1)$ respectively. This is done according to equation (2).

$$p_{ij}(t+1) = \begin{cases} (1-\gamma)(1-\beta)p_{ij}(t) & j = v \\ \gamma + (1-\gamma)\frac{\beta}{k-1} + (1-\gamma)(1-\beta)p_{ij}(t) & j = v \\ (1-\gamma)\frac{\beta}{k-1} + (1-\gamma)(1-\beta)p_{ij}(t) & \text{otherwise.} \end{cases} \tag{2}$$

Lastly, values in the probability matrix are reduced by multiplying them by a coefficient ρ if they are above a defined threshold p_0. This is done so that older and potentially misleading decisions can be forgotten.

6.4 Multi-stage Optimisation

As can be seen in Figure 4, the last significant change to the original method has been to use multi-stage optimisation, specifically, fairness prioritised and compactness prioritised optimisation stages are both used. This means that, for the majority of the algorithm, the solutions are optimised solely for fairness during the DBLS and how well they are able to represent the electorate, but near the end of the execution, the algorithm changes to solely focus on compactness.

Although the main goal of the model is to optimise for fairness, which it is still able to do, implementing a compactness stage allowed for the algorithm to produce much more reasonable results. This ultimately follows the previously mentioned concept by Kaufman, King and Komisarchik, who note that despite many US state constitutions stating legislative boundaries should be "compact," there usually is no formal definition, and they instead rely on the idea people should recognise gerrymandering when they see it [20]. Using two stages of optimisation is therefore a good compromise that allows for the algorithm to produce fairer constituency boundaries that still appear as reasonable and logical solutions.

7 Fitness Metrics

The design of a suitable fitness metric can be incredibly important in the design of any optimisation algorithm and can have a large impact on its performance and the solutions it produces. Therefore, a not insignificant proportion of the project was spent on testing and implementing a decent fitness metric for redistricting. Specifically, although measuring the compactness of a constituency is fairly trivial, more research had to be done into the development of a qualitative measure for fairness that can be applied to parliamentary systems, which to the author's knowledge, has not been done before. Therefore, the seats-votes (SV) difference was formulated for use in this project, as outlined in section 7.2.

This results in the fitness metric $f(S)$ for a solution S as seen in equation (3), where the average party SV difference is denoted by v and the average constituency compactness denoted by c. The algorithm's priority regarding each of these factors can also be tuned using the parameters λ and μ

respectively. Multi-objective optimisation techniques could be applied to this problem, however, using the two-stage optimisation makes this unnecessary as each factor is optimised in isolation. On the other hand, if more optimisation factors were to be added, this may no longer be the case.

$$f(S) = \frac{\lambda(1 - v) + \mu \cdot c}{\lambda + \mu} \tag{3}$$

7.1 Compactness Measures

As found in a study comparing compactness measures by Barnes and Solomon [21], the choice of which compactness measure to use is fairly subjective, however, it was decided that a version of the Reock score [18] would be used for the model. This calculates compactness as the ratio of the constituency's area to the area of the smallest circle that would completely enclose it, i.e the bounding circle. However, a slight modification to the calculation was made for this model due to the complexity and computational cost of repeatedly finding the bounding circle of a polygon, a challenge called the "smallest-circle problem" [39].

Therefore, in this case, the compactness is measured as the ratio of the constituency's area to that of a circle with a diameter as the length or width of the constituency, whichever is the largest. If this value is above 1, which cannot be the case when using the real bounding circle, the compactness is taken as the reciprocal of this value. The average compactness for constituencies in the dataset used by this model is 0.49996, with the most compact constituency being "Bristol South" (0.8969) and the least compact being "Orkney and Shetland" (0.0588), both of which can be seen in Figure 6.

Figure 6: Left: The most compact constituency, "Bristol South" (0.8969). Right: The least compact constituency, "Orkney and Shetland" (0.0588).

7.2 Fairness Measures

7.2.1 Fairness in Parliamentary Systems

As previously mentioned, a majority of the literature surrounding research into redistricting and gerrymandering is done using the US as a case study, even though many other countries use similar FPTP electoral systems, such as the UK. This could be because many state congressional districts in the US are drawn by the governing party [40] instead of by independent commissions as is done in the UK, meaning that partisan bias may be more of an issue. For example, Canada used to have similar problems of partisan bias in electoral boundaries, until they started using independent commissions in the 1960s [41]. However, despite commissions being established in some US states, gerrymandering still remains an issue, such as in California where Democrats were still able to influence the commission to favour their party during the 2010 redistricting cycle [42]. Therefore, when attempting to measure the fairness of parliamentary maps, some extra considerations have to be taken, specifically, any metric must be able to handle any number of parties and potentially many more seats. Scaling a metric for a much larger number of seats is not too difficult, however, the computational cost may become a factor depending on the metric used, although this was not an issue within this project.

On the other hand, measuring fairness when only using 2 parties is comparatively much easier than systems with many parties, such as the 13 used within the redistricter model, as parties can be directly compared to one another. For example, the efficiency gap is simply the ratio between two parties' wasted votes, which cannot scale to any number of parties. As well as this, in parliamentary systems, it may be common to have a few larger parties alongside several smaller parties, as seen in

Figure 7, who may have a significant, but ultimately too widespread, level of support, making it more difficult for them win seats in government. Therefore, in order to best represent the electorate, it would be desirable to design a fitness metric in a way that can give equal respect to large and small parties. Of course, the UK Parliament complicates this further by having national parties for countries within the UK, such as Scotland's SNP or Wales' Plaid Cymru, who must also be represented fairly within Parliament. Some other measures, such as the median-mean difference [43] which says that large differences in the mean and median district party vote shares denote potential bias, could handle parliamentary systems much better as the number of parties is not specified, but this does not use comparisons to proportion vote in its calculations.

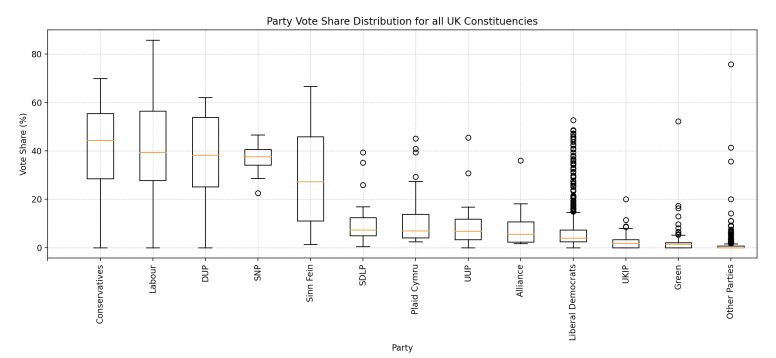


Figure 7: A graph showing the distributions of each party's vote share in all constituencies where they ran a candidate.

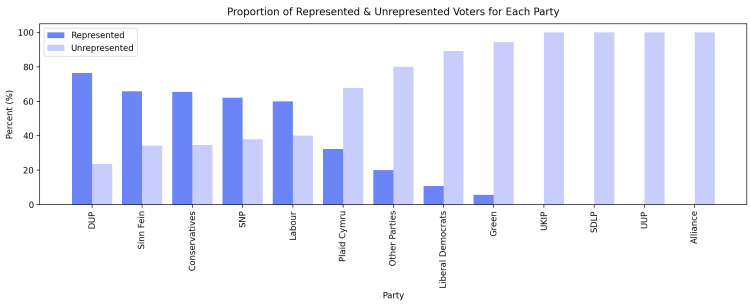


Figure 8: A graph showing the proportion of UK voters who are "represented" in Parliament, meaning they are in a constituency where their party won, against those who are "unrepresented," meaning they voted for a party that did not win their constituency.

7.2.2 The Efficiency Gap & the Wasted-vote Score

A few different metrics were tested before it was decided that the average SV difference would be used to measure fairness, mainly, this involved applications of the concept of a "wasted vote," as defined in the efficiency gap. However, this was found to not apply well, due to the efficiency gap's assumption that a party wins a seat if they gain 50% of the vote. As mentioned in section 2.1, this is not always the case in parliamentary systems, which means that whilst votes wasted by being cast towards a losing candidate can easily be denoted "wasted," calculating surplus votes can be more challenging.

Keeping this threshold at 50% was found to disadvantage smaller parties during testing, as this assumes that all parties are equally as likely to win. Similarly, another assumption that using $\frac{1}{n}$, where n is the number of parties, resulting in $\frac{1}{13}$ for the redistricter model, also does not hold due to the uneven distribution of votes between parties as seen in Figure 7. Using $\frac{1}{13}$ as the threshold for a surplus vote would massively over-calculate the wasted votes for small parties, for example, looking at Brighton Pavilion, the only seat won by the Green party who received 57.2% of the vote there, all votes over 8% would be denoted as surplus. Later testing adapted this into the "wasted-vote score,"

which used a dynamic surplus threshold, taken as the party's overall proportional vote. However, this did not produce suitable results, as using the same example, the threshold would then be at 1.6%, and the model would be encouraged to essentially "*crack*" the votes until constituencies only have 1.6% vote share for the greens.

Another method developed under a similar concept was to examine the proportions of represented and unrepresented voters, where represented voters are those who voted for the candidate who won their constituency and unrepresented voters who voted for a candidate who lost their constituency. The proportions of each of these classes of voters for each party can be seen in Figure 8. This was ultimately not fully completed, but showed promising results and could be researched further to create a new measure that combines the proportions of both classes and n parties, potentially in the same way that the seat and vote shares are combined as in the final method.

7.2.3 Seats-Votes Difference

The seats-votes difference was the final method developed for use in the redistricting model and was used over the others mentioned due to how it can directly compare to the proportional vote. The efficiency gap itself has been criticised as a metric not only for being fairly unstable in competitive races but also for the way it can penalise proportionality as noted by Bernstein and Duchin, who calculated parties with a 60:40 split in the proportional vote and seat share would be found to be 8% biased towards the losing party [44]. So, as this project has held the aim that a fair map is one that can best represent the electorate, a "Seats-Votes difference" measure for fairness is therefore proposed that simply calculates the average difference between a party's overall vote share and its seat share in government, divided by their vote share.

This has the advantage that it scales well for any number of parties, is easy to calculate and directly rewards proportionality, where the seat share of a party is the same as its proportional vote. A formalised definition of the measure can be seen in equation 4, where n is the number of parties, p is the set of proportional votes for each party and s is the set of seat shares for each party, both of length n. Appendix B shows a full breakdown of party SV differences, although the UK was found to have a national average SV difference of 64.99%, meaning a fairness score of 0.3501 out of 1.

$$v = \frac{1}{n}\left(\sum_{i=0}^{n} \frac{|p_i - s_i|}{p_i}\right) \tag{4}$$

8 Evaluation & Results

8.1 Overview

As mentioned in section 2.2, the model was to be evaluated on how well it is able to improve the fairness of the current UK boundaries, with some evaluation also done on the proposed 2023 boundaries. The experiments performed are therefore outlined as follows. First, several parameter tuning experiments have been carried out using the current boundaries, specifically focusing on the reinforcement learning and descent-based local search parameters. Performance on other factors, including the algorithm runtime and ability to capture communities of interest, are evaluated and some potential limitations are given. Then, further evaluation of the 2023 boundaries and the algorithm's performance when generating completely new boundaries are outlined. Lastly, a discussion surrounding the SV difference as a fairness measure is given, mainly through a comparison to the efficiency gap.

8.2 Model Evaluation

8.2.1 Current Boundaries

Overall, the model evaluation using current boundaries was performed over 106 runs, covering 24 experiments, and showed that the model was able to produce an improvement in the solution's fairness, starting at 0.3501, in all runs. For example, as seen in table 1, over the 5 runs in the control experiments, each saw an average improvement in fairness of 1.5% after 45 stage 1 and 5 stage 2

iterations. This was using reinforcement learning parameters of $\alpha = 0.8,\ \beta = 0.6,\ \gamma = 0.8$, which were used as the base parameters during all other experiments. It can also be noted that although all runs showed an improvement in fairness, only some produced an increase in constituency compactness, except for when explicitly prioritising compactness as seen in table 2. An example of the fitness metrics throughout one of the control experiments can be seen in Figure 9.

Fitness Metric	***Run 1***	***Run 2***	***Run 3***	***Run 4***	***Run 5***	***Average***	***Change***
SV Difference	0.363729	0.36808	0.360212	0.369794	0.364258	0.3652146	+0.0151046
Compactness	0.498137	0.493512	0.494746	0.495945	0.495827	0.4956334	-0.0043286

Table 1: Final solution fairness and compactness for the control experiments, with 45 stage 1 iterations and 5 stage 2 iterations. Here, $\alpha = 0.8,\ \beta = 0.6,\ \gamma = 0.8$ and 100 steps were taken during the DBLS.

Optimisation Priority	***Metric***	***Average***	***Change***
Fairness ($\lambda = 1.0$)	SV Difference	0.36352	+0.01341
	Compactness	0.493032	-0.00693
Compactness ($\mu = 1.0$)	SV Difference	0.3623824	+0.0122724
	Compactness	0.5029076	+0.0029456

Table 2: Final solution SV difference and compactness score after 50 iterations when prioritising solely for fairness or compactness, using $\lambda = 1.0,\ \mu = 0.0$ and $\lambda = 0.0,\ \mu = 1.0$ respectively. All other parameters were the same as in the control experiments.

The fairest set of boundaries produced by the model can be seen to the right of figure 10, which shows the model's boundaries in comparison to the current real-world boundaries. This run showed an improvement of 7% from the current boundaries, to produce a fairness score of 0.4201 and was achieved using reinforcement learning parameters of $\alpha = 0.9,\ \beta = 0.1,\ \gamma = 0.9$. Notable seat differences in the proposed map include; the Ulster Unionist Party (UUP) gaining a seat in Northern Ireland, Plaid Cymru gaining a seat and Labour gaining 2, whilst the DUP and SNP each lost 1 and 2 seats respectively. For example, Labour, who have 262 seats using the current boundaries and 264 in the optimised map, had their SV difference decrease from 0.6% to 0.3%, whilst the SNP's SV difference decreased from 77.2% to 68.7% by losing 2 seats. Then, taking the UUP as an example, there is a high concentration of supporters in the southwest of Northern Ireland even though they do not have any seats in Parliament, however, giving them a seat reduced their SV difference from 100% to 46%. The high concentration of voters meant the model was able to give the UUP a seat fairly often, which could signify some level of partisan bias in the current Northern Ireland constituency map.

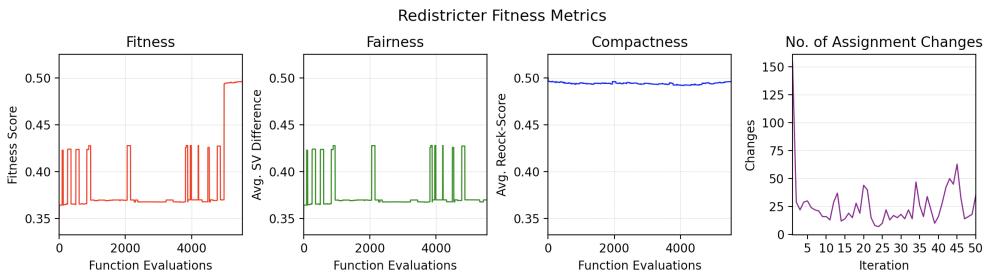


Figure 9: Fitness metrics over time during a control experiment (run 4), showing the combined fitness score, the average SV difference, the average Reock Score and the number of changes in ward assignment during the selection method for each iteration.

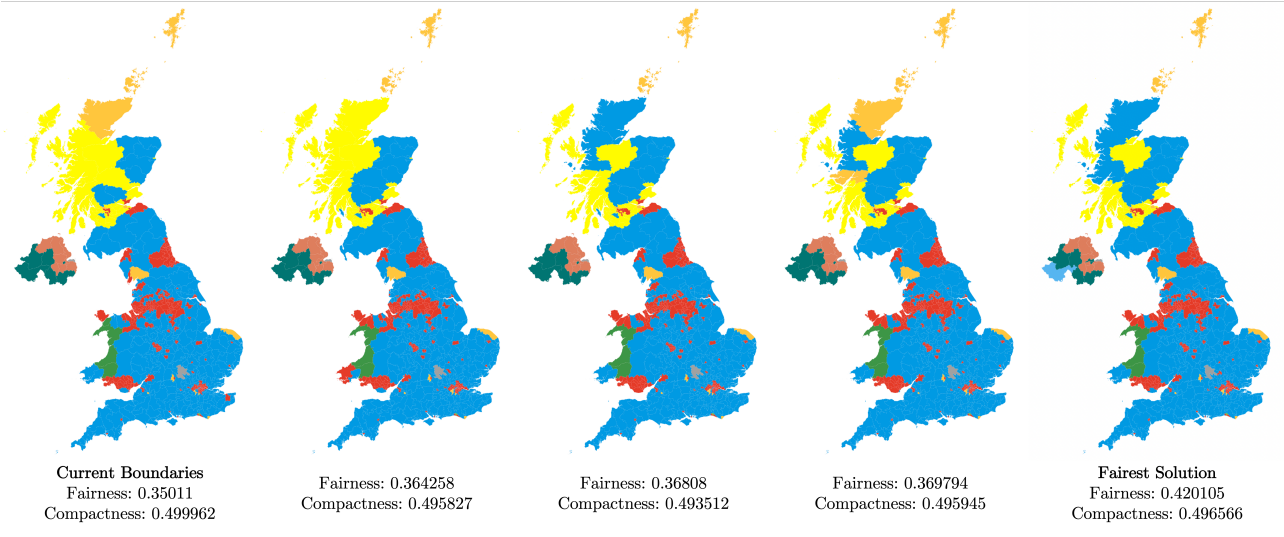


Figure 10: A comparison of the current constituency boundaries and four maps generated during the control experiments, along with their respective fairness and compactness scores. The furthest left was the fairest generated in all experiments, with $\alpha = 0.9,\ \beta = 0.1,\ \gamma = 0.9$.

Significant experimentation was also carried out on the reinforcement learning parameters, with the averaged results over 5 runs shown in appendix C. However, although some of the better results were produced with higher reward and compensation factors, and lower penalisation factors, there was no clear trend in the model's performance. Similarly, increasing the iterations and number of steps made during the DBLS seemed to produce no clear pattern in the fairness of the maps generated by the model. There are likely two main reasons that this is the case; a limitation of the SV difference in its application to redistricting and the nature of the current boundaries themselves.

Firstly, as the SV difference solely uses comparisons of seat share to the proportional vote, it is not able to recognise a favourable swap unless there is a change in seats, which of course does not always happen. This means that during the DBLS, the model may struggle to optimise for seat share when not using a metric such as the efficiency gap or the proportion of represented to unrepresented voters, as these would change with each swap. This, alongside the selection method, explains the dramatic increases and decreases in the fairness score that can be seen in Figure 9, which would be due to the large impact that the selection method has on the fitness of a solution. This means that as a result, whilst the SV difference may still perform well as a method for evaluating the fairness of an election, it is less able to detect small changes in a constituency map.

The other potential limitation of the algorithm's performance is the current constituency boundaries, because, as would be the case with any map, there is likely a limit to what can be done by swapping individual wards. In order to test this, the algorithm was run as a gradient descent algorithm, which can be done by only running the model for 1 iteration and dramatically increasing the steps made during the DBLS, in this case to 100,000. The results, which can be seen in appendix D, show the limitation of this approach as the fitness was never able to increase above 0.428 in any attempt. This likely means that, despite the fact that the reinforcement learning approach was able to achieve competitive results during the execution of some of the experiments, there must be some inherent limit with the use of the current boundaries as the initial solution.

Despite these somewhat marginal changes, it is worth reiterating that the algorithm was still able to improve the fairness of the current constituency map in every experiment. Even a few seats changing hands to parties that would otherwise be significantly under-represented in Parliament would demonstrate a positive shift in how the electorate is being represented. As well as this, a solution to the high level of randomness and fitness peaks may be to simply terminate the algorithm early if the fitness has reached a given threshold. Otherwise, it also may be the case that smaller changes in the map are desirable as current boundaries would likely reflect long-standing cultural or historical ties which could evidently be upset by drastically changing the boundary of every constituency.

Alongside the main goal of optimising boundaries for fairness, there are some other important factors to consider, these mainly include execution time and the algorithm's ability to preserve communities of interest.

Typically, the redistricting algorithm took around a minute per iteration, although this is evidently dependent on many factors, especially the hardware used. The largest factors in this are the mapping of solutions, the selection method and the number of steps in the DBLS, although, the impact of most of these can be reduced if required. Firstly, the implementation of the algorithm has the ability to map solutions at the beginning of each iteration and can compile them into a video once the algorithm has terminated. This was found to be incredibly useful for tracking the performance over time, although, this can, of course, be skipped if runtime becomes an issue, and the same effectively goes for the number of steps during the DBLS which could also be decreased. However, the selection method is slightly more limited as the algorithm needs to perform a few iterations to assign all wards whilst conserving contiguity. The number of iterations before contiguity is ignored could be lowered, however, this would not typically be advised. Ultimately, the algorithm was found to execute in a reasonable time and could easily be tuned or terminated early if the solution's fitness has reached a given threshold, which is especially the case in comparison to the traditional, non-algorithmic, methods for redistricting, where the algorithm would represent a significant improvement.

The other factor, the preservation of communities of interest, is much harder to optimise for and evaluate, and therefore represents a significant challenge for algorithmic redistricting. For example, many papers within the literature explicitly mention the limitation that their method cannot consider communities or note how this should be implemented in the future, such as the optimal compactness method [5] or the antimander project [23]. On the other hand, the redistricter proposed by this paper aims to preserve communities of interest through two methods; the use of electoral wards as the geographic base unit, and limiting swaps in the DBLS. Firstly, as mentioned in section 4, electoral wards are typically used by the UK boundary commissions, not only because they are well-defined and well-understood, but also because they somewhat follow communities of interest, such as local government boundaries or cultural ties within an area. Therefore, the use of electoral wards means that boundaries generated by the redistricter should also follow this idea. Secondly, as mentioned in section 6.2, ward swaps are limited within each region within the country so that constituencies do not spread too far, which should also help towards the preservation of communities. Although, this could very easily be extended in the future by using further subdivisions, such as at the county level.

8.2.2 2023 Boundaries

Alongside experiments on the current boundaries, the model was also tested on those from the initial boundary proposal for the 2023 boundary review, released in June 2021. Using the same metrics as above, the proposed map was found to be 6% less fair than the current map, with a fairness score of only 0.2986, as opposed to 0.3501. The compactness score decreased as well, but only marginally, from 0.49996 to 0.4916. The model also predicted that the Conservatives would gain 6 seats from the changes, whilst Labour, the Liberal Democrats, the DUP, and Plaid Cymru would all lose seats, notably with Plaid Cymru losing 2 of their 3 current seats. It is also worth noting that the Conservatives would supposedly have gained a total of 327 seats, meaning they would have above the 325 needed to hold a majority, something they were not able to achieve in the real 2017 election. This is ultimately in agreement with many of the original criticisms of the proposals, including how the map would benefit the Conservative party and neglect Scotland and Wales, although the model's predictions were not as bad as others. For example, Electoral Calculus' model predicted that the Conservatives would gain 13 seats and Labour would lose 8 [45], however, this would of course use more recent elections data to make predictions. A detailed outline of the model's predicted seat changes is in appendix E.

The boundaries themselves can be seen in Figure 11, alongside the optimised maps generated by the redistricting algorithm. This was somewhat able to improve the fairness score, with the best improvement only being to 0.3274 which is still below the current boundaries. However, the Conservatives would have one less seat and Labour 4 more than compared to the current boundaries. However, this potentially highlights how the initial boundaries affect the model's performance as the model was not able to have Plaid Cymru regain their seats.

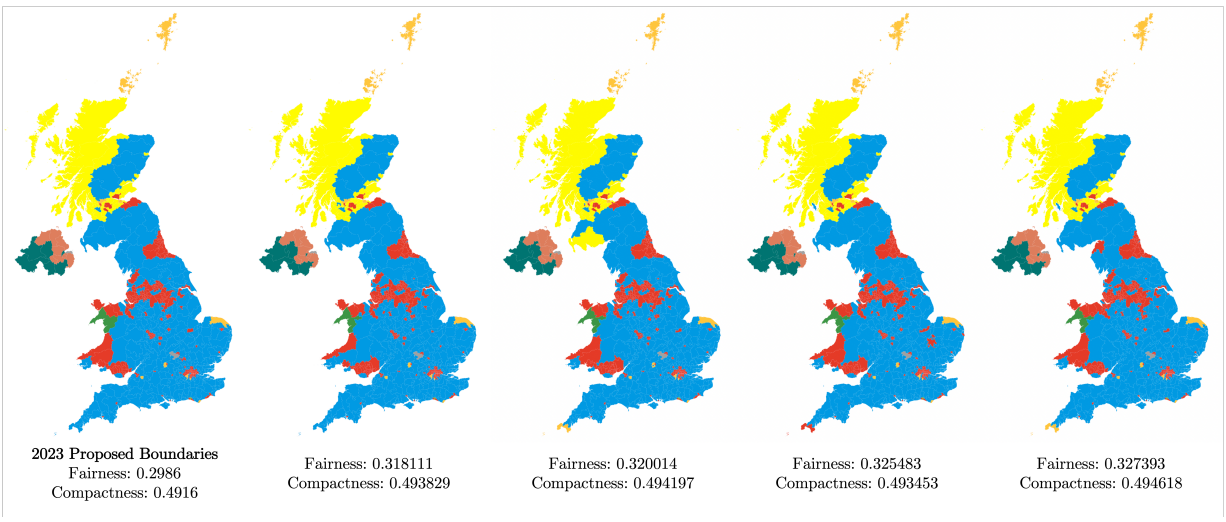


Figure 11: A comparison of the proposed 2023 boundaries and four fairness optimised versions, along with their respective fairness and compactness scores. These were generated using the same parameters from the control experiments, $\alpha = 0.8$, $\beta = 0.6$, $\gamma = 0.8$.

8.2.3 Generation of New Boundaries

Some limited experimentation was carried out to evaluate how the model performed when generating boundaries completely from scratch, which is very easy to achieve by initialising all ward-constituency probability values in the probability matrix as $1/n$, where n is the number of constituencies in the ward's country. The results can be seen in Figure 12, and show how well the model is able to converge onto a reasonable solution within only a few iterations. Although, as mentioned previously, smaller changes may be desirable when reviewing constituency boundaries, and they are only updated and not redrawn every 5 years, the fact that the model is able to converge is promising for future applications of the algorithm. Despite the fact the fitness values achieved were not as high as the current boundaries, this means further experimentation should be carried out to evaluate the model's performance over more iterations and using multiple runs.

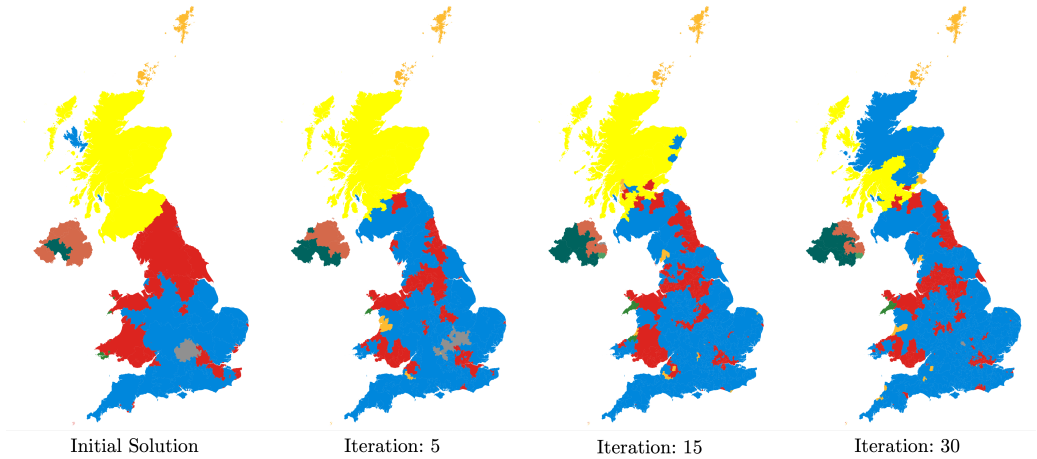


Figure 12: Boundaries generated from scratch over 30 iterations, demonstrating the algorithm's convergence on a reasonable solution.

8.3 SV Difference Evaluation

As mentioned before, the goal of the SV difference is to directly compare the difference between a party's seat share and vote share, meaning that it is able to evaluate the fairness of an election by

directly comparing the proportional vote and rewarding proportionality as a result. Again taking the example by Bernstein and Duchin, an election between two parties with a 60:40 split in the vote share and seat share would have the efficiency gap finding the election 8% biased towards the losing party, whilst the SV difference would show the election was completely fair. Of course, some of the SV difference's advantages have already been discussed, such as its simplicity and ability to apply to a parliamentary system, as well as its limitations, mainly its handling of small changes in constituency boundaries, however, the metric has been further evaluated through a comparison to the efficiency gap.

In order to compare the two, the average SV difference was calculated for all 50 US states for results from the 2016 congressional elections, with Figure 13 comparing the results to a dataset of efficiency gap values for the same elections [46]. Generally, the average SV difference tended to be larger than the advantage measured by the efficiency gap, with one of the more notable examples being Idaho, which had an efficiency gap advantage of 13.7% towards the Republicans, but a SV difference of 73.5%. In this case, the state only had the 20th highest efficiency gap bias even though the Republicans won both of the 2 seats despite only having 68% of the popular vote. Using the SV difference, Idaho would be the 11th most biased state, which should be more logical as 32% of the electorate is not represented by either congress member.

Overall, if the belief that a fair FPTP election is one where the seat share closely mirrors the vote share, then the average SV difference performs well as it is able to highlight elections where a large percent of the electorate is not being accurately represented. However, it may be true that results may evidently cannot be as nuanced when there is only one seat in the election, which would be because states with more seats are able to more easily have closer seat and vote shares, for example, in comparison to states with only one district. In the case of a single-district state where all of one party's voters inherently cannot be represented, this could be interpreted in two ways; one stance would say that this is a fair election as the one party with the majority of votes won the seat, whilst the other would say the SV difference highlights that a large percent of the electorate cannot be represented when only using one seat.

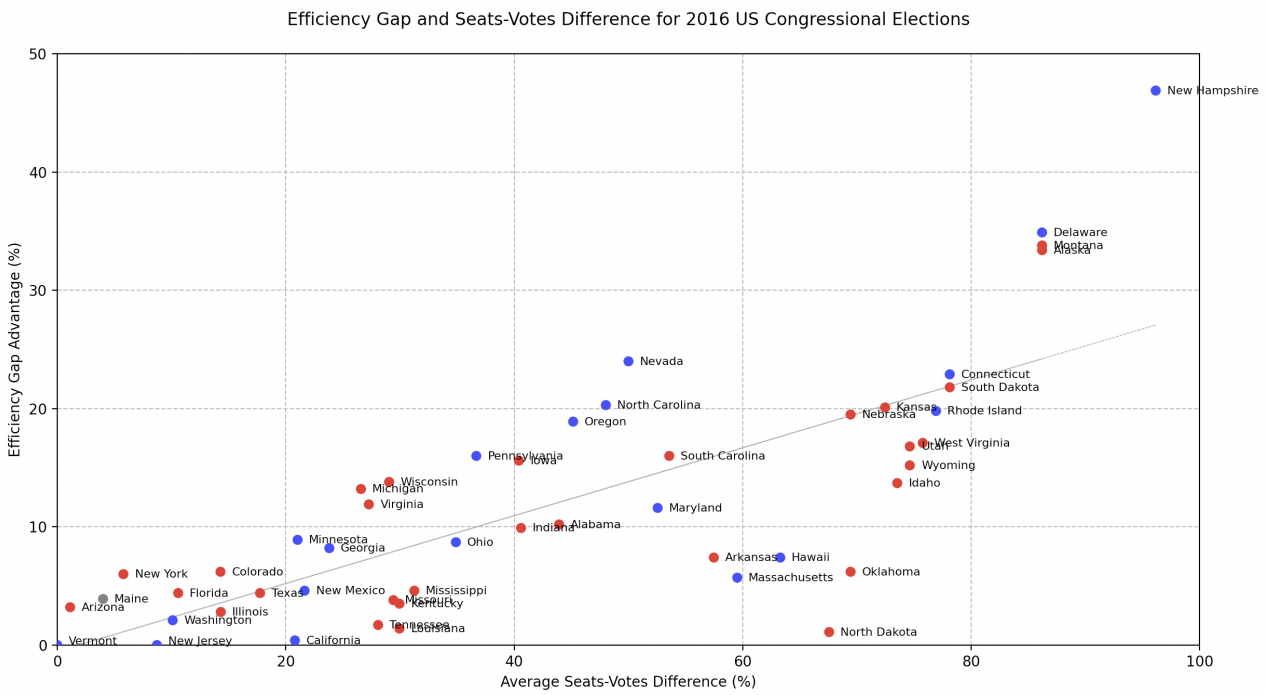


Figure 13: A comparison of the efficiency gap and SV difference for states in the 2016 US congressional election, with the efficiency gap values calculated by McGlone and Needham [46]. States are colour coded by the party that won the most seats in the state.

9 Conclusions

9.1 Future Work

There are many ways that the redistricter algorithm could be expanded, although, in the author's opinion, the main areas of potential future research or experimentation would include:

1. ***2021 Census Data:*** Currently, the model uses the most recent UK census data from 2011 and election results from the 2017 general election, however, this obviously could be updated to use more recent datasets. Due to the timing of the project, results from the 2021 UK census could not be used as the final release of data by the ONS is not set to take place until March 2023. Therefore, after this point, the model could be evaluated again along with results from the more recent 2019 general election or any subsequent elections.
2. ***Fairness Metric:*** It is clear from this project that further research should be undertaken to develop a fairness metric for parliamentary systems that would be more suitable for a redistricting algorithm such as this. This could be done through an expansion and subsequent evaluation of either the wasted vote score or the ratio of represented to unrepresented voters.
3. ***Improvement Algorithm:*** As was noted by Zhou, Hao and Duval, the authors of the original paper, a more powerful algorithm could be used in place of the DBLS. For example, a population-based approach could be taken in order to better optimise for fairness, potentially mixing the optimal versions of different regions to create the improved solution. This could also make the algorithm more easily parallelised, helping to improve the efficiency and runtime, as the current method could not be easily parallelised, except when multiple entirely separate runs were performed simultaneously during experimentation.
4. ***Socio-demographic Factors:*** Other factors could be added into the improvement algorithm and fitness function, such as some kind of demographic measure that aims to ensure that communities of interest are preserved and highlighted in proposed maps. The demographic data is already needed as part of the linear regression model and the creation of the ward-level election results, but it is not currently used within the actual redistricting algorithm. However, as more factors are added to the fitness function, multi-objective optimisation techniques may be required in order to ensure that maps are optimised correctly as the current multi-stage optimisation may no longer be sufficient.

9.2 Final Conclusions

It should be clear from the variety of redistricting algorithms and fairness measures, along with the large number of factors to be considered and optimised, that the problem of drawing or updating constituency or electoral district boundaries is a very complex one with no objective solution. Whilst purely quantitative methods are able to accurately distribute a population amongst a series of compact districts, these methods often fail to produce results that can best represent the electorate as they do not consider the other factors that influence the fairness of a map. On the other hand, those that directly take election results and fairness into consideration should be able to ensure that accurate representation remains a primary goal of a strong democracy.

Ultimately, ensuring this was possible was one of the main intentions behind this project, and I am very confident that these results fulfil all of these ambitions. Even though changes were fairly marginal, the algorithm was consistently able to improve the fairness of current and proposed maps and, with some of the future expansions as outlined above, could be built upon and improved even further. Then similarly, The SV difference was found to perform well towards aspects of the goal, mainly in rewarding proportionality between the seat share and vote share, and should be excellent in its application as a fairness metric for elections in parliamentary systems.

Overall, whilst the political landscape in the UK is constantly changing, looking to the future, I believe that aiming for proportionality and ensuring the electorate is fairly and accurately represented by their government could only be beneficial. Therefore, through the implementation and application of this model, as well as the fairness metric, this project should hopefully demonstrate that fairness can and should be considered a higher priority within our electoral systems.

References

[1] F.H. 'How Britain draws its electoral boundaries'. In: *The Economist* (Sept. 2018). URL: https://www.economist.com/the-economist-explains/2018/09/20/how-britain-draws-its-electoral-boundaries (visited on 01/05/2022).

[2] Andrew Witherspoon and Sam Levine. 'These maps show how Republicans are blatantly rigging elections'. In: *The Guardian* (Nov. 2021). URL: https://www.theguardian.com/us-news/ng-interactive/2021/nov/12/gerrymander-redistricting-map-republicans-democrats-visual (visited on 01/05/2022).

[3] Neil Johnston and Elise Uberoi. *Constituency boundary reviews and the number of MPs*. Research Briefing 5929. London: House of Commons Library, Nov. 2021, pp. 46–48.

[4] Warren D. Smith. *Splitline districtings of all 50 states + DC + PR*. July 2011. URL: https://rangevoting.org/SplitLR.html (visited on 01/05/2022).

[5] Brian Olson. *BDistricting*. Nov. 2021. URL: https://bolson.org/dist/ (visited on 01/05/2022).

[6] Benjamin Fifield, Michael Higgins, Imai Kosuke and Alexander Tarr. 'Automated Redistricting Simulation Using Markov Chain Monte Carlo'. In: *Journal of Computational and Graphical Statistics* 29.4 (Oct. 2020), pp. 715–728. DOI: 10.1080/10618600.2020.1739532.

[7] Kevin Baas. *Auto-Redistricter*. Sept. 2018. URL: https://github.com/happyjack27/autoredistrict/releases (visited on 19/04/2022).

[8] Wes Gurnee and David B. Shmoys. 'Fairmandering: A column generation heuristic for fairness-optimized political districting'. In: *arViv:2103.11469* (June 2021).

[9] Andrew Reynolds, Ben Reilly and Andrew Ellis. *Electoral system design: the new international IDEA handbook*. Handbook series. Stockholm, Sweden: International Institute for Democracy and Electoral Assistance, 2005. ISBN: 978-91-85391-18-9.

[10] Erik J. Engstrom. *Partisan Gerrymandering and the Construction of American Democracy*. 4th ed. Ann Arbor: University of Michigan Press, 2016. ISBN: 978-0-472-02952-5.

[11] Christopher Hewitt. 'Catholic Grievances, Catholic Nationalism and Violence in Northern Ireland during the Civil Rights Period: A Reconsideration'. In: *The British Journal of Sociology* 32.3 (1981), pp. 362–380. DOI: 10.2307/589283.

[12] Nicholas Stephanopoulos. 'The Causes and Consequences of Gerrymandering 2020 Redistricting: Mapping a New Political Decade Symposium'. In: *William & Mary Law Review* 59.5 (2018), pp. 2115–2158.

[13] Ferran Martinez i Coma and Ignacio Lago. *Majoritarian electoral systems are more prone to gerrymandering than proportional systems*. June 2016. URL: http://eprints.lse.ac.uk/80662/ (visited on 01/05/2022).

[14] Nicholas Stephanopoulos and Eric McGhee. 'Partisan Gerrymandering and the Efficiency Gap'. In: *University of Chicago Law Review* 82.2 (Mar. 2015), pp. 831–900.

[15] Henri Theil. 'The Desired Political Entropy'. In: *American Political Science Review* 63.2 (June 1969), pp. 521–525. DOI: 10.1017/S0003055400262382.

[16] Jonathan P. Kastellec, Andrew Gelman and Jamie P. Chandler. 'Predicting and Dissecting the Seats-Votes Curve in the 2006 U.S. House Election'. In: *Political Science & Politics* 41.01 (Jan. 2008), pp. 139–145. DOI: 10.1017/S1049096508080207.

[17] Mark Monmonier. 'Hook, Line and Gerrymander'. In: *Mercator's World* 5.4 (July 2000), p. 36.

[18] Ernest C. Reock. 'A Note: Measuring Compactness as a Requirement of Legislative Apportionment'. In: *Midwest Journal of Political Science* 5.1 (Feb. 1961), p. 70. DOI: 10.2307/2109043.

[19] Daniel D. Polsby and Robert D. Popper. 'The Third Criterion: Compactness as a Procedural Safeguard against Partisan Gerrymandering'. In: *Yale Law & Policy Review* 9.2 (1991). Publisher: Yale Law & Policy Review, Inc., pp. 301–353.

[20] Aaron R. Kaufman, Gary King and Mayya Komisarchik. 'How to Measure Legislative District Compactness If You Only Know It When You See It'. In: *American Journal of Political Science* 65.3 (July 2021), pp. 533–550. DOI: 10.1111/ajps.12603.

[21] Richard Barnes and Justin Solomon. 'Gerrymandering and Compactness: Implementation Flexibility and Abuse'. In: *Political Analysis* 29.4 (Oct. 2021), pp. 448–466. DOI: 10.1017/pan.2020.36.

[22] Aaron Bycoffe, Ella Koeze, David Wasserman and Julia Wolfe. *The Atlas Of Redistricting*. Jan. 2018. URL: https://projects.fivethirtyeight.com/redistricting-maps/ (visited on 01/05/2022).

[23] Joel Simon and Joel Lehman. 'Antimander: open source detection of gerrymandering though multi-objective evolutionary algorithms'. In: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*. Cancún, Mexico: ACM, July 2020, pp. 51–52. DOI: 10.1145/3377929.3398156.

[24] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal and T. Meyarivan. 'A fast and elitist multiobjective genetic algorithm: NSGA-II'. In: *IEEE Transactions on Evolutionary Computation* 6.2 (Apr. 2002), pp. 182–197. DOI: 10.1109/4235.996017.

[25] Ying-Che Hung and Liang-Yü Chen. 'Using Intelligent Clustering to Implement Geometric Computation for Electoral Districting'. In: *ISPRS International Journal of Geo-Information* 8.9 (Aug. 2019), p. 369. DOI: 10.3390/ijgi8090369.

[26] Yangming Zhou, Jin-Kao Hao and Béatrice Duval. 'Reinforcement learning based local search for grouping problems: A case study on graph coloring'. In: *Expert Systems with Applications* 64 (Dec. 2016), pp. 412–422. DOI: 10.1016/j.eswa.2016.07.047.

[27] Encyclopedia Britannica. *Parliament*. May 2017. URL: https://www.britannica.com/topic/Parliament (visited on 01/05/2022).

[28] Donald D. Searing. 'The Role of the Good Constituency Member and the Practice of Representation in Great Britain'. In: *The Journal of Politics* 47.2 (June 1985), pp. 348–381. DOI: 10.2307/2130887.

[29] *Parliamentary Constituencies Act 1986*. c. 56, Parliament of the United Kingdom, Publisher: Queen's Printer of Acts of Parliament. Nov. 1986.

[30] Boundary Commission for England. *Guide to the 2023 Review of Parliamentary constituencies*. May 2021. URL: https://boundarycommissionforengland.independent.gov.uk/2023-review/guide-to-the-2023-review-of-parliamentary-constituencies/page/3/ (visited on 01/05/2022).

[31] Electoral Calculus. *Regression Analysis*. URL: https://www.electoralcalculus.co.uk/services_polling.html (visited on 14/04/2022).

[32] Office for National Statistics. *Nomis - Official Labour Market Statistics*. URL: https://www.nomisweb.co.uk/ (visited on 14/04/2022).

[33] Scotland's Census. *Scotland's Census*. en. URL: https://www.scotlandscensus.gov.uk/ (visited on 14/04/2022).

[34] Northern Ireland Statistics and Research Agency. *Northern Ireland Statistics and Research Agency (NISRA)*. en. URL: https://www.nisra.gov.uk/statistics/census (visited on 14/04/2022).

[35] The British Election Study. *Explaining Voter Volatility: A summary of the British Election Study*. Tech. rep. The British Election Study, Oct. 2019. URL: https://www.britishelectionstudy.com/wp-content/uploads/2019/10/Explaining-Voter-Volatility.pdf (visited on 01/05/2022).

[36] Richard M. Karp. 'Reducibility among Combinatorial Problems'. In: *Complexity of Computer Computations*. Ed. by Raymond E. Miller, James W. Thatcher and Jean D. Bohlinger. Boston, MA: Springer US, 1972, pp. 85–103. DOI: 10.1007/978-1-4684-2001-2_9.

[37] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi. 'Optimization by Simulated Annealing'. In: *Science* 220.4598 (May 1983), pp. 671–680. DOI: 10.1126/science.220.4598.671.

[38] V. Černý. 'Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm'. In: *Journal of Optimization Theory and Applications* 45.1 (Jan. 1985), pp. 41–51. DOI: 10.1007/BF00940812.

[39] James Joseph Sylvester. 'A question in the geometry of situation'. In: *Quarterly Journal of Pure and Applied Mathematics* 1.1 (1857), pp. 79–80.

[40] Nick Corasaniti, Reid J. Epstein, Taylor Johnston, Rebecca Lieberman and Eden Weingart. 'How Maps Reshape American Politics'. In: *The New York Times* (Nov. 2021). URL: https://www.nytimes.com/interactive/2021/11/07/us/politics/redistricting-maps-explained.html (visited on 01/05/2022).

[41] C. P. Hoffman. *The Gerrymander and the Commission: Drawing Electoral Districts in the United States and Canada*. SSRN Scholarly Paper ID 2272836. Rochester, NY: Social Science Research Network, 2006.

[42] Olga Pierce and Jeff Larson. *How Democrats Fooled California's Redistricting Commission*. Dec. 2011. URL: https://www.propublica.org/article/how-democrats-fooled-californias-redistricting-commission?token=WXm9z79e0T7-hUjzwsfdbQxryFu6WxDL (visited on 01/05/2022).

[43] Michael D. McDonald and Robin E. Best. 'Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases'. In: *Election Law Journal: Rules, Politics, and Policy* 14.4 (Dec. 2015), pp. 312–330. DOI: 10.1089/elj.2015.0358.

[44] Mira Bernstein and Moon Duchin. 'A formula goes to court: Partisan gerrymandering and the efficiency gap'. In: *American Mathematical Society* 64.9 (2017), pp. 1020–1024. DOI: http://dx.doi.org/10.1090/noti1573.

[45] Electoral Calculus. *New Constituency Boundaries for 2023*. Apr. 2020. URL: https://www.electoralcalculus.co.uk/boundaries2023.html (visited on 01/05/2022).

[46] Daniel McGlone and Esther Needham. *The Most Gerrymandered States Ranked by Efficiency Gap and Seat Advantage*. en-US. July 2017. URL: https://www.azavea.com/blog/2017/07/19/gerrymandered-states-ranked-efficiency-gap-seat-advantage/ (visited on 20/04/2022).

Appendices

Party	***2017 Seats***	***Model Seats***	***2017 Vote (%)***	***Model Vote (%)***
Conservatives	317	321	42.4	41.91
Labour	262	260	40.0	40.45
SNP	35	33	3.0	3.03
Liberal Democrats	12	13	7.4	7.32
DUP	10	10	0.9	0.92
Sinn Fein	7	7	0.7	0.79
Plaid Cymru	4	3	0.5	0.50
Green	1	1	1.6	1.60
UKIP	0	0	1.8	1.77
SDLP	0	0	0.3	0.31
UUP	0	0	0.3	0.28
Alliance	0	0	0.2	0.23
Other Parties	2	2	0.9	0.89

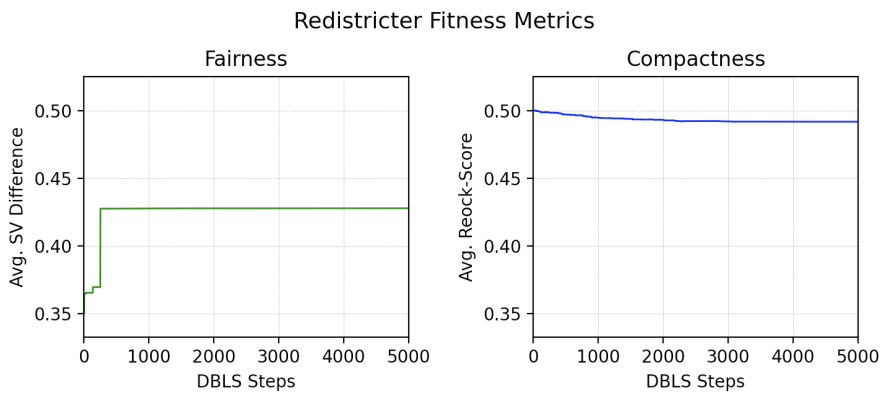
Appendix A: The number of parliamentary seats won by each party and their vote share in the 2017 general election, according to the real results and the linear regression model results.

Party	***National***	***England***	***Wales***	***Scotland***	***Northern Ireland***
Conservatives	16.873	23.986	39.850	23.036	100
Labour	0.673	0.182	42.644	56.392	-
Liberal Democrats	74.713	80.647	100	2.536	-
Green	90.552	89.949	100	100	100
UKIP	100	100	100	100	-
Plaid Cymru	19.267	-	5.430	-	-
SNP	77.183	-	-	60.493	-
DUP	60.788	-	-	-	53.433
Sinn Fein	40.776	-	-	-	34.337
SDLP	100	-	-	-	100
UUP	100	-	-	-	100
Alliance	100	-	-	-	100
Other Parties	64.035	78.641	100	100	74.159
Average	64.989%	62.234%	69.703%	63.208%	82.741%

Appendix B: The seats-votes difference for each party, taken as the absolute difference between their seat and vote shares, divided by their vote share, separated by each country.

Parameters			***Metric***	***Average***	***Change***
α (Reward)	*β (Penalisation)*	*γ (Compensation)*			
0.9	0.1	0.9	SV Difference	0.3746518	+0.0245418
			Compactness	0.4970132	-0.0029488
0.8	0.2	0.8	SV Difference	0.3641342	+0.0140242
			Compactness	0.4961882	-0.0037738
0.7	0.3	0.7	SV Difference	0.3755	+0.02539
			Compactness	0.495159	-0.004803
0.6	0.4	0.6	SV Difference	0.3662868	+0.0161768
			Compactness	0.4952904	-0.0046716
0.5	0.5	0.5	SV Difference	0.3641534	+0.0140434
			Compactness	0.496127	-0.003835
0.4	0.6	0.4	SV Difference	0.3656728	+0.0155628
			Compactness	0.4955172	-0.0044448
0.3	0.7	0.3	SV Difference	0.3648708	+0.0147608
			Compactness	0.4963978	-0.0035642
0.2	0.8	0.2	SV Difference	0.365893	+0.015783
			Compactness	0.4945108	-0.0054512

Appendix C: Final solution SV difference and compactness scores after 50 iterations for different reward, penalisation and compensation factors. All other parameters were the same as in the control experiments.



Appendix D: The fairness and compactness scores for each step during the gradient descent run, using a hill climbing algorithm. This demonstrates how the fairness is not able to go above 0.428 (an improvement of 7.2%) by using only ward swaps, at least with the current boundaries as the initial solution.

Party	***2017 Seat Share***	***2023 Seat Share***	***Difference***
Conservatives	321	327	+6
Labour	260	258	-2
SNP	33	33	0
Liberal Democrats	13	11	-2
DUP	10	9	-1
Sinn Fein	7	8	+1
Plaid Cymru	3	1	-2
Green	1	1	0
UKIP	0	0	0
SDLP	0	0	0
UUP	0	0	0
Alliance	0	0	0
Other Parties	2	2	0

Appendix E: A comparison of seats gained using the model's 2017 dataset and the model's predicted seat share using the 2023 proposed boundaries. Here, the 2017 seat shares are those generated using the linear regression model, with differences to the real 2017 seat share outlined in appendix A.