# Evaluating Redistricting of Electoral Areas

Tiago Almeida
tiagomcalmeida@tecnico.ulisboa.pt
Instituto Superior Técnico, Lisboa, Portugal

January 2021

**Abstract**

Since the 1960s several approaches have been proposed with the goal of preventing the redrawing of electoral districts in such way that they are beneficial to a certain party or political faction (gerrymandering).

In this work, a new compact Boolean formulation is proposed for solving the electoral districting problem and avoid gerrymandering. This formulation satisfies all common criteria for building electoral maps, in particular the contiguity and representation criteria. Moreover, a new compactness measure is also proposed that does not depend on geographic centers. Additionally, an incomplete formulation is also devised to be used in some problem instances.

Experimental results are obtained by drawing electoral maps for continental Portugal. Results show that the proposed formulations are more effective than previous ones in drawing the electoral districts in Portugal. Moreover, based on results from previous elections, several gerrymandering scenarios are devised, showing that electoral outcomes can be twisted depending on the drawing of the electoral maps.

**Keywords:** Political Districting, Gerrymandering, Multi-Objective Combinatorial Optimization

## 1. Introduction

Most European democracies elect their members for the legislative body through a closed list party system. In this case, the elected members of each party are defined through a proportional representation (PR) system (such as the D'Hondt method [10]) that makes the distribution according to a predefined ordered list. Hence, in these voting systems, people do not vote on a particular candidate but in a party-list.

A different electoral system is one based on single-member electoral districts. In this system, the territory is split into a given number of electoral districts (e.g. the number of members to be elected) and each electoral district elects just one person. The voting system usually used in these districts is the first-past-the-post (FPTP) where the person with the most votes wins the district, even if a majority (50% of the votes) was not attained. The FPTP is used in countries such as the United States of America or the United Kingdom. Although less common, instant-runoff voting (e.g. Australia) and the two-round system (e.g. France) are also a possibility and these systems guarantee a majority to the winner. Alternatively, a parallel voting system in which two voting systems are combined (usually FPTP and PR) can be used, this is the case in countries such as Russia or Italy.

There are several works to build electoral district maps based on clustering [3, 4, 14, 25, 31], local search [15, 33] or metaheuristics [2, 5, 9, 17, 20]. The focus of these works is mainly on trying to maximize the compactness of the electoral districts by optimizing the distance to the geographic center of the electoral district. However, this compactness metric might still produce electoral districts with *odd* shapes.

Integer Linear Programming (ILP) models have also been proposed for electoral districting [13, 16, 32] that try to maximize compactness. However, a common difficulty of these ILP models is to ensure the contiguity of the electoral districts. In this paper, we extend previous work in several ways: (i) new and more compact models to ensure the satisfaction of contiguity constraints; (ii) new measure of compactness based on the size of the frontiers between districts; (iii) new symmetry breaking techniques that improve upon previously proposed models, as well as our own models; (iv) new multi-objective combinatorial optimization (MOCO) models to generate possible gerrymandering scenarios; (v) extensive experimental results carried out using real data from continental Portugal. Overall, we improve upon the state of the art by being able to build compact and contiguous electoral district maps more effectively than previous approaches. Moreover, we also show how one can generate electoral maps that introduce bias towards a political party and present the effects of gerrymandering on electoral results and electoral map design.

## 2. Background

This section reviews several voting systems, defines the electoral districting problem, as well as the related issue of gerrymandering. Next, previous work using incomplete and complete methods is briefly described, as we refer to the literature for a more complete survey [26].

### 2.1. Electoral Districting and Voting Systems

An electoral district is a territorial subdivision of a country which elects members to the legislative body of that country.[1] The number of elected members in each district can vary depending on the electoral system in place.

The electoral system is the set of rules that determines when the elections take place, who can vote and the voting system (how candidates are elected). There are many voting systems, but the most used are party-list proportional representation (e.g. the Portuguese and Spanish systems for legislative elections), first-past-the-post voting (also known as *winner takes all*, used in the United States of America), the two-round system (common in head of state elections such as the French or Portuguese presidential elections) and ranked voting where voters rank several candidates by preference (e.g. used to elect members to the legislature of Australia) [7, 11].

A single-member district is an electoral district where only one candidate is elected to a body with multiple members. The voting system usually used in these districts is the first-past-the-post (FPTP) where the person with the most votes is the winner even if a majority was not achieved. Although less common, instant-runoff voting (IRV) (e.g. Australia) and the two-round system (e.g. France) are also a possibility and these systems guarantee a majority to the winner. One argument against single-member districts is that they tend to favor two-party systems (Duverger's law) [7, 29] resulting in fewer minority parties in parliament. In order to introduce some proportional representation, some countries use parallel voting that combines single-member districts with party-list proportional representation.

In order to create new electoral maps, electoral districts must be created by joining territorial units (indivisible) to form clusters. There are multiple criteria to evaluate the quality of the new map but the core ones are:

1. Each elected member should represent approximately the same number of people. Ideally each elected member would represent the theoretical best value $B$ of people where $B$ equals the total number of electors divided by the number of officials to be elected.

2. All the new electoral districts should be contiguous. An electoral district $ED$ is contiguous if, and only if, to go from any point $A_1$ inside $ED$ to any other point $A_2$ inside $ED$ as well, it is not necessary to leave $ED$.

3. All districts should be compact meaning that odd shapes (see examples in Figure 1) should not exist.[2] To measure compactness multiple ideas have been proposed such as the Schwartzberg score [30], the Reock score [24] or simply summing the Euclidean distances between the geographical centers (centroids) of each territorial unit inside a district [22, 34].

[1]There can also be (possibly different) electoral districts within administrative subdivisions of a country to elect local legislatures (e.g. state elections in the United States of America).

[2]Odd shapes in electoral districting are usually associated with Gerrymandering

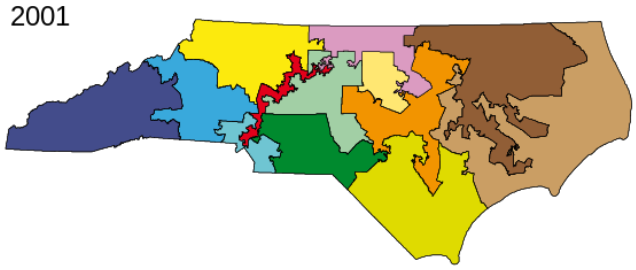


**Figure 1:** North Carolina Congressional Districts in 2001. Extracted from [12]

Although less crucial, there are other criteria (sometimes used in political districting) such as the conformity to administrative boundaries, that is, the respect of existing administrative subdivisions (used, as much as possible, in the United Kingdom) and the respect of natural boundaries in cases where mountains or rivers may be a problem for the contiguity of the districts [28].

The term gerrymandering first appeared in 1812 associated with electoral districting in the Boston area, and is often defined as the practice of redrawing the boundaries of an electoral district in order to make it more beneficial to a certain party or political faction. Although less usual, gerrymandering can also be used to increase/decrease the voting power of a racial minority (racial gerrymandering) [19]. As a result of gerrymandering, the electoral districts often end up with odd shapes. Figure 1 illustrates the practice of gerrymandering, in particular the electoral district in red.

### 2.2. Electoral Districting Methods

Most approaches to the electoral districting problem use incomplete methods. Hill climbing methods are one of the swapping-based methods to have been applied to the problem [33]. Since these methods [15, 33] only perform swaps that result in an improvement over the previous map, they can easily converge to a local optimal solution. Hence, an alternative is to use simulated annealing methods [5, 9, 20] that try random swaps to find improved approximations. Clustering algorithms have also been widely used for electoral districting [3, 4, 14, 25, 31] where a graph model for territorial representation is used.

Considering that the electoral districting problem is NP-Complete [1], all complete methods are exponential in the worst case, unless $P = NP$. Garfinkel and Nemhauser [13] were the first to use an exact approach for the political districting problem. They use enumeration techniques and the algorithm is divided in two stages. In the first one, they construct all sets of feasible districts (districts which are contiguous and meet the population requirements). In the second stage, they solve a set cover problem to choose the solutions that minimize the population deviation with respect to the ideal value.

The first Integer Linear Programming (ILP) model was proposed by Hess [16]. Let $n$ and $k$ denote respectively the total number of territorial units and the number of electoral districts to be drawn. The goal is to identify $k$ territorial units as the centers of the $k$ districts and each

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} d_{i,j}^2 \, p_i \, x_{i,j} \quad (1)$$

$$\sum_{j=1}^{n} x_{i,j} = 1 \qquad \forall i \in \{1 \dots n\} \quad (2)$$

$$\sum_{j=1}^{n} x_{j,j} = k \quad (3)$$

$$L x_{j,j} \leq \sum_{i=1}^{n} p_i \, x_{i,j} \leq U x_{j,j} \qquad \forall j \in \{1 \dots n\} \quad (4)$$

**Figure 2:** Objective function and constraints in Hess et al. [16].

territorial unit must be assigned to exactly one district center. The model has $n^2$ binary variables $x_{i,j}$ where $i, j \in \{1, \dots, n\}$ and $x_{i,j} = 1$ if unit $i$ is assigned to district with center at unit $j$, and 0 otherwise. A variable $x_{j,j}$ equals 1 if, and only if, unit $j$ is chosen as the center of a district. Let $d_{i,j}$ denote the distance between units $i$ and $j$ and let $p_i$ denote the population in territorial unit $i$, where $i, j \in \{1, \dots, n\}$. Finally, let $L$ and $U$ define the lower and upper bound on the population allowed in each district. Hence, the ILP model can be build as shown in Figure 2.

The objective function in the Hess model (1) considers the Euclidean distance between the centers of the territorial units and the population in each unit. It penalizes adding territorial units with large populations to a district where they are far away from center. The constraints are as follows: (i) a territorial unit must be assigned to only one district center (2), (ii) there must be exactly $k$ district centers (3), (iii) the population limits inside each electoral district must be between the lower and upper bounds (4). Although not in the original model, some adaptations [23, 27, 32] add that $x_{i,j} \leq x_{j,j}, \forall i, j \in \{1 \dots n\}$, meaning that if a territorial unit $i$ is assigned to a district centered at $j$, then $j$ must be the center of a district.

Observe that the model in Figure 2 does not guarantee contiguity of the electoral districts. While trying to optimize the overall compactness of the districts [16, 21] might produce some contiguous ones, this is not guaranteed for all and a post-processing algorithm (e.g. local search) must be applied to try to fix this issue.

Defining a model that satisfies the contiguity restriction on the electoral districts is not trivial. However, different techniques have been proposed [18], including by extending the Hess model [32]. Recently, Validi et al. [32], present and test four different approaches to add contiguity to the Hess model. In this paper we briefly review the flow-based formulation using Boolean variables, the authors name it the MCF model and refer to the original paper for the remaining formulations [32].

In the MCF formulation, the authors start by creating a bi-directional version of the contiguity graph $G = (V, E)$ by replacing each undirected edge $\{i, j\} \in E$ by its directed counterparts $(i, j)$ and $(j, i)$. The set of edges pointing away from vertex $i$ is denoted by $\delta^+(i)$ while, inversely, the set of edges pointing towards vertex $i$ is denoted $\delta^-(i)$. Let $f_{i,j}^{a,b}$ be a Boolean variable that denotes

$$f^{a,b}(\delta^+(b)) - f^{a,b}(\delta^-(b)) = x_{a,b} \qquad \forall a \in V \setminus \{b\}, \forall b \in V \quad (5)$$

$$f^{a,b}(\delta^+(i)) - f^{a,b}(\delta^-(i)) = 0 \qquad \forall i \in V \setminus \{a, b\}, \forall a \in V \setminus \{b\}, \forall b \in V \quad (6)$$

$$f^{a,b}(\delta^-(b)) = 0 \qquad \forall a \in V \setminus \{b\}, \forall b \in V \quad (7)$$

$$f^{a,b}(\delta^-(i)) \leq x_{i,b} \qquad \forall a, i \in V \setminus \{b\}, \forall b \in V \quad (8)$$

**Figure 3:** Contiguity Constraints for the Hess model [32].

if edge $(i, j) \in E$ is on the path to vertex $a$ from its district center $b$. If this is the case, then $f_{i,j}^{a,b} = 1$. Otherwise, $f_{i,j}^{a,b}$ must be assigned value 0. In practice, one can interpret the direct graph as a flow network and $f_{i,j}^{a,b}$ denotes the flow passing at edge $(i, j)$ considering $b$ as the source and $a$ as the sink. Moreover, let $f^{a,b}(S)$ be a shorthand for $\sum_{(i,j) \in S} f_{i,j}^{a,b}$ for a given $S \subseteq E$. Hence, by adding the constraints in Figure 3 to the Hess model, the new formulation only generates contiguous electoral districts.

In more detail, constraints (5) states that if a territorial unit $a$ has $b$ as its district center, then the flow coming out of $b$ to $a$ must equal 1. Moreover, constraints (6) ensure the flow conservation on the network. Finally, if $b$ is a district center, then its incoming flow is 0 (7) and if $i$ belongs to the flow path from $b$ to $a$, then its flow is limited to 1 (8).

### 3. Multi-Objective Combinatorial Optimization Model

In this section we present our formulations of the electoral districting problem as multi-objective combinatorial optimization model. Optimizations are proposed in Sections 3.4 and 3.5.

#### 3.1. Shortest-Path contiguity formulation

We start by presenting a new formulation for the electoral districting problem based on shortest-paths in the contiguity graph of territorial units. Although this formulation guarantees that all districts are contiguous, it is not complete in the sense that some feasible electoral districting maps are not considered. However, it is simple, effective and it seldom fails to include the global optimal solution. Moreover, even in those cases, it manages to find compact solutions, making it a valid option to use in scenarios where more complex complete formulations fail to produce an answer.

Let $N$ denote the set of territorial units (TUs) numbered from 1 to $n$. We say that two TUs are adjacent (or neighbors) if they share a common border. Let $N_j$ denote the set of neighbors of TU $j$. If two TUs $i$ and $j$ are adjacent, then $i \in N_j$ and $j \in N_i$. Let $\mathcal{K}$ denote the set of electoral districts to be build numbered from 1 to $k$.

Consider two sets of Boolean variables where each $x_{i,j}$ denotes if TU $j \in N$ is assigned to electoral district $i \in \mathcal{K}$ and each $b_{i,j,j'}$ denotes if TUs $j$ and $j'$ both belong to electoral district $i \in \mathcal{K}$.

Figure 4 contains the base formulation for the simple model. In more detail, the constraints are as follows: (i)

$$\max \sum_{i\in\mathcal{K}}\sum_{j\in N}\sum_{j'\in N_j} L_{j,j'} b_{i,j,j'} \quad (9)$$

$$\sum_{i\in\mathcal{K}} x_{i,j} = 1 \qquad \forall j \in N \quad (10)$$

$$\sum_{j\in N} x_{i,j} \geq 1 \qquad \forall i \in \mathcal{K} \quad (11)$$

$$L \leq \sum_{j\in N} v_j \cdot x_{i,j} \leq U \qquad \forall i \in \mathcal{K} \quad (12)$$

$$\neg(x_{i,j} \wedge x_{i,j'}) \implies \neg b_{i,j,j'} \qquad \forall i \in \mathcal{K},\ j \in N, j' \in N_j \quad (13)$$

**Figure 4:** Base formulation.

each TU must belong to a single electoral district (10), (ii) each electoral district must have at least one TU (11), (iii) the number of voters in each electoral district must be at least $L$ and at most $U$ (12), and (iv) if two neighbouring TUs $j$ and $j'$ are not in the same electoral district $i \in \mathcal{K}$ then $b_{i,j,j'}$ is set to 0 (13). Finally, let $L_{j,j'}$ be the border length between neighbouring TUs $j$ and $j'$. Our objective function (9) aims at maximizing the total length of the internal border lengths (i.e. minimizing the border length between different electoral districts).

Observe that our goal is also to obtain compact electoral districts, but using a different metric than other approaches [16, 32]. While other formulations try to minimize the total distance between the geographic centers of TUs to the geographic center of the electoral district, we focus on the border length as a proxy measure for the shape of the electoral district. Moreover, in some cases, the geographic center might be outside the territorial unit due to: (i) odd shape territorial unit or (ii) territorial units that contain enclaves. Therefore, the border length is also a valid measure for compactness and that is confirmed by our experimental results.

Note that the base formulation in Figure 4 must be extended with additional constraints in order to guarantee contiguity of the electoral districts. For that, consider the contiguity graph of the territorial units where a weight of 1 is defined for every edge between two adjacent territorial units. Given this weighted graph, one can compute the matrix $D$ of shortest distances between all pairs of graph vertexes (TUs) [6] in polynomial time. Based on the information from the shortest paths between all pairs of territorial units, one can define that between two non-neighbouring TUs $j$ and $j'$ that belong to the same electoral district, then one of its shortest paths in the contiguity graph must also belong to the electoral district. This can be achieved with the following constraints:

$$\forall i \in \mathcal{K}, j, j' \in V, j' \notin N_j \ :$$

$$(x_{i,j} \wedge x_{i,j'}) \implies \left( \bigvee_{j''\in N_{j'}, D_{j,j''} < D_{j,j'}} (x_{i,j''}) \right) \quad (14)$$

Notice that if two non-neighbouring TUs $j$ and $j'$ belong to the same electoral district $i$, then there must exist at least another territorial unit $j''$ in electoral district $i$ such that $j''$ is neighbour of $j'$ and the distance between $j$ and $j''$ is lower than the distance between $j$ and $j'$ ($D_{j,j''} < D_{j,j'}$). In other words, a shortest path between $j$ and $j'$ must exist inside the electoral district. Finally, note that (14) only applies if $j$ and $j'$ belong to the same electoral district $i$. Otherwise, the constraint is trivially satisfied.

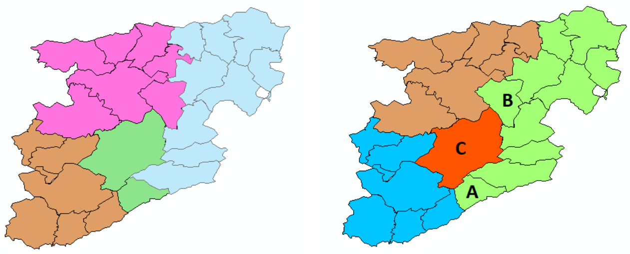


**Figure 5:** Redistricting the same area with the shortest path constraints (left map) and the excluded optimal solution (right map).

In most cases, it is beneficial for compactness to include shortest paths between territorial units. Therefore, it is rare that the most compact solution does not include a shortest path between territorial units in the electoral district. However, in Figure 5 we present one scenario where the optimal solution for model in Figure 4 is excluded due to the shortest path constraints in (14).

Note that the shortest path between territorial units A and B is 2 (passing through C). Therefore, according to the shortest path constraints in (14), the only way for A and B to be in the same district is to also contain C. However, in this context, since C is a high density territorial unit, this is not feasible and the optimal solution (Figure 5, right) is not considered due to (14). Although it is not easy to visually identify which one is better, the map on the right has a 5.22% increase on the objective function,

### 3.2. Tree-based contiguity formulation

The previous encoding allows to generate electoral maps that are contiguous, but it excludes some feasible solutions. As a result, the optimal solution might also be excluded. In this section, alternative contiguity constraints are used to ensure that the territorial units assigned to each electoral district form a tree in the contiguity graph. As a result, all electoral districts will be contiguous.

Recall the base formulation in Figure 4. In order to represent the tree, for each electoral district we extend this formulation with the following new sets of variables:

- $r_j$ denotes if territorial unit $j \in N$ is the root of a district
- $p_{j,j'}$ denotes if territorial unit $j' \in N_j$ is predecessor of $j \in N$
- $d_{j,l}$ denotes if territorial unit $j \in N$ is at depth $l$ in the tree

In this case, observe that only neighbors can be considered as predecessors of a given territorial unit $j$ in variables $p_{j,j'}$. Moreover, for the special case of the root of a given tree, those vertexes do not have a predecessor. Variables $d_{j,l}$ encode the depth in the tree for each vertex

$$\left(\sum_{j' \in N_j} p_{j,j'}\right) + r_j = 1 \qquad \forall j \in N \tag{15}$$

$$(x_{i,j} \land \neg x_{i,j'}) \implies \neg p_{j',j} \qquad \forall i \in \mathcal{K}, j \in N, j' \in N_j \tag{16}$$

$$p_{j,j'} + p_{j',j} \leq 1 \qquad \forall j \in N, j' \in N_j \tag{17}$$

$$(x_{i,j} \land x_{i,j'} \land r_j) \implies \neg r_{j'} \qquad \forall i \in \mathcal{K}, j, j' \in N, j \neq j' \tag{18}$$

$$\sum_{i=1}^{m} d_{j,i} = 1 \qquad \forall j \in N \tag{19}$$

$$r_j \implies d_{j,0} \qquad \forall j \in N \tag{20}$$

$$(p_{j,j'} \land d_{j',l}) \implies d_{j,l+1} \qquad \forall j \in N, j' \in N_j, l \in \mathcal{M} \setminus \{m\} \tag{21}$$

$$d_{j,m} \implies \neg p_{j',j} \qquad \forall j \in N, j' \in N_j \tag{22}$$

$$\sum_{j \in N} r_j = k \tag{23}$$

**Figure 6:** Tree-based formulation for contiguity of electoral districts.

$j$. The root node is considered to be at depth 0 and the depth is increased by one at each level of the tree. Let $\mathcal{M} = \{0, \ldots, m\}$ denote the set of possible depths in a tree where $m$ is the maximum number of territorial units that can be assigned to an electoral district minus 1. Considering there are limits on the number of voters in each electoral district, one can easily calculate a priori a proper value for $m$. In the worst case, for a problem instance with $n$ territorial units and $k$ electoral districts, one can safely define that $m = n - k$.

The contiguity constraints of the tree-based encoding are presented in Figure 6. In more detail, the constraints are as follows: (i) each territorial unit $j$ must have a neighbor as predecessor in the tree that represents a given electoral district or be itself the root (15), (ii) if two neighbors $j$ and $j'$ belong to different electoral districts, then they cannot have a predecessor relation (16), (iii) for each pair of neighbors $j$ and $j'$, the predecessor relation can only be established in one direction (17), (iv) in each electoral district, there can only be one root node (18), (v) each territorial unit can only be assigned a depth (19), (vi) a root node is always at depth 0 (20), (vii) the depth of a territorial unit is one more than its predecessor (21), (viii) the nodes at depth $m$ must be leafs of the tree, i.e. these nodes cannot be predecessors of other nodes (22), (ix) the number of roots must be exactly the same number as the number of electoral districts (23).

We note that tree-based contiguity constraints have been previously used. Kotthoff et al. [18] propose to encode a tree with a proper ranking among all territorial units inside the electoral district. However, our encoding shows that a topological sorting of units is not necessary.

### 3.3. Gerrymandering constraints

The encodings proposed in the previous sections are solely concerned with optimizing compactness. However, it is often the case that gerrymandering can be used to produce electoral maps favoring some political party (or being less favorable to some other party). In this section, we extend the previous encodings by considering which party is likely to win each electoral district in a first-past-the-post (FPTP) electoral system based on results from previous elections.

Let $\mathcal{P}$ denote the set of political parties to be considered in the election. In order to add gerrymandering objectives, we consider the following new sets of variables:

- $l_{i,p,p'}$ denotes if party $p \in \mathcal{P}$ is likely to lose the election at electoral district $i$ to party $p' \in \mathcal{P}$
- $l_{i,p}$ denotes if party $p \in \mathcal{P}$ is likely to lose the election at electoral district $i$ against at least one of the other parties

Let $v_{p,j}$ be an estimate of the votes to be obtained from party $p \in \mathcal{P}$ in territorial unit $j \in N$. Observe that this estimate can be obtained from previous election results. The procedure to best determine this estimation is out of scope of this paper, since our focus is on the problem formulations. In order to determine if a political party $p \in \mathcal{P}$ has fewer votes than another party $p' \in \mathcal{P}$ in electoral district $i \in \mathcal{K}$, the following constraints can be added:

$$\forall i \in \mathcal{K},\ \forall p, p' \in \mathcal{P}, p \neq p' : \quad \left(\sum_{j \in N} (v_{p',j} - v_{p,j}) x_{i,j} \geq 1\right) \iff l_{i,p,p'} \tag{24}$$

Hence, to determine if a political party $p \in \mathcal{P}$ loses the elections in electoral district $i \in \mathcal{K}$, the following is added:

$$\forall i \in \mathcal{K},\ \forall p \in P : l_{i,p} \iff \bigvee_{p' \in P, p \neq p'} l_{i,p,p'} \tag{25}$$

Finally, in order to gerrymander the building of the electoral districts, one can define an objective function that maximizes the number of electoral district wins of a certain party $p \in \mathcal{P}$. For that, one can define an objective function that minimizes the sum of all $l_{i,p}$ variables:

$$\text{Minimize} \sum_{i \in \mathcal{K}} l_{i,p} \tag{26}$$

Note that if the goal is to minimize the electoral district wins, then we simply have to create a maximizing objective function instead of a minimizing one. Furthermore, by joining the base formulation from Figure 4, the contiguity constraints from Figure 6, the party constraints (24) and (25) with the objective function (26), then we define a Multi-Objective Combinatorial Optimization (MOCO) formulation that gerrymanders the electoral district and maximizes compactness. Experimental results from this MOCO formulation applied to the case of electoral districting of Portugal are presented in section 4.

### 3.4. Additional Techniques

A trivial upper bound on the maximum depth $m$ of each tree that represents a district, in the tree-based model, is to consider that the district with the most TUs is composed of the TUs with fewer voters. Let $U$ denote the maximum number of voters to be assigned to a district and let $v_j$ denote the number of voters in TU $j$. Furthermore, let $O$ define the ordered list of the $n$ TUs indexes in a non-decreasing order of voters. Therefore, one can define $m$ as follows:

$$m = max\{u : \sum_{j=1}^{u} v_{O[j]} \leq U\} \tag{27}$$

Clearly, there can not exist a district with $u+1$ TUs, since it would not satisfy constraint (12). In the worst case, the units form a graph that corresponds to a linked list. Notwithstanding, in this case, if we consider the root of a district to be a TU in the middle of the list, there is still always a solution since our constraints defined in Section 3.2 do not preclude it. Hence, we can safely divide $m$ by 2 without removing any feasible district.

Drawing from Validi et al. [32], it is also possible to determine that some pairs of TUs cannot be in the same district. Consider the following weighted directed graph $G_w = (N, E)$ defined as follows:

- For each TU $j \in N$ there is a corresponding vertex $j \in V$
- For each pair of adjacent TU $j$ and $j'$, we define two edges $(j, j')$ and $(j', j)$ in $E$ such that $w(j, j') = v_{j'}$ and $w(j', j) = v_j$

Let $\delta(j, l)$ denote the shortest path from $j$ to $l$ in $G_w$. If $v_j + \delta(j, l) > U$, then TUs $j$ and $l$ cannot be in the same district. Note that it would require more voters than the upper limit $U$ in order for these TUs to be in the same contiguous district. Hence, the following constraint could be safely added to our formulation:

$$\forall i \in \mathcal{K} : x_{i,j} \implies \neg x_{i,l} \tag{28}$$

The same principle can also be applied in the Hess model, resulting in Equation 29 requiring TUs $j$ and $l$ to always be in different districts.

$$\forall i \in N : x_{j,i} \implies \neg x_{l,i} \tag{29}$$

### 3.5. Symmetry constraints

Symmetry constraints are simply additional constraints that allow to remove equivalent solutions.

Consider four TUs $A, B, C, D$ and two districts. Suppose that a solution is reached where $A, B$ form district 1 and $C, D$ form district 2. Notice that if we were to assign $A, B$ to district 2 and $C, D$ to district 1, then we would have essentially the same solution. We say that these two solutions are symmetric to the district assignment.

In order to avoid these symmetries between districts, one can add constraints that cut these symmetric assignments, just allowing one of them to be a solution to our model. Observe that the model remains valid, since we are only cutting symmetric solutions. In our domain, we can add constraints such that the number of voters in each district is non-decreasing according to the district identifier. Hence, let $v_j$ be the number of voters in TU $j$. Then, we can add the following constraints.

$$\forall i \in \mathcal{K} \setminus \{k\} : \sum_{j \in N} v_j \cdot x_{i,j} \leq \sum_{j \in N} v_j \cdot x_{i+1,j} \tag{30}$$

Observe that there are also symmetries inside districts. Consider the same four TUs $A, B, C, D$ to be divided between the same two districts. It is equivalent assigning $A, B$ to district 1 and $C, D$ to district 2 with $A$ and $C$ as roots or the same assignment with $B$ and $D$ as roots and $A$ and $C$ as leaves.

A possible solution to avoid inner district symmetries is to add a constraint allowing only the TU with the lowest ID to be the root of the district. It can be defined as follows:

$$\forall i \in \mathcal{K}, \forall j \in N, j' \in N_j, j < j' : \; p_{j,j'} \implies \neg r_{j'} \tag{31}$$

Equation 31 can be easily adapted to the Hess model [16] (see Figure 2) in order to cut the same type of symmetries and significantly increase the performance.

$$\forall j, j' \in V, j < j' : \; x_{j,j'} \implies \neg x_{j',j'} \tag{32}$$

Note that this optimization is only valid if the objective function used to maximize compactness does not depend on the root of the districts.

## 4. Results & discussion

In this section, our formulations proposed in section 3 are compared with previous approaches in order to test its efficiency. Therefore, section 4.1 describes the benchmark instances used in the experimental evaluation using a scenario of electoral districting in Portugal. Next, section 4.2 considers the problem of maximizing compactness in electoral districting and section 4.3 tests real-world scenarios where a party tries to gerrymander the electoral maps to favor itself.

The results presented use real data from continental Portugal.[3] The continental part is divided into 18 regions where each region elects a given number of officials to the national parliament proportional to its population. Regarding the computational infrastructure, all results were obtained on four Intel Xeon Silver 4110 processor running Debian Linux with 64GB of RAM.

### 4.1. Electoral Districting in Portugal

Portugal uses a party-list proportional representation system. However, in the context of this work, we study the scenario of changing it to a parallel voting system. In fact, this was precisely the system proposed by the two major Portuguese parties in 1998 [2, 8] making it a strong candidate for a future electoral system revision in Portugal.

In a parallel electoral system a percentage of the votes is distributed using single-member districts and the

[3] We do not consider electoral districting of the archipelagos of Azores and Madeira, since a different approach would need to be applied.

| ID | Region Name | Electoral Districts |
|---|---|---|
| 01 | Aveiro | 8 |
| 02 | Beja | 2 |
| 03 | Braga | 10 |
| 04 | Bragança | 2 |
| 05 | Castelo Branco | 2 |
| 06 | Coimbra | 5 |
| 07 | Évora | 2 |
| 08 | Faro | 5 |
| 09 | Guarda | 2 |
| 10 | Leiria | 5 |
| 11 | Lisboa | 24 |
| 12 | Portalegre | 1 |
| 13 | Porto | 20 |
| 14 | Santarém | 5 |
| 15 | Setúbal | 9 |
| 16 | Viana do Castelo | 3 |
| 17 | Vila Real | 3 |
| 18 | Viseu | 4 |
| Total | | 112 |

**Table 1:** Identifiers, names and number of electoral districts to be created for each region under a parallel voting system.

rest using a party-list proportional representation method (most commonly, the D'Hondt method) at the national level. The idea behind it is bringing the voters closer to politics by voting directly to elect a parliament member who is typically more concerned and connected with their local electoral area, while maintaining some proportionality through the national circle at the party level (avoiding the tendency to create a two-party system of a first-past-the-post (FPTP) system).

In the parallel voting system scenario, we consider the number of single-member districts to be created in each Portuguese region[4] as half the current number of elected officials, rounded up. Hence, the total number of single-member districts becomes 112 in continental Portugal as shown in Table 1.

To generate the electoral maps (instances) tested in the next sections we follow a set of rules, typical for redistricting, which are: (i) The number of people registered to vote in each electoral district (ED) must not diverge more than 25% from the theoretical best value, (ii) the new electoral districts should be as compact as possible, (iii) all electoral districts must be contiguous, (iv) there must be conformity to administrative boundaries, the largest possible administrative divisions should be kept whenever possible without disregarding the first rule.

The population margin is difficult to set. The bigger the margin from a theoretical ideal value, the less equal is the voting power between electoral districts. Some countries such as the United States of America prefer a margin as low as possible. However, this is only possible because administrative boundaries are ignored in favor of census tracts. Other countries set the margin limit at values such as 10% (Italy, Australia or Ukraine). In our scenarios, the maximum margin value is set at 25%, as used in countries such as Canada or Germany. Moreover, it was the maximum margin value presented in one of the propositions for Portugal back in 1998 [2, 8].

[4]We refer to the Portuguese first-level administrative divisions (called *districtos* in Portuguese) as regions to avoid confusion with electoral districts.

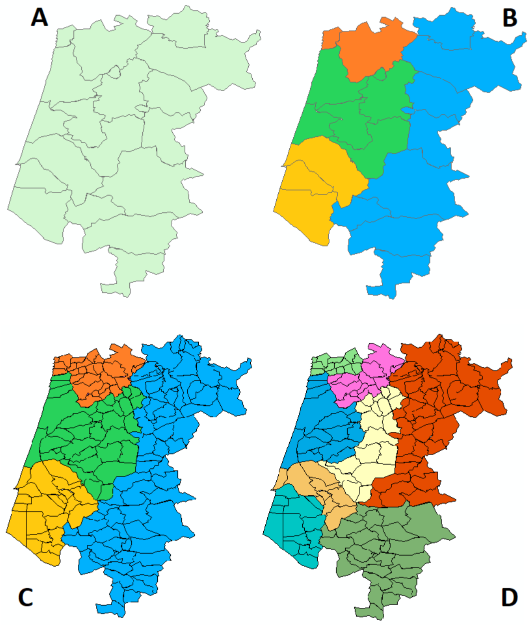


**Figure 7:** Steps to redistrict at the parish level.

Whenever possible, we try to preserve the municipalities inside each Portuguese region. However, that is not possible in most of the cases, either because the number of electoral districts is larger than the number of municipalities of the region, or because the population differences between municipalities does not allow us to join them while respecting the population similarity rule. In the aforementioned cases, we start by splitting the region in as many different areas as possible at the municipality level. Next, each of these areas is then redistricted at the civil parish level to define the electoral map of a region.

Figure 7 exemplifies our approach. Consider the region of Aveiro (number 01) and its 19 municipalities (map A in Figure 7). This region is to be split into 8 electoral districts. However, this problem is unsatisfiable at the municipality level, since there is a municipality with 125534 registered voters and the upper population bound for each electoral district is only 100883. Hence, we first split the region into 4 areas (map B in Figure 7) where each area contains roughly the same population (i.e. inside the maximum margin of 25%), is contiguous and maximizes compactness. Next, we consider the civil parishes inside each of these areas as the territorial units (map C in Figure 7). Finally, districting of each of these 4 areas into 2 single-member districts is performed to achieve our final electoral map for the region of Aveiro (map D in Figure 7).

After this first step, we end up with 35 areas of continental Portugal where a given number of electoral districts are to be defined. These 35 areas are characterized in Table 2 and define the benchmark instances used to evaluate the proposed ideas of the paper.

In the next sections, we provide experimental results in order to answer the following research questions: (i) can we improve previous models with the additional techniques from Sections 3.4 and 3.5? (ii) How does our tree-

| Instance Number | Portuguese Region | Territorial Units | Electoral Districts |
|---|---|---|---|
| 1 | 01 | 22 | 2 |
| 2 | 01 | 25 | 2 |
| 3 | 01 | 33 | 2 |
| 4 | 01 | 67 | 2 |
| 5 | 02 | 14 | 2 |
| 6 | 03 | 34 | 2 |
| 7 | 03 | 37 | 2 |
| 8 | 03 | 53 | 2 |
| 9 | 03 | 70 | 2 |
| 10 | 03 | 153 | 2 |
| 11 | 04 | 12 | 2 |
| 12 | 05 | 11 | 2 |
| 13 | 06 | 64 | 2 |
| 14 | 06 | 91 | 3 |
| 15 | 07 | 14 | 2 |
| 16 | 08 | 16 | 5 |
| 17 | 09 | 14 | 2 |
| 18 | 10 | 57 | 2 |
| 19 | 10 | 53 | 3 |
| 20 | 11 | 15 | 6 |
| 21 | 11 | 15 | 6 |
| 22 | 11 | 24 | 6 |
| 23 | 11 | 80 | 6 |
| 24 | 12 | 15 | 1 |
| 25 | 13 | 11 | 4 |
| 26 | 13 | 15 | 4 |
| 27 | 13 | 25 | 4 |
| 28 | 13 | 43 | 4 |
| 29 | 13 | 149 | 4 |
| 30 | 14 | 21 | 5 |
| 31 | 15 | 12 | 4 |
| 32 | 15 | 43 | 5 |
| 33 | 16 | 10 | 3 |
| 34 | 17 | 14 | 3 |
| 35 | 18 | 24 | 4 |

**Table 2:** Correspondence between instance number, Portuguese region, the number of territorial units and districts to be created.

| Instance Number | Time without Optimizations | Time with Optimizations |
|---|---|---|
| 1 | 768.84 | 9.47 |
| 2 | 10800 | 30.49 |
| 3 | 10800 | 148.31 |
| 4 | 10800 | 10800 |
| 5 | 8.73 | 0.56 |
| 6 | 10800 | 501.70 |
| 7 | 10800 | 525.94 |
| 8 | 10800 | 10800 |
| 9 | 10800 | 10800 |
| 10 | 10800 | 10800 |
| 11 | 3.68 | 0.66 |
| 12 | 0.94 | 0.05 |
| 13 | 10800 | 10800 |
| 14 | 10800 | 10800 |
| 15 | 8.74 | 0.51 |
| 16 | 6.87 | 0.51 |
| 17 | 9.20 | 0.85 |
| 18 | 10800 | 2241.33 |
| 19 | 10800 | 10800 |
| 20 | 6.59 | 0.66 |
| 21 | 4.41 | 0.15 |
| 22 | 10800 | 46.14 |
| 23 | 10800 | 10800 |
| 24 | 0.20 | 0.04 |
| 25 | 1.25 | 0.12 |
| 26 | 252.93 | 4.33 |
| 27 | 5486.67 | 7.33 |
| 28 | 10800 | 10800 |
| 29 | 10800 | 10800 |
| 30 | 8106.49 | 7.52 |
| 31 | 3.72 | 0.85 |
| 32 | 10800 | 10800 |
| 33 | 1.78 | 0.76 |
| 34 | 14.27 | 0.91 |
| 35 | 10800 | 4.60 |

**Table 3:** Comparison between the execution times of the Hess model [16] with the contiguity constraints by Validi et al. [32] with and without the optimizations proposed in Sections 3.4 and 3.5.

based formulation compare with previously proposed formulations? (iii) The shortest-path based formulation can just provide an approximation. How far are these results from the optimal? What about the performance? (iv) Can we solve the districting problem for Portugal just considering compactness? (v) Can we gerrymander the political districts? What flip can we make to the electoral results?

#### 4.2. Maximizing compactness

In order to answer the first research question, we implemented the the original Hess model [16] (see Figure 2) with the contiguity constraints by Validi et. al. [32] (see Figure 3), abbreviated simply as *Hess model* in the next sections. Next, the optimizations proposed in Sections 3.4 and 3.5 were introduced. Both implementations are evaluated in the 35 benchmark instances defined in the previous section where the goal in each instance is solely to optimize compactness. Therefore, both formulations were evaluated using CPLEX 12.6.0 [5] with a time limit of three hours (10800 seconds).

Table 3 presents the results for the Hess formulation with and without the proposed techniques. Whenever that limit was exceeded, the background of the respective cell is highlighted. Otherwise, the time spent to find the optimal solution is shown. The experimental results clearly show the effectiveness of cutting symmetries in the Hess formulation. Not only can we increase the number of solved instances within the time limit, but also greatly improve the performance in the remaining instances. Observe that the performance never worsens and in some cases improves by two orders of magnitude. Although not shown, in instances where the time limit is reached, a better solution is usually obtained when using the techniques proposed in the paper.

Next, we compare the improved Hess model with the proposed tree-based (see section 3.2) and shortest path-based (see section 3.1) formulations. Table 4 shows the computational time for each approach (in seconds). Whenever the time limit is reached, that entry's background is highlighted. Moreover, the fastest computational time between the optimized Hess model and the tree-based model is also highlighted in bold.

Overall, it is clear that the proposed tree-based formulation improves upon the Hess model being able to solve more instances. Moreover, when both formulations are able to find an optimal solution, the tree-based formulation is usually much faster. Finally, observe that CPLEX is able to solve all instances of the shortest-path based formulation. However, in some cases the optimal solution of the shortest-path model is worse, since this formulation excludes some feasible solutions. Nevertheless, as can be seen in the last column of Table 4, for this set of instances, the quality of the solution using the tree-based model never increases by more than 6% with regard to the solution found by the shortest-path formulation. Moreover, for the 5 instances where the tree-based

[5] https://www.ibm.com/analytics/cplex-optimizer

| Instance Number | Hess Model | Tree-based Model | Shortest Path Model | Quality Increase |
|---|---|---|---|---|
| 1 | 9.47 | **0.47** | 0.48 | 0.00% |
| 2 | 30.49 | **0.82** | 0.30 | 0.00% |
| 3 | 148.31 | **8.62** | 0.52 | **1.33%** |
| 4 | 10800 | 10800 | 1.96 | TLE |
| 5 | 0.56 | **0.38** | 0.26 | 0.00% |
| 6 | 501.7 | **5.47** | 0.64 | 0.00% |
| 7 | 525.94 | **1.34** | 0.75 | 0.00% |
| 8 | 10800 | **26.32** | 1.61 | 0.00% |
| 9 | 10800 | **1882.42** | 3.28 | **1.80%** |
| 10 | 10800 | 10800 | 4.80 | TLE |
| 11 | 0.66 | **0.56** | 0.34 | 0.00% |
| 12 | 0.05 | **0.04** | 0.03 | **5.15%** |
| 13 | 10800 | **575.96** | 1.77 | **1.29%** |
| 14 | 10800 | 10800 | 18.44 | TLE |
| 15 | 0.51 | **0.38** | 0.01 | **6.13%** |
| 16 | **0.51** | 1.01 | 0.81 | 0.00% |
| 17 | 0.85 | **0.42** | 0.02 | 0.00% |
| 18 | 2241.33 | **20.07** | 0.74 | 0.00% |
| 19 | 10800 | **35.20** | 5.20 | **0.67%** |
| 20 | **0.66** | 5.63 | 0.71 | 0.00% |
| 21 | **0.15** | 0.69 | 0.70 | 0.00% |
| 22 | **46.14** | 53.26 | 11.42 | 0.00% |
| 23 | 10800 | 10800 | 8939.96 | TLE |
| 24 | 0.04 | **0.01** | 0.01 | 0.00% |
| 25 | **0.12** | 0.40 | 0.66 | 0.00% |
| 26 | 4.33 | **1.32** | 0.37 | 0.00% |
| 27 | 7.33 | **3.97** | 0.98 | 0.00% |
| 28 | 10800 | **711.26** | 6.08 | **0.72%** |
| 29 | 10800 | 10800 | 1839.54 | TLE |
| 30 | 7.52 | **4.89** | 1.28 | 0.00% |
| 31 | 0.85 | **0.48** | 0.24 | 0.00% |
| 32 | 10800 | **225.03** | 19.21 | **0.56%** |
| 33 | 0.76 | **0.49** | 0.32 | 0.00% |
| 34 | 0.91 | **0.41** | 0.42 | 0.00% |
| 35 | 4.60 | **2.20** | 0.53 | **5.22%** |

**Table 4:** Execution times of the Hess optimized model and our proposed Tree-based (complete) and Shortest-path (incomplete) models for all instances. Highlighted in orange are cases where the time limit (3 hours) was exceeded (TLE). In bold is the fastest model between the Hess model and the Tree-based model.

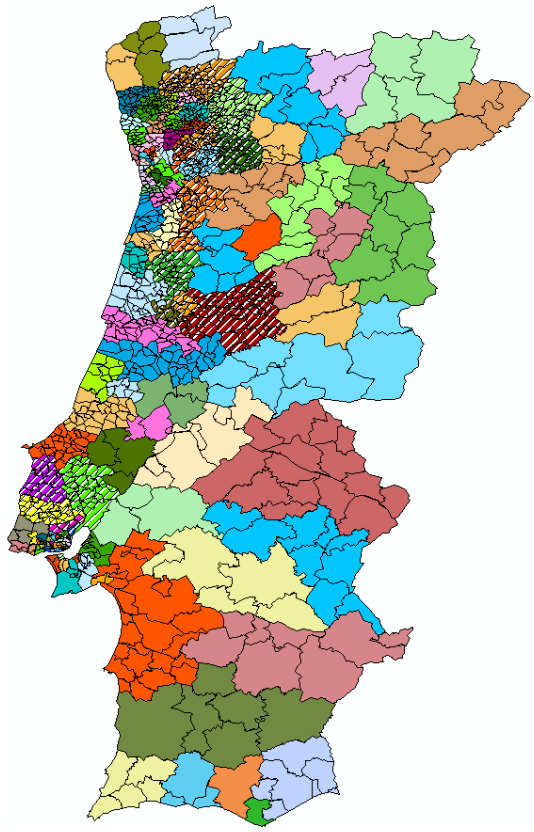

**Figure 8:** Complete electoral map for continental Portugal under a parallel voting system.

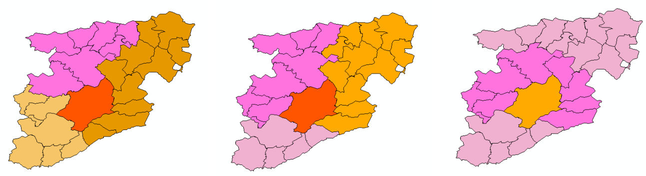

**Figure 9:** Pareto front for the region of Viseu (18) at the municipality level. In pink EDs won by party PS, in orange EDs won by party PSD.

model reached the time limit, CPLEX was not able to find an integer solution. As a result, for these instances, there is no comparison on the quality of solutions.

The benchmark instances used to evaluate the different problem formulations represent real-world areas of Portugal. Hence, it is possible to draw an electoral map of Portugal under a parallel voting system. As a result, Figure 8 presents the complete electoral map for continental Portugal. Note that all the electoral districts are contiguous meaning that a change in color also represents a change in district. There are 15 striped electoral districts. A district is striped whenever the tree-based model exceeded our time limits and the presented results are the ones obtained with the shortest-path model which may not be the global optimum solution for that particular region. Overall, there are no major compactness issues (such as the ones in Figure 1) thus showing the capabilities of the proposed objective function for compactness, as well as the formulations proposed in this paper.

### 4.3. Gerrymandering Electoral Maps

Since the establishment of democratic elections in Portugal in the 1970's, the political spectrum has been dominated by two major political parties: Socialist Party (PS) and the Social Democratic Party (PSD). Hence, the drawing of electoral maps is bound to be constrained by people from these two parties that might decided to use other criteria than compactness. In this section, we consider a scenario where a political party tries to gerrymander the electoral maps to its favour based on previous electoral results and evaluate the overall impact on the number of elected officials for parliament.

In order to maximize the results of a party, for each instance in Table 2 two more instances were generated, also creating biased maps towards each of the main Portuguese parties using the formulations in Section 4.3. To solve multi-objective combinatorial optimization problems we use the Sat4jMoco solver[6].

In our results, given that the associated colours of parties PS and PSD are, respectively, pink and orange, districts colored in shades of pink are wins for party PS and colored in shades of orange are wins for party PSD.

There are 3 possible redistricting options for the region of Viseu. These 3 points of the Pareto front are the trade-off between electoral district wins (for party PS) and compactness. On the left map in Figure 9 only compactness is maximized and represents the most favourable distribu-

[6] `https://gitlab.ow2.org/sat4j/moco`

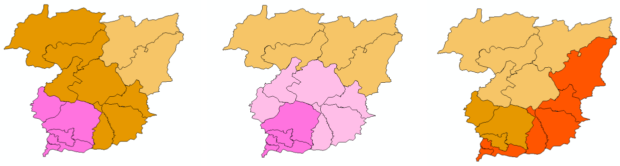

**Figure 10:** Different results for the region of Vila Real (17). In pink EDs won by party PS and in orange EDs won by party PSD.

tion towards PSD. On this map the border length between TUs in the same district is 614891 meters and it goes down as the number of ED wins for party PS increases, first to 611282 meters (center map) and then to 594286 meters (right map). These values represent a 0.59% and a 3.35% decrease in compactness to the map that only maximizes compactness. In the most biased map the pink districts starts forming a snake around the orange district which might be an indicator of gerrymandering.

For the region of Vila Real the results are presented in Figure 10. Optimizing only the compactness (left map in Figure 10), the orange party wins 2 districts and the compactness measure is 440210 meters. If we are also trying to maximize the wins of the pink party (center map), it is possible to make PS win 2 out of the 3 districts but the length of the border between TUs in the same district goes down to 426060 meters (a 3.21% decrease). Finally, maximizing the wins of PSD along with compactness (Figure 10, right map), PSD would win all the seats but we can immediately notice that one of the districts has an extremely odd shape and that is confirmed by our compactness measure now being only 371788 meters (a 15.54% decrease to the unbiased version).

The districting was done, independently, for all Portuguese regions, and despite a complete map not being presented, Table 5 presents the complete electoral results (number of seats won by each party) in each area.

In the 2019 elections (data used for the biased distributions) PS received 36.65% of the national vote while PSD received 27.90%. [7] Consequentially, at the national level, PS is expected to be the party with the most parliament members under a FPTP system. This prediction is confirmed by the complete electoral results at the national level (Table 5). They show that under an unbiased distribution PS would win 84 out of 112 seats (75%) and PSD the remaining 28 seats (25%). However, using gerrymandering, in the most favorable distribution towards PS, it is possible to make it win 13 more seats, bringing its total to 97 (86.61%). Contrarily, under an unfavorable distribution it can also lose 8 seats to the opposing party leaving it with 76 members in parliament out of the 112 distributed using the FPTP voting system (67.86%).

## 5. Conclusions

In this work we develop a novel and compact multi-objective combinatorial optimization model with the ability of creating electoral districts that are: compact, contiguous and give voters similar popular representation. Two formulations to create contiguous districts are presented, a complete one (capable of finding global optimal solutions) and an incomplete, but dramatically faster version. Optimizations that can also be adapted to classical approaches to the redistricting problem are also presented. The developed model is versatile and can also be used to create gerrymandered maps using data from previous elections, allowing the study of the effects of gerrymandering in electoral maps.

Following the propositions to change the Portuguese electoral system to one using single-member districts, the experimental results are obtained drawing electoral maps for continental Portugal. The results show that the proposed model performs better than a classical approach in this task and can deliver compact and contiguous solutions within the population boundaries while conforming with current administrative divisions.

Additionally, using the gerrymandering capacities of the model and official electoral results from previous legislative elections, the impact of gerrymandering is studied. The results show that gerrymandering significantly affects the compactness of the electoral districts, creating odd shapes, and has the power to completely change the electoral results.

| Electoral Results - Section 4.3 | | | | | | |
|---|---|---|---|---|---|---|
| District | Unbiased | | Max PS | | Max PSD | |
| | PS | PSD | PS | PSD | PS | PSD |
| 1 | 1 | 1 | 1 | 1 | 0 | 2 |
| 2 | 2 | 0 | 2 | 0 | 2 | 0 |
| 3 | 2 | 0 | 2 | 0 | 1 | 1 |
| 4 | 0 | 2 | 1 | 1 | 0 | 2 |
| 5 | 2 | 0 | 2 | 0 | 2 | 0 |
| 6 | 1 | 1 | 2 | 0 | 1 | 1 |
| 7 | 1 | 1 | 2 | 0 | 1 | 1 |
| 8 | 2 | 0 | 2 | 0 | 2 | 0 |
| 9 | 0 | 2 | 0 | 2 | 0 | 2 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 |
| 11 | 0 | 2 | 0 | 2 | 0 | 2 |
| 12 | 2 | 0 | 2 | 0 | 2 | 0 |
| 13 | 2 | 0 | 2 | 0 | 1 | 1 |
| 14 | 3 | 0 | 3 | 0 | 3 | 0 |
| 15 | 2 | 0 | 2 | 0 | 2 | 0 |
| 16 | 5 | 0 | 5 | 0 | 5 | 0 |
| 17 | 2 | 0 | 2 | 0 | 1 | 1 |
| 18 | 1 | 1 | 2 | 0 | 1 | 1 |
| 19 | 1 | 2 | 2 | 1 | 0 | 3 |
| 20 | 6 | 0 | 6 | 0 | 6 | 0 |
| 21 | 5 | 1 | 5 | 1 | 5 | 1 |
| 22 | 5 | 1 | 6 | 0 | 3 | 3 |
| 23 | 6 | 0 | 6 | 0 | 6 | 0 |
| 24 | 1 | 0 | 1 | 0 | 1 | 0 |
| 25 | 2 | 2 | 3 | 1 | 2 | 2 |
| 26 | 4 | 0 | 4 | 0 | 4 | 0 |
| 27 | 3 | 1 | 3 | 1 | 3 | 1 |
| 28 | 3 | 1 | 4 | 0 | 3 | 1 |
| 29 | 3 | 1 | 4 | 0 | 3 | 1 |
| 30 | 4 | 1 | 4 | 1 | 4 | 1 |
| 31 | 4 | 0 | 4 | 0 | 4 | 0 |
| 32 | 5 | 0 | 5 | 0 | 5 | 0 |
| 33 | 1 | 2 | 2 | 1 | 1 | 2 |
| 34 | 1 | 2 | 2 | 1 | 0 | 3 |
| 35 | 1 | 3 | 3 | 1 | 1 | 3 |
| **TOTAL** | 84 | 28 | 97 | 15 | 76 | 36 |

**Table 5:** Electoral results for the districting in Section 4.3. Cells highlighted in yellow represent instances where the time limit was exceeded, hence the presented results are the best solution found after 3 hours.

[7] https://www.eleicoes.mai.gov.pt/legislativas2019/territorio-nacional.html

## References


[1] M. Altman. Is automation the answer: the computational complexity of automated redistricting. *Rutgers Computer and Law Technology Journal*, 23:81–142, 1997.

[2] F. Bação, V. S. Lobo, and M. Painho. Applying genetic algorithms to zone design. *Soft Comput.*, 9(5):341–348, 2005.

[3] L. Bodin. A districting experiment with a clustering algorithm. *Annals of the New York Academy of Sciences*, 219:209 – 214, 12 2006.

[4] N. Bottman, W. Essig, and S. Whittle. Why weight? a cluster-theoretic approach to political districting. *The UMAP Journal*, 28(3), 2007.

[5] M. H. Browdy. Simulated annealing: An improved computer model for political redistricting. *Yale Law & Policy Review*, 8(1):163–179, 1990.

[6] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, Second Edition*. The MIT Press and McGraw-Hill Book Company, 2001.

[7] G. W. Cox. *Making votes count: strategic coordination in the world's electoral systems*. Cambridge University Press, 1997.

[8] M. B. da Cruz. A revisão falhada do sistema eleitoral. *Análise Social*, 35(154/155):45–53, 2000.

[9] S. J. D'Amico, S. Wang, R. Batta, and C. M. Rump. A simulated annealing approach to police district design. *Computers & OR*, 29(6):667–684, 2002.

[10] V. D'Hondt. *Système pratique et raisonné de représentation proportionnelle*. librairie C. Muquardt, Merzbach et Falk, éditeurs, 1882.

[11] D. M. Farrell. *Electoral systems: A comparative introduction*. Macmillan International Higher Education, 2011.

[12] Furfur. Wikimedia commons north carolina congressional districts 1992-2001, 2017.

[13] R. S. Garfinkel and G. L. Nemhauser. Optimal political districting by implicit enumeration techniques. *Management Science*, 16(8):B495–B508, 1970.

[14] O. Guest, F. J. Kanayet, and B. C. Love. Gerrymandering and computational redistricting. *Journal of Computational Social Science*, 2(2):119–131, Jul 2019.

[15] B. Hayes. Machine politics. *American Scientist*, 84(6):522–526, 1996.

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

[17] D. Joshi, L. Soh, and A. Samal. Redistricting using constrained polygonal clustering. *IEEE Trans. Knowl. Data Eng.*, 24(11):2065–2079, 2012.

[18] L. Kotthoff, B. O'Sullivan, S. S. Ravi, and I. Davidson. Complex Clustering Using Constraint Programming: Modelling Electoral Map Creation. In *14th International Workshop on Constraint Modelling and Reformulation*, Sept. 2015.

[19] D. Lublin. *The paradox of representation: Racial gerrymandering and minority interests in Congress*. Princeton University Press, 1999.

[20] W. Macmillan. Redistricting in a gis environment: An optimisation algorithm using switching-points. *Journal of geographical systems*, 3(2):167–180, 2001.

[21] A. Mehrotra, E. L. Johnson, and G. L. Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1114, 1998.

[22] R. G. Niemi, B. Grofman, C. Carlucci, and T. Hofeller. Measuring compactness and the role of a compactness standard in a test for partisan and racial gerrymandering. *The Journal of Politics*, 52(4):1155–1181, 1990.

[23] J. Oehrlein and J.-H. Haunert. A cutting-plane method for contiguity-constrained spatial aggregation. *Journal of Spatial Information Science*, 2017, 12 2017.

[24] E. C. Reock. A note: Measuring compactness as a requirement of legislative apportionment. *Midwest Journal of Political Science*, 5(1):70–74, 1961.

[25] F. Ricca, A. Scozzari, and B. Simeone. Weighted voronoi region algorithms for political districting. *Math. Comput. Model.*, 48(9-10):1468–1477, 2008.

[26] F. Ricca, A. Scozzari, and B. Simeone. Political districting: from classical models to recent approaches. *Ann. Oper. Res.*, 204(1):271–299, 2013.

[27] F. Ricca, A. Scozzari, and B. Simeone. Political districting: From classical models to recent approaches. *Annals of Operations Research*, 204:271–299, 04 2013.

[28] F. Ricca and B. Simeone. Political districting: Traps, criteria, algorithms and tradeoffs. *Ricerca Operativa AIRO*, 27:81–119, 01 1997.

[29] W. H. Riker. The two-party system and duverger's law: An essay on the history of political science. *American Political Science Review*, 76(4):753–766, 1982.

[30] J. E. Schwartzberg. Reapportionment, gerrymanders, and the notion of compactness. *Minnesota Law Review*, 50(1701):443, 1966.

[31] L. Svec, S. Burden, and A. Dilley. Applying voronoi diagrams to the redistricting problem. *The UMAP Journal*, 28(3):313–329, 2007.

[32] H. Validi, A. Buchanan, and E. Lykhovyd. Imposing contiguity constraints in political districting models. *Optimization Online*, 2020.

[33] W. Vickrey. On the prevention of gerrymandering. *Political Science Quarterly*, 76(1):105–110, 1961.

[34] H. P. Young. Measuring the compactness of legislative districts. *Legislative Studies Quarterly*, 13(1):105–115, 1988.