Capturing Uncertainty in Districting Problems: Towards More Comprehensive Modeling Frameworks

Francisco Saldanha-da-Gama

Departmento de Estatística e Investigação Operacional,
Centro de Matemática Aplicações Fundamentais e Investigação Operacional,
Faculdade de Ciências, Universityof Lisbon, Portugal

Seminar presented at
The University of Hong Kong
Systems Analytics Global Leaders' Seminars
March 23, 2022

Capturing Uncertainty in Districting Problems: Towards More Comprehensive Modeling Frameworks

Francisco Saldanha-da-Gama

Departmento de Estatística e Investigação Operacional,
Centro de Matemática Aplicações Fundamentais e Investigação Operacional,
Faculdade de Ciências, Universityof Lisbon, Portugal

Seminar presented at
The University of Hong Kong
Systems Analytics Global Leaders' Seminars
March 23, 2022

Collaboration:

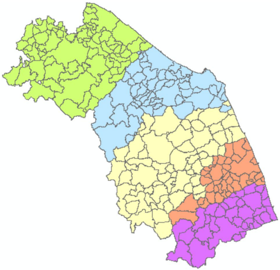
Antonio Diglio, Stefan Nickel, Juanjo Peiró, Anika Pomes, Carmela Piccolo

Outline

Outline

Districting

A Districting Problem (DP) consists of grouping a set of Territorial Units (TUs) into a given number of larger geographic clusters, called districts.



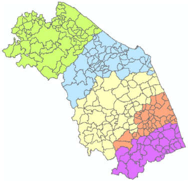
Some planning criteria are usually considered in districting problems:

INTEGRITY BALANCING COMPACTNESS CONTIGUITY

Planning criteria

Integrity

Each TU is assigned exactly to one district.



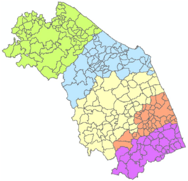
Balancing

Given a certain "activity measure" for a TU (e.g. demand), balancing expresses the need for districts of similar size with respect to the considered measure(s).

Planning criteria

Compactness

Districts are created in such a way that the TUs in each district are “close” to each other.



Contiguity

Do not cross other districts to move between TUs of the same district (no enclaves or exclaves).

Applications and related literature

DPs have applications in many areas:

- **Political districting**.
  Garfinkel and Nemhauser (1970), Ricca and Simeone (2008), Ricca et al. (2013).
- **Strategic service planning and management (e.g. in health care)**.
  Benzarti et al. (2013), Blais et al. (2003), Mostafayi Darmian et al. (2021).
- **School systems**.
  Bruno et al. (2016), Caro et al. (2004), Ferland and Guénette (1990), Schoepfle and Church (1991).
- **Energy and power distribution networks**.
  Bergey et al. (2003), de Assis et al. (2014), Yanık et al. (2014).

Applications and related literature

DPs have applications in many areas:

- **Design of police districts**.
  D'Amico et al. (2002).
- **Waste collection**.
  Mourão et al. (2009), Ríos-Mercado and Bard (2019).
- **Transportation**.
  Bruno and Laporte (2002), Tavares et al. (2007).
- **Design of commercial areas to assign sales forces**.
  Ríos-Mercado and López-Pérez (2013), Zoltners and Sinha (2005).
- **Distribution Logistics**.
  Konur and Geunes (2019), Zhong et al. (2007).
- **...**

Outline

The model by Hesse et al. (1965)

Hess et al. (1965) introduced a seminal optimization model for districting.

Create p districts according to three criteria:

(i) Integrity;

(ii) Balancing;

(iii) Compactness.

The number of districts p is given beforehand.

The model by Hesse et al. (1965)

Some notation:

I set of basic territorial units (TUs).

d_i activity level in TU $i \in I$.

$\downarrow$

"demand"

ℓ_{ij} distance between TU i and TU j ($i, j \in I$).

The model by Hesse et al. (1965)

Ensuring BALANCING

The demand in the districts should be similar:

$$w(D_1) \cong w(D_2) \cong \ldots \cong w(D_p)$$

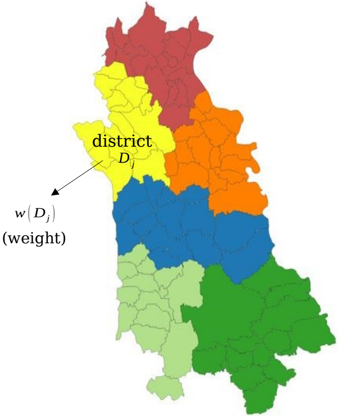
How to ensure this?

μ average demand per district.

$$\mu = \frac{1}{p} \sum_{i \in I} d_i.$$

α maximum deviation allowed w.r.t. μ.

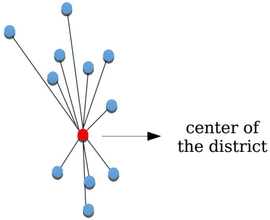
(The demand in a district should not deviate from μ more than $100 \times \alpha\%$.)

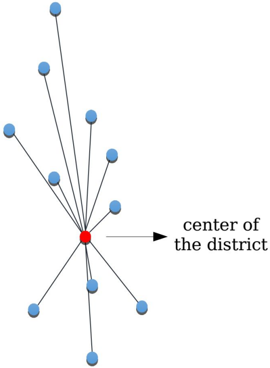


The model by Hesse et al. (1965)

Ensuring COMPACTNESS

Which one is more compact?





The model by Hesse et al. (1965)

Ensuring COMPACTNESS

ℓ_{io}, distance between a TU i and the “central TU”, say o, of its district.

$c_{io} \leftarrow \ell_{io}$ or $c_{io} \leftarrow \ell_{io}^2$

$\downarrow$ $\downarrow$

“Cost” for including TU i in the district that has o as the central TU.

$$\sum_{i \in D_o} c_{io} \quad \rightarrow \quad \text{Measure of compactness for the district.}$$

The model by Hesse et al. (1965)

$\surd$ Every district will have a TU representing it ("center" of the district).

$\surd$ TU i is assigned to TU j $\equiv$ TU i is in the district represented by TU j.

Decision variables:

$$x_{ij} = \begin{cases} 1, & \text{if TU } i \text{ is assigned to TU } j, \\ 0, & \text{otherwise,} \end{cases} \qquad \forall i, j \in I.$$

For some $j \in I$:

$$x_{jj} = 1 \quad \equiv \quad \text{TU } j \text{ is assigned to itself.}$$

$$\equiv \quad \text{TU } j \text{ is the representative—center—of its district.}$$

The model by Hesse et al. (1965)

$$
\begin{aligned}
\min \quad & \sum_{i \in I} \sum_{j \in I} c_{ij} x_{ij}, & \\
\text{s. t.} \quad & \sum_{j \in I} x_{ij} = 1, & \forall i \in I, \\
& \sum_{j \in I} x_{jj} = p, & \\
& (1-\alpha)\, \mu\, x_{jj} \leq \sum_{i \in I} d_i x_{ij} \leq (1+\alpha)\, \mu\, x_{jj}, & \forall j \in I, \\
& x_{ij} \leq x_{jj}, & \forall i, j \in I, \\
& x_{ij} \in \{0, 1\}, & \forall i, j \in I.
\end{aligned}
$$

The model by Hesse et al. (1965)

$$
\begin{aligned}
\min \quad & \sum_{i\in I}\sum_{j\in I} c_{ij}x_{ij}, && && \text{Compactness}\\
\text{s. t.} \quad & \sum_{j\in I} x_{ij} = 1, && \forall i \in I, && \text{Integrity}\\
& \sum_{j\in I} x_{jj} = p, && && p \text{ Districts}\\
& (1-\alpha)\,\mu\, x_{jj} \le \sum_{i\in I} d_i x_{ij} \le (1+\alpha)\,\mu\, x_{jj}, && \forall j \in I, && \text{Balancing}\\
& x_{ij} \le x_{jj}, && \forall i, j \in I, && \text{Allocation}\\
& x_{ij} \in \{0,1\}, && \forall i, j \in I. && \text{Integrity}
\end{aligned}
$$

Discrete p-median problem with balancing constraints!

The model by Hesse et al. (1965)

$$\min \quad \sum_{i\in I}\sum_{j\in I} c_{ij}x_{ij}, \qquad \text{Compactness}$$

$$\text{s. t.} \quad \sum_{j\in I} x_{ij} = 1, \qquad \forall i \in I, \qquad \text{Integrity}$$

$$\sum_{j\in I} x_{jj} = p, \qquad p \text{ Districts}$$

$$(1-\alpha)\,\mu\, x_{jj} \le \sum_{i\in I} d_i x_{ij} \le (1+\alpha)\,\mu\, x_{jj}, \qquad \forall j \in I, \qquad \text{Balancing}$$

$$x_{ij} \le x_{jj}, \qquad \forall i,j \in I, \qquad \text{Necessary?}$$

$$x_{ij} \in \{0,1\}, \qquad \forall i,j \in I. \qquad \text{Integrity}$$

Discrete p-median problem with balancing constraints!

Outline

Demand changes

But...in practice demand (activity level) may change!

- expansion of urban areas;
- migration flows;
- economic conditions;
- ...

How to hedge against such changes?

Depends on the problem we are solving!

- Re-districting;
- Multi-period districting;
- Districting under uncertainty;
- ...

Demand changes

Re-districting

- $\surd$ The demand changing has already occurred;
- $\surd$ A reactive posture is assumed;
- $\surd$ A so-called redistricting problem is solved.

$$\downarrow$$

Optimization problem aiming at redesigning existing districts.

Demand changes

Multi-period districting

- $\surd$ The demand has still not changed;
- $\surd$ Demand changes can be predicted for some future—planning horizon;
- $\surd$ A proactive posture is assumed;
- $\surd$ A multi-period districting problem is solved.

$$\downarrow$$

A plan (districting/redistricting) is devised to cope with the varying demand.

Demand changes

Districting under uncertainty

- √ Demand is uncertain—cannot be forecasted accurately;
- √ A proactive posture is assumed;
- √ Different paradigms can be considered:
  - **Online optimization**
    Typically for short term uncertainty.
  - **Robust optimization**
    Uncertainty sets are considered for the demand.
  - **Stochastic Programming**
    Uncertainty is described by some probability distribution.

Demand changes

Districting under uncertainty

- $\surd$ Demand is uncertain—cannot be forecasted accurately;;
- $\surd$ A proactive posture is assumed;
- $\surd$ Different paradigms can be considered:
  - Online optimization
    Typically for short term uncertainty.
  - Robust optimization
    Uncertainty sets are considered for the demand.
  - Stochastic Programming
    Uncertainty is described by some probability distribution.

    The case we consider hereafter!

Outline

Assumptions

We seek to plan for the organization of a geographical area into districts.

- √ A here-and-now decision has to be made about the districts;
- √ Demand is stochastic;
  We denote the corresponding random vector by $\boldsymbol{\xi} = [d_1, \ldots, d_{|I|}]$.
- √ The CDF of $\boldsymbol{\xi}$ is known (e.g. estimated using historical data).

Decisions

Here-and-now decision:

a territory design

feasible?

Under demand uncertainty:

- looking for a solution that is feasible for all the possible future observations of the demand may be impossible...

  $\downarrow$

  Non-balanced solutions!

- Even if it is possible, it may render a solution too much “fatness”.
- What can we do?

Decisions

Here-and-now decision:

a territory design

$\longrightarrow$

demand becomes known

$\longrightarrow$

Some recourse action is implemented:

A second-stage decision is made.

Outline

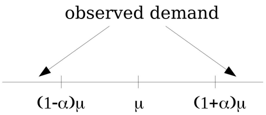
Unbalanced first-stage solutions

After demand is observed we may conclude that the initial solution is infeasible.

$\downarrow$

Unbalanced solution.

$\downarrow$



Extraordinary actions may be taken to overcome shortage or surplus at the districts (w.r.t. the established threshold):

- $\surd$ Handling possible shortage at some districts;
  e.g. downsizing work force.
- $\surd$ Handling possible surplus at some districts;
  e.g. outsourcing.

Additional notation

We need extra variables and extra cost parameters:

ψ_j demand surplus in district $j \in I$ w.r.t. the maximum threshold.

φ_j demand shortage in district $j \in I$ w.r.t. the minimum threshold.

g_j unit cost in district $j \in I$ for demand above the maximum threshold.

h_j unit cost in district $j \in I$ for demand below the minimum threshold.

Two-stage stochastic programming model

$$
\begin{aligned}
\min \quad & \sum_{i \in I} \sum_{j \in I} c_{ij} x_{ij} + \mathcal{Q}(\mathbf{x}), \\
\text{s. t.} \quad & \sum_{j \in I} x_{ij} = 1, && \forall i \in I, \\
& \sum_{j \in I} x_{jj} = p, \\
& x_{ij} \leq x_{jj}, && \forall i, j \in I, \\
& x_{ij} \in \{0, 1\}, && \forall i, j \in I,
\end{aligned}
$$

with

$$\mathcal{Q}(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\xi}}[Q(\mathbf{x}, \boldsymbol{\xi})].$$

Two-stage stochastic programming model

$$
\begin{aligned}
Q(\mathbf{x}, \boldsymbol{\xi}) = \min \quad & \sum_{j \in I} g_j \psi_j(\xi) + \sum_{j \in I} h_j \varphi_j(\xi), \\
\text{s. t.} \quad & (1-\alpha)\,\mu\, x_{jj} \leq \sum_{i \in I} d_i(\xi) x_{ij} + \varphi_j(\xi) - \psi_j(\xi) \leq (1+\alpha)\,\mu\, x_{jj}, \\
& \qquad\qquad\qquad \forall j \in I, \\
& \psi_j(\xi) \geq 0, \qquad \forall j \in I, \\
& \varphi_j(\xi) \geq 0, \qquad \forall j \in I.
\end{aligned}
$$

If the support of the random vector $\boldsymbol{\xi}$, say $\boldsymbol{\Xi}$, is finite than we can go further in terms of modeling.

We call scenario to a realization of the random vector $\boldsymbol{\xi} = [d_1, \ldots, d_{|I|}]$.

If $\boldsymbol{\Xi}$ is finite we can index the different scenarios in a finite set, say $S = \{1, \ldots, |\boldsymbol{\Xi}|\}$.

Extensive form of the deterministic equivalent

We can now index in S the demand and the second-stage decision variables:

d_{is}, demand of TU $i \in I$ under scenario $s \in S$.

ψ_{js}, demand surplus in district $j \in I$ under scenario $s \in S$.

φ_{js}, demand shortage in district $j \in I$ under scenario $s \in S$.

Additionally,

π_s, probability associated with scenario $s \in S$.

$$\pi_s \geq 0 \; (s \in S) \quad \text{and} \quad \sum_{s \in S} \pi_s = 1.$$

$\hat{\mu}$, reference value to be used in the balancing constraints.

$$\hat{\mu} = \frac{1}{p} \sum_{s \in S} \left(\pi_s \sum_{i \in I} d_{is} \right).$$

Two-stage stochastic programming model

$$
\begin{aligned}
\min \quad & \sum_{i \in I} \sum_{j \in I} c_{ij} x_{ij} + \sum_{s \in S} \left[\pi_s \sum_{j \in I} \left(g_j \psi_{js} + h_j \varphi_{js} \right) \right], && \\
\text{s. t.} \quad & \sum_{j \in I} x_{ij} = 1, && \forall i \in I, \\
& \sum_{j \in I} x_{jj} = p, && \\
& x_{ij} \leq x_{jj}, && \forall i, j \in I, \\
& (1-\alpha)\, \hat{\mu}\, x_{jj} \leq \sum_{i \in I} d_{is} x_{ij} + \varphi_{js} - \psi_{js} \leq (1+\alpha)\, \hat{\mu}\, x_{jj}, && \forall j \in I,\ s \in S, \\
& x_{ij} \in \{0,1\}, && \forall i, j \in I, \\
& \psi_{js} \geq 0, && \forall j \in I,\ s \in S, \\
& \varphi_{js} \geq 0, && \forall j \in I,\ s \in S.
\end{aligned}
$$

Outline

Additional second-stage decisions

The set of second-stage decisions can be enriched.

In addition to handling possible shortages and surplus at the districts we may become more proactive.

$\downarrow$

$\surd$ Move TUs across districts.

Decisions

Here-and-now decision:

a territory design

Second-stage decisions:

Shortages / Surplus / Re-districting

$\longrightarrow$

demand becomes known

$\longrightarrow$

Additional notation

We keep assuming uncertainty captured by a finite set of scenarios.

Additional notation:

r_{ijs}, cost for re-assigning TU i to TU j under scenario s, $\forall i, j \in I$, $s \in S$.

$$w_{ijs} = \begin{cases} 1, & \text{if TU } i \text{ is assigned to district } j \text{ under scenario } s, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i, j \in I,\ s \in S.$$

$$v_{ijs} = \begin{cases} 1, & \text{if } w_{ijs} = 1 \text{ and } x_{ij} = 0, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i, j \in I,\ s \in S.$$

Enriched model

$$\min \quad \sum_{i\in I}\sum_{j\in I} c_{ij}x_{ij} + \sum_{s\in S}\pi_s \left(\sum_{i\in I}\sum_{i\in I} r_{ijs}v_{ijs} + \sum_{j\in I}\left(g_j\psi_{js} + h_j\varphi_{js}\right)\right),$$

$$\text{s. t.} \quad \sum_{j\in I} x_{ij} = 1, \qquad \forall i \in I,$$

$$\sum_{j\in I} x_{jj} = p,$$

$$x_{ij} \le x_{jj}, \qquad \forall i,j \in I,$$

$$(1-\alpha)\,\hat{\mu}\,x_{jj} \le \sum_{i\in I} d_{is}w_{ijs} + \varphi_{js} - \psi_{js} \le (1+\alpha)\,\hat{\mu}\,x_{jj}, \qquad \forall j \in I,$$

$$x_{ij} \in \{0,1\}, \qquad \forall i,j \in I,$$

$$\psi_{js} \ge 0, \qquad \forall j \in I,\ s \in S,$$

$$\varphi_{js} \ge 0, \qquad \forall j \in I,\ s \in S,$$

Enriched model

$$\min \quad \sum_{i\in I}\sum_{j\in I} c_{ij}x_{ij} + \sum_{s\in S}\pi_s\left(\sum_{i\in I}\sum_{i\in I} r_{ijs}v_{ijs} + \sum_{j\in I}\left(g_j\psi_{js} + h_j\varphi_{js}\right)\right),$$

$$\text{s. t.} \quad \sum_{j\in I} w_{ijs} = 1, \qquad i\in I,\ s\in S,$$

$$\sum_{l\neq j} w_{jls} \le 1 - x_{jj}, \qquad j\in I,\ s\in S,$$

$$v_{ijs} \ge w_{ijs} - x_{ij}, \qquad i,j\in I,\ s\in S, \qquad \checkmark \quad v_{ijs} \ge \max\{0, w_{ijs} - x_{ij}\}$$

$$v_{ijs} \ge 0, \qquad i,j\in I,\ s\in S, \qquad \checkmark \quad v^*_{ijs} = \max\{0, w^*_{ijs} - x^*_{ij}\}$$

$$w_{ijs} \in \{0,1\}, \qquad i,j\in I,\ s\in S.$$

Outline

Some extensions of practical relevance:

- Maximum solution dispersion;
- Limitation in the number of reassignments in the second stage;
- Similarity w.r.t. the initial districting.

Outline

Dispersion

Territory dispersion indicates the maximum distance between a TU and the center of the district it is assigned to.

Suppose that there is a maximum desirable dispersion, say $l_{\max}$, for the districting solutions being devised.

$\downarrow$

For the first-stage solution this can be easily ensured:

$x_{ij} \leftarrow 0$, for every $i, j \in I$ such that $\ell_{ij} > l_{\max}$.

Dispersion

A maximum value for the dispersion may also be imposed for the second-stage solution.

For every $i, j \in I$ such that $\ell_{ij} > l_{\max}$ and for every $s \in S$ we must set:

$$x_{ij} \leftarrow 0,$$

$$w_{ijs} \leftarrow 0,$$

$$v_{ijs} \leftarrow 0.$$

For every instance of the problem there is a maximum absolute dispersion that makes sense to impose: the one resulting from solving the stochastic model.

Outline

Number of reassignments

In a stochastic setting, adapting a solution to the occurring scenarios may disfavor compactness and thus contiguity.

Natural extension to the problem:

$\downarrow$

Imposing a limit on the number of reassignments in the second stage.

m, maximum number of allowed reassignments.

Additional constraint:

$$\sum_{i \in I} \sum_{j \in I} v_{ijs} \leq m, \ \forall s \in S.$$

Outline

Similarity w.r.t. an initial plan

In some applications, it may be relevant to guarantee a certain degree of "similarity" w.r.t. some territory organization already planned.

In a stochastic setting we can consider similarity constraints associated with the second stage.

$\downarrow$

Ensure that we redesign districts keeping a certain degree of similarity w.r.t. the first-stage districting plan.:

$$\sum_{i \in I} x_{ij} w_{ijs} \geq \gamma \sum_{i \in I} x_{ij}, \quad \forall j \in I,\ s \in S.$$

$\gamma \in [0, 1]$, minimum proportion of TUs to keep together in the same district according to the first-stage plan.

Similarity w.r.t. an initial plan

Additional constraints:

$$\sum_{i\in I} x_{ij} w_{ijs} \geq \gamma \sum_{i\in I} x_{ij}, \quad \forall j \in I,\, s \in S.$$

Linearizarion:

$$v_{ijs} \leq w_{ijs}, \quad \forall i, j \in I,\, s \in S.$$

$$x_{ij} + v_{ijs} \leq 1, \quad \forall i, j \in I,\, s \in S.$$

The non-linear constraints can now be reformulated as:

$$\sum_{i\in I} \left(w_{ijs} - v_{ijs}\right) \geq \gamma \sum_{i\in I} x_{ij}, \quad \forall j \in I,\, s \in S.$$

Outline

Real geographical data

Province of Novara (Northern Italy)
88 municipalities.

√ Demand generated randomly.

Base demand: $d_i \sim U[1, 10]$.

Scenario 1:
20% below base demand.

Scenario 2:
base demand.

Scenario 3:
20% above base demand.

Real geographical data

Province of Novara (Northern Italy)
88 municipalities.

√ Four probability distributions:

Scenario	Probability distribution (k) 1	2	3	4
1	1/3	2/3	1/6	1/6
2	1/3	1/6	2/3	1/6
3	1/3	1/6	1/6	2/3

√ Assignment costs:

$$c_{ij}^k = \ell_{ij} \sum_{s \in S} \pi_{sk} d_{is}, \quad i, j \in I.$$

√ Reassignment costs:

$$r_{ijs} = \omega d_{is} \ell_{ij}, \quad i, j \in I,\ s \in S.$$

Outline

One stochastic solution: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

Stochastic model

Optimal solution—first stage:

Optimal solution—second stage?

We have three scenarios!

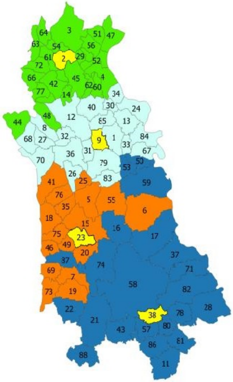
$\downarrow$

One second-stage solution for each scenario.

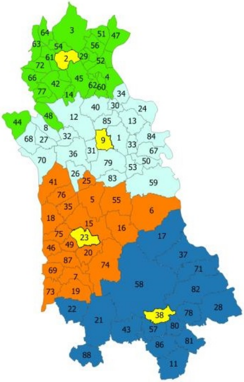
One stochastic solution: $p = 4,\ k = 3,\ \omega = 1,\ \alpha = 0.2$

Stochastic model—second stage solutions

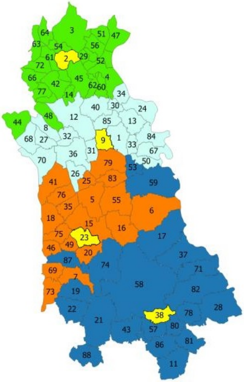
Scenario 1:



Scenario 2:



Scenario 3:



Optimal value of the stochastic model: obj(SP)=3463.20.

The Expected Value Problem: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

Expected Value (EV) Problem

Deterministic solution obtained by replacing the uncertain demand by their expected values.
($=$ scenario 2 in this example)

$\downarrow$

Territory design

$\surd$ Solution $\hat{x}_{ij}$, $i, j \in I$.

Is this a good solution for the stochastic problem?

The Expected Value Problem: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

(EEV) Expected cost (in the stochastic model) setting the solution obtained for the EV Problem

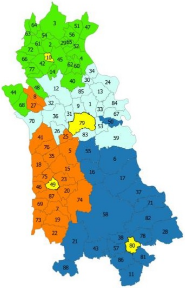
Take again the stochastic model but:

$\surd$ fixing the first-stage districting decision according to the values

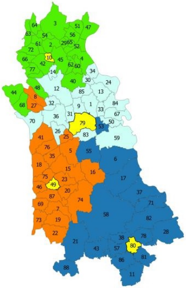
$$\hat{x}_{ij},\ i, j \in I;$$

$\surd$ Solving the model to find the w- and v-variables in all scenarios.

scenario 1



scenario 3



Obj(EEV)= 3595.15

but... Obj(SP)= 3463.20

Two initial districting solutions: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

A "deterministic" perspective:

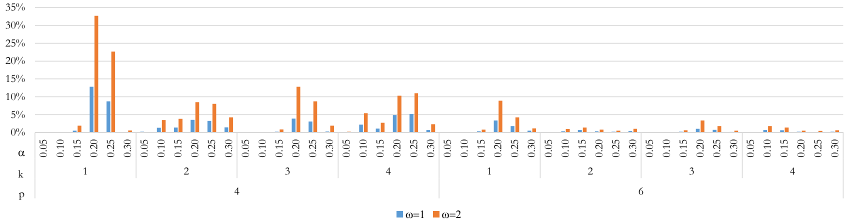
Final solution value: 3595.15

A "stochastic" perspective:

Final solution value: 3463.20

Additional results: $100 \times \mathrm{VSS}/\mathrm{SP}$

The value of the stochastic solution: $\mathrm{VSS} = \mathrm{EEV} - \mathrm{SP}$



Additional results: $100 \times$VSS/SP

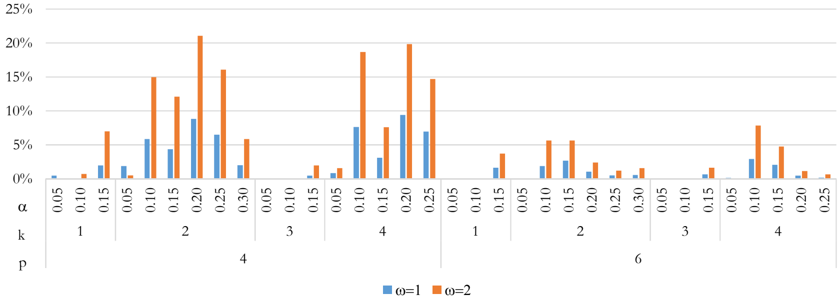
The value of the stochastic solution:

- The optimal solution to the expected value problem provides a better approximation for the stochastic problem when $\omega = 1$.
  - Relocations cost more when $\omega = 2$;
  - The expected value solution is not sensitive to ω since no relocations are made for a single-scenario problem.
- The VSS decreases with an increase in p—more districts are considered:
  - the expected value problem provides a better approximation to the stochastic problem $\rightarrow$ we typically observe fewer reassignments required.
- The uniform distribution seems to dominate the others.
- The intermediate values of α give raise to higher values of VSS:
  - for small α the expected value problem provides a good approximation to the problem since the penalty costs become dominant.
  - For larger α there's not much to rearrange in the solution.

Additional results: $100 \times \mathrm{CVSS}/\mathrm{SP}$

The compactness value of the stochastic solution (CVSS):

(Like VSS but ignoring the penalty costs for shortages and surplus.)



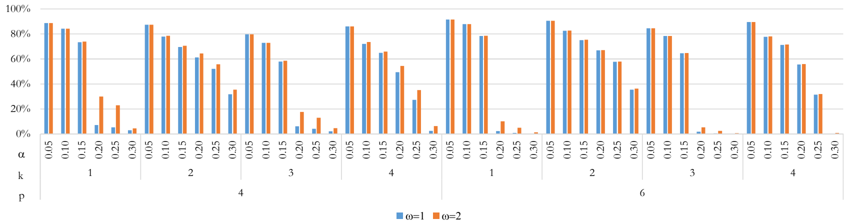
Additional results: $100\times$CVSS/SP

The compactness value of the stochastic solution (CVSS):

- It is important to compute because the penalty costs may blur the results.
- The VSS values are smaller in terms of compactness...
- ...but the expected value solution provides a poor approximation to the stochastic problem.

Additional results: $100 \times \text{EVPI}/\text{SP}$

Expected Value of the Perfect Information: $\text{EVPI} = \text{SP} - \text{WS}$



Additional results: $100 \times$EVPI/SP

Expected Value of the Perfect Information:

- Capturing uncertainty in our districting problem is of great relevance.
- The EVPI is rather insensitive to the number of districts.
- The EVPI shows a decreasing trend with α.

  The lower this parameter the higher the risk of observing demand surplus or shortages.

  $\downarrow$

  A decision maker is willing to pay a higher price to know perfect information about the future!

Performance of the stochastic model

So far we focused the results on a single set of geographical data.

$\downarrow$

88 municipalities in the province of Novara (Northern Italy).

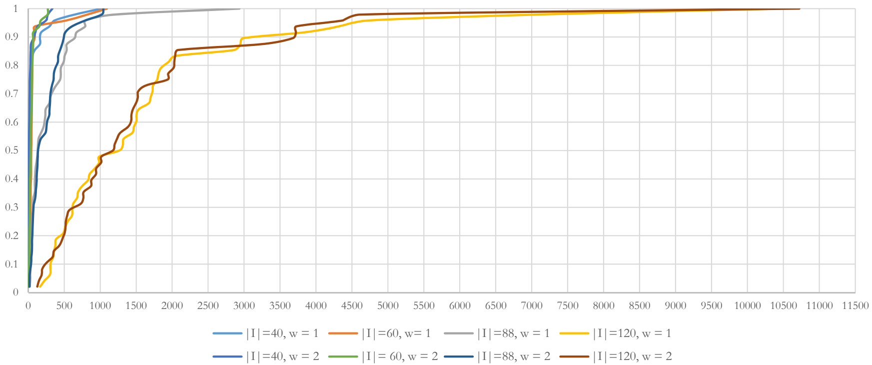
To assess the performance of the models we extended the instance set generating instances with 120, 60, and 40 TUs.

The stochastic model was solved using an off-the-shelf solver.

Additional results: computing time

Computing time (seconds):

(CPLEX, run on an Intel(R) Celeron(R) with 1.50GHz, 4GB RAM.)



Additional results: computing time

- Our stochastic districting problem can be solved up to proven optimality using tools that are available to most practitioners.
- There is still room for considering more comprehensive models.

Outline

Maximum dispersion: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

$$l_{\max} \leftarrow \lfloor \max_{i,j \in I,\, s \in S} \{\ell_{ij} \tilde{w}_{ijs}\} \rfloor = 31.$$

($\tilde{w}_{ijs}$: second-stage districting using the base stochastic model.)

$l_{\max}$	SP	EEV	$100 \times \frac{VSS}{SP}$	WS	$100 \times \frac{EVPI}{SP}$	CPU (sec.)
∞	3463.21	3597.21	3.87	3249.18	6.18	534.33
31.00	3463.99	3597.21	3.85	3249.18	6.20	300.44
27.00	3464.43	3601.97	3.97	3249.18	6.21	313.25
26.00	3464.91	3601.97	3.96	3249.18	6.23	401.59
25.00	3465.68	3601.97	3.93	3249.18	6.25	122.50
22.00	3474.10	3620.88	4.23	3249.18	6.63	293.56
18.00	3479.85	6731.63	93.45	3249.18	12.02	662.77
16.00	3693.27	7080.40	91.71	3249.18	50.08	222.99
15.00	6551.23	7099.49	8.37	3270.69	50.00	86.98
14.00	8890.62	The corresponding EEV is infeasible				39.08

Maximum number of reallocations: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

$$m_{\max} \leftarrow \max_{s\in S} \left\{ \sum_{i\in I} \sum_{j\in I} \tilde{v}_{ijs} \right\} = 7.$$

($\tilde{v}_{ijs}$: re-allocations using the base stochastic model.)

m	SP	EEV	$100 \times \frac{VSS}{SP}$	WS	$100 \times \frac{EVPI}{SP}$	CPU (sec.)	$\max_{s\in S} \left\{ \sum_{i,j\in I} v_{ijs} \right\}$
∞	3463.21	3597.21	3.87	3249.18	6.18	534.33	7
6	3464.00	3629.55	4.78	3249.18	6.20	519.21	5
4	3465.67	5332.71	53.87	3249.18	6.25	209.87	4
3	3469.52	6961.19	100.64	3249.18	6.35	236.44	3
2	3473.69	8625.23	148.30	3249.18	6.46	664.57	2
1	3480.25	10398.22	198.78	3249.18	6.64	461.14	1
0	3498.56	12175.12	248.00	3249.18	7.13	27.00	0

First row ($m = \infty$): our base solution.

Last row ($m = 0$): No reallocations allowed.

Similarity w.r.t. initial plan: $p = 4$, $k = 3$, $\omega = 1$, $\alpha = 0.2$

$$\gamma^* \leftarrow \min_{j \in I,\, s \in S} \left\{ \frac{\sum_{i \in I} \tilde{x}_{ij} \tilde{w}_{ijs}}{\sum_{i \in I} \tilde{x}_{ij}} \right\} = 0.84.$$

($\tilde{x}_{ij}$, $\tilde{w}_{ijs}$: values obtained using the base stochastic model.)

γ	SP	EEV	$100 \times \dfrac{VSS}{SP}$	WS	$100 \times \dfrac{EVPI}{SP}$	CPU (sec.)
0.00	3463.21	3597.21	3.87	3249.18	6.18	534.33
0.85	3464.00	3606.23	4.11	3249.18	6.20	516.58
0.90	3468.80	4184.58	20.63	3249.18	6.33	1405.86
0.95	3473.69	8718.12	150.98	3249.18	6.46	850.02
1.00	3498.56	12175.12	248.00	3249.18	7.13	2095.78

First row ($\gamma = 0.00$): our base stochastic solution.

Last row ($\gamma = 1.0$): full similarity (no reallocations allowed).

Outline

Overview of this seminar

- Basic concepts of districting problems were reviewed;
- Different ways for hedging against changing demand we discussed;
- A two-stage stochastic programming modeling framework was presented for capturing stochastic demand.
- The stochastic model was enriched by capturing additional features of practical relevance.
- The results of a series of computational tests performed using real geographical data were discussed.

Some ideas for further research

This is an area in which much work remains to be done!

- Multi-period variants of the problem;
- Stochastic models for risk-averse decision makers;
- Robust Optimization models;
- ...

References

E. Benzarti, E. Sahin, Y. Dallery. Operations management applied to home care services: Analysis of the districting problem. *Decision Support Systems*, 55:587–598, 2013.

B.K. Bergey, C.T. Ragsdale, M. Hoskote. A simulated annealing genetic algorithm for the electrical power districting problem. *Annals of Operations Research*, 121:33–55, 2003.

M. Blais, S.D. Lapierre, G. Laporte. Solving a home-care districting problem in an urban setting. *Journal of the Operational Research Society*, 54:1141–1147, 2003.

G. Bruno, E. Esposito, A. Genovese, C. Piccolo. Institutions and facility mergers in the italian education system: models and case studies. *Socio-Economic Planning Sciences*, 53:23–32, 2016.

G. Bruno and G. Laporte. An interactive decision support system for the design of rapid public transit networks. *INFOR: Information Systems and Operational Research*, 40:111–118, 2002.

F. Caro, T. Shirabe, M. Guignard, A. Weintraub. School redistricting: Embedding GIS tools with integer programming. *Journal of the Operational Research Society*, 55:836–849, 2004.

I. Correia, F. Saldanha-da-Gama. Facility location under uncertainty. G. Laporte, S. Nickel, F. Saldanha-da-Gama (editors) *Location Science*, Springer International Publishing, 2nd edition, 2019, Chapter 11, pp 185–213.

S.J. D'Amico, S.-J. Wang, R. Batta, C.M. Rump. A simulated annealing approach to police district design. *Computers & Operations Research*, 29:667–684, 2002.

L.S. de Assis, P.M. Franca, F.L. Usberti. A redistricting problem applied to meter reading in power distribution networks. *Computers & Operations Research*, 41:65–75, 2014.

References (cont.)

A. Diglio, S. Nickel, F. Saldanha-da-Gama. Towards a stochastic programming modeling framework for districting. *Annals of Operations Research*, 292:249–285, 2020.

A. Diglio, J. Peiró, C. Piccolo, F. Saldanha-da-Gama. Solutions for districting problems with chance-constrained balancing requirements. *Omega*, 103:102430, 2021.

J. Ferland and G. Guénette. Decision support system for the school districting problem. *Operations Research*, 38:15–21, 1990.

R. Garfinkel, G. Nemhauser. Districting by Implicit Enumeration Techniques. *Management Science*, 16(8):B495-B508, 1970.

S.W. Hess, J.B. Weaver, H.J. Siegfeldt, J.N. Whelan, P.A. Zitlau. Nonpartisan political redistricting by computer. *Operations Research*, 138:998–1006, 1965.

J. Kalcsics, R.Z. Ríos-Mercado. Districting Problems. G. Laporte, S. Nickel, F. Saldanha-da-Gama (editors) *Location Science*, Springer International Publishing, 2nd edition, 2019, Chapter 25, pp 705–743.

D. Konur, J. Geunes. Integrated districting, fleet composition, and inventory planning for a multi-retailer distribution system. *Annals of Operations Research*, 273:527–559, 2019.

S. Mostafayi Darmian, M. Fattahi, E. KeyvanshokoohS. An optimization-based approach for the healthcare districting under uncertainty. *Computers & Operations Research*, 135:105425, 2021.

M.C. Mourão, A.C. Nunes, C. Prins. Heuristic methods for the sectoring arc routing problem. *European Journal of Operational Research*, 196:856–868, 2009.

References (cont.)

F. Ricca, B. Simeone. Local search algorithms for political districting. *European Journal of Operational Research*, 189:1049–1426, 2008.

F. Ricca, A. Scozzari, B. Simeone. Political districting: from classical models to recent approaches. *Annals of Operations Research*, 204:271–299, 2013.

R.Z. Ríos-Mercado, J.F. López-Pérez. Commercial territory design planning with realignment and disjoint assignment requirements. *Omega*, 41:525–535, 2013.

R.Z. Ríos-Mercado, J.F. Bard. An exact algorithm for designing optimal districts in the collection of waste electric and electronic equipment through an improved reformulation. *European Journal of Operational Research*, 276:259–271, 2019.

O. Schoepfle and R. Church. A new network representation of a "classic" school districting problem. *Socio-Economic Planning Sciences*, 25:189–197, 1991.

P. Tavares, F. Figueira, J. Mousseau, V. Roy. Multiple criteria districting problems. The public transportation network pricing system of the Paris region. *Annals of Operations Research*, 154: 69–92, 2007.

S. Yanık, Ö. Sürer, B. Öztayşi. Designing sustainable energy regions using genetic algorithms and location-allocation approach. *Energy*, 97:161–172, 2016.

H. Zhong, R.W. Hall, M. Dessouky. Territory planning and vehicle dispatching with driver learning. *Omega*, 41:74–89, 2007.

A.A. Zoltners and P. Sinha. Sales territory design: thirty years of modeling and implementation. *Marketing Science*, 24:313–331, 2005.

Capturing Uncertainty in Districting Problems: Towards More Comprehensive Modeling Frameworks

Francisco Saldanha-da-Gama

Departmento de Estatística e Investigação Operacional,
Centro de Matemática Aplicações Fundamentais e Investigação Operacional,
Faculdade de Ciências, Universityof Lisbon, Portugal

Seminar presented at
The University of Hong Kong
Systems Analytics Global Leaders' Seminars
March 23, 2022

Thank you for your attention!