Strategic design of public transport networks, frequencies and bus sizes

Andrés Fielbaum

TU Delft

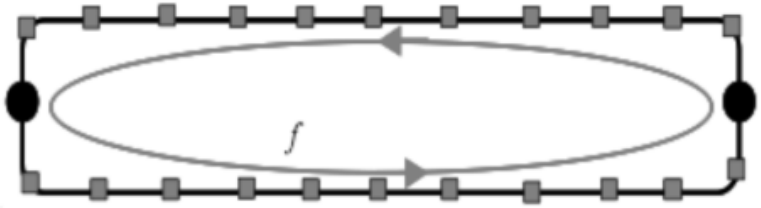
27/09/2021

Introduction

The basic single-line single-period model

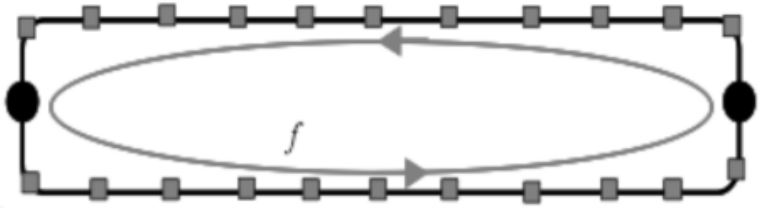
The basic model

In the classic single-line single-period model, Y passengers are distributed accross L which is completed in T, they all travel l and take t to board/alight.



The basic model

In the classic single-line single-period model, Y passengers are distributed accross L which is completed in T, they all travel l and take t to board/alight.



$$VRC = VRC_O + VRC_U = B(c_B + c_K K) + Yp_v\bar{t_v} + Yp_w\bar{t_w}$$

- B=fleet size, K=capacity of each vehicle, $\bar{t_v}, \bar{t_w}$ average in-vehicle and waiting times
- c_B, c_K, p_w, p_v exogenous cost-related parameters

Solutions of the simple model

Everything can be expressed as a function of the frequencies $VRC = Af + G/f + W$, yielding:

$$f^* = \sqrt{\frac{Y}{Tc_B}\left(\frac{p_w}{2} + tY\frac{l}{L}(p_v + c_K)\right)}, K^* = \frac{Yl}{Lf^*}$$

Solutions of the simple model

Everything can be expressed as a function of the frequencies $VRC = Af + G/f + W$, yielding:

$$f^* = \sqrt{\frac{Y}{Tc_B}\left(\frac{p_w}{2} + tY\frac{l}{L}(p_v + c_K)\right)}, K^* = \frac{Yl}{Lf^*}$$

- Economies of scale: waiting time decreases, fleet per passenger decreases.
- Diseconomies of scale: Buses get larger, inducing longer times at stops.

Economies of scale prevail; all these effects get exhausted as Y increases.

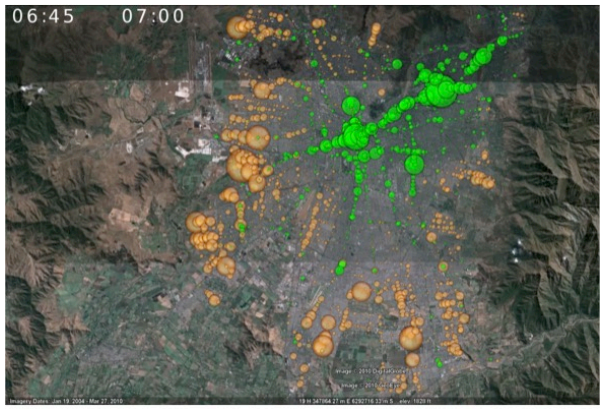
Relevance of the basic model

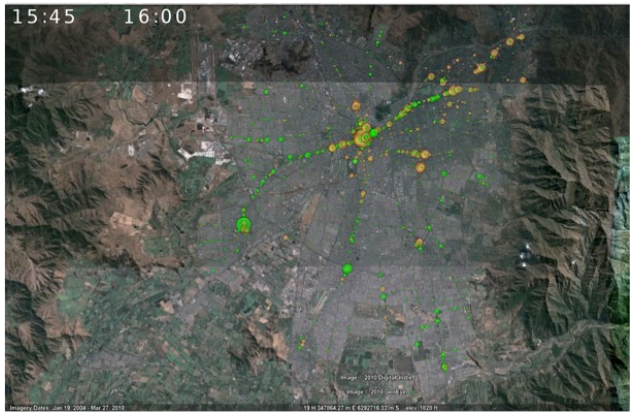
"By means of a simple bus line model it is possible to show that social cost minimisation results in a pattern of service characteristics which is **radically different from most present services**, mainly in these respects: given the demand, **more buses should be running, and the buses should be much smaller**." (Jansson, 1980).

Extension to two periods

A one-fleet system

Motivation





Source: Munizaga and Palma (2012)

Jansson (1984) tried to optimize the system under this scheme, but he could only do it by assuming equal frequencies accross periods $f_P = f_N$. Other authors have considered more than one period for other types of problems.

Posing the problem

We assume that Y_i, T_i and l_i depend on the period ($i = P, N$). Each period lasts E_i, and we distinguish between capital and operational costs. Now we have to minimize

Posing the problem

We assume that Y_i, T_i and l_i depend on the period ($i = P, N$). Each period lasts E_i, and we distinguish between capital and operational costs. Now we have to minimize

$$VRC_2 = \text{máx}(B_P, B_N)(c_{BC} + Kc_{KC}) + (B_P E_P + B_N E_N)(c_{BO} + Kc_{KO})$$

$$+ \sum_{i=P,N} E_i Y_i (p_v \bar{t_{vi}} + p_w \bar{t_{wi}})$$

$$s.t. \;\; \text{a)} K \geq \frac{Y_P l_P}{f_P L}, \text{b)} K \geq \frac{Y_N l_N}{f_N L}$$

Posing the problem

We assume that Y_i, T_i and l_i depend on the period ($i = P, N$). Each period lasts E_i, and we distinguish between capital and operational costs. Now we have to minimize

$$VRC_2 = \text{máx}(B_P, B_N)(c_{BC} + Kc_{KC}) + (B_P E_P + B_N E_N)(c_{BO} + Kc_{KO})$$

$$+ \sum_{i=P,N} E_i Y_i (p_v \bar{t_{vi}} + p_w \bar{t_{wi}})$$

$$s.t. \;\; \text{a)} K \geq \frac{Y_P l_P}{f_P L}, \text{b)} K \geq \frac{Y_N l_N}{f_N L}$$

Everything is written as a function of f_P, f_N. At least one restriction must be active. $\text{máx}(B_P, B_N) = B_P$.

Size given by the peak

Let us first analyze the case in which the peak restriction is active but the off-peak one is not: $K = \frac{Y_P l_P}{f_P L}$. This yields

$$VRC_2 = A_P f_P + A_N f_N + \frac{G_P}{f_P} + \frac{G_N}{f_N} + \delta \frac{f_N}{f_P} + U$$

$$(VRC = Af + \frac{G}{f} + W)$$

Size given by the peak

Let us first analyze the case in which the peak restriction is active but the off-peak one is not: $K = \frac{Y_P l_P}{f_P L}$. This yields

$$VRC_2 = A_P f_P + A_N f_N + \frac{G_P}{f_P} + \frac{G_N}{f_N} + \delta \frac{f_N}{f_P} + U$$

$$(VRC = Af + \frac{G}{f} + W)$$

F.O.C. $\Rightarrow$ equations of degree 5.

Size given by the peak

Let us first analyze the case in which the peak restriction is active but the off-peak one is not: $K = \frac{Y_P l_P}{f_P L}$. This yields

$$VRC_2 = A_P f_P + A_N f_N + \frac{G_P}{f_P} + \frac{G_N}{f_N} + \delta \frac{f_N}{f_P} + U$$

$$(VRC = Af + \frac{G}{f} + W)$$

F.O.C. $\Rightarrow$ equations of degree 5. But:

$$f_P^* = \sqrt{\frac{G_P + \delta f_N^*}{A_P}}, f_N^* = \sqrt{\frac{G_N}{A_P + \delta / f_P^*}}$$

$$\left(f_1^* = \sqrt{\frac{G}{A}} \right)$$

Analytical results

- f_P^* is larger than in the isolated case, because we want smaller buses as they also run at the off-peak.
- f_N^* might be larger or smaller than in the isolated case: no capital costs ($\uparrow$), but larger buses and frequency does not impact their size ($\downarrow$).

Analytical results

- f_P^* is larger than in the isolated case, because we want smaller buses as they also run at the off-peak.
- f_N^* might be larger or smaller than in the isolated case: no capital costs ($\uparrow$), but larger buses and frequency does not impact their size ($\downarrow$).

Crossed effects can also be proved:

- $\frac{\partial f_P^*}{\partial Y_N} > 0$, because vehicles' size becomes more important.
- $\frac{\partial f_N^*}{\partial Y_P} < 0$ because vehicles get larger.

Analytical results

- f_P^* is larger than in the isolated case, because we want smaller buses as they also run at the off-peak.
- f_N^* might be larger or smaller than in the isolated case: no capital costs ($\uparrow$), but larger buses and frequency does not impact their size ($\downarrow$).

Crossed effects can also be proved:

- $\frac{\partial f_P^*}{\partial Y_N} > 0$, because vehicles' size becomes more important.
- $\frac{\partial f_N^*}{\partial Y_P} < 0$ because vehicles get larger.

Buses full at the off-peak

We now analyze when buses are full at the off-peak (i.e. its constraint is active).

Buses full at the off-peak

We now analyze when buses are full at the off-peak (i.e. its constraint is active).

- A similar analysis with A_P, G_P, A_N, G_N shows that the peak constraint must be active as well: buses always run full at the peak.

Buses full at the off-peak

We now analyze when buses are full at the off-peak (i.e. its constraint is active).

- A similar analysis with A_P, G_P, A_N, G_N shows that the peak constraint must be active as well: buses always run full at the peak.
- Then $\frac{Y_P l_p}{f_P^* L} = K = \frac{Y_N l_N}{f_N^* L} \Rightarrow$ everything can be expressed as a function of f_N^* and explicit solutions can be found

$$f_N = \sqrt{\frac{tY_P c_{KC} Y_N \frac{l_N}{L} + \frac{p_w Y_N (E_N + E_P l_N / l_P)}{2} + [E_P Y_P + E_N Y_N] tY_N (l_N/L)(c_{KO} + p_v)}{T_N E_N c_{BO} + T_P c_{BC} Y_P l_P / (Y_N l_N) + T_P E_P c_{BO} Y_P l_P / (Y_N l_N)}}$$

Buses full at the off-peak

We now analyze when buses are full at the off-peak (i.e. its constraint is active).

- A similar analysis with A_P, G_P, A_N, G_N shows that the peak constraint must be active as well: buses always run full at the peak.
- Then $\frac{Y_P l_p}{f_P^* L} = K = \frac{Y_N l_N}{f_N^* L} \Rightarrow$ everything can be expressed as a function of f_N^* and explicit solutions can be found

$$f_N = \sqrt{\frac{tY_P c_{KC} Y_N \frac{l_N}{L} + \frac{p_w Y_N (E_N + E_P l_N / l_P)}{2} + \left[E_P Y_P + E_N Y_N\right] tY_N (l_N/L)(c_{KO} + p_v)}{T_N E_N c_{BO} + T_P c_{BC} Y_P l_P / (Y_N l_N) + T_P E_P c_{BO} Y_P l_P / (Y_N l_N)}}$$

- Analytical results depend on the value of the parameters.

Numerical results show that, regardless of the case:

Numerical results show that, regardless of the case:

- Peak frequency is larger and off-peak is smaller than the corresponding isolated periods cases, which means that optimal frequencies are very different, unlike Janson's assumption.

Numerical results show that, regardless of the case:

- Peak frequency is larger and off-peak is smaller than the corresponding isolated periods cases, which means that optimal frequencies are very different, unlike Janson's assumption.
- The capacity of the buses lies in-between the respective isolated ones.

Numerical results show that, regardless of the case:

- Peak frequency is larger and off-peak is smaller than the corresponding isolated periods cases, which means that optimal frequencies are very different, unlike Janson's assumption.
- The capacity of the buses lies in-between the respective isolated ones.
- The more similar the hourly flows, the larger both frequencies are.

Numerical results show that, regardless of the case:

- Peak frequency is larger and off-peak is smaller than the corresponding isolated periods cases, which means that optimal frequencies are very different, unlike Janson's assumption.
- The capacity of the buses lies in-between the respective isolated ones.
- The more similar the hourly flows, the larger both frequencies are.

Extension to two periods

A two-fleets system

The general idea

Which is the best way to design for a two-periods scheme? A reasonable strategy is allowing for two fleets of different size:

- Both running at both periods.

The general idea

Which is the best way to design for a two-periods scheme? A reasonable strategy is allowing for two fleets of different size:

- Both running at both periods.
- Each fleet runs at only one period.

The general idea

Which is the best way to design for a two-periods scheme? A reasonable strategy is allowing for two fleets of different size:

- Both running at both periods.
- Each fleet runs at only one period.
- ***Small* buses running alone at the off-peak, *large* buses complementing at the peak**.

The general idea

Which is the best way to design for a two-periods scheme? A reasonable strategy is allowing for two fleets of different size:

- Both running at both periods.
- Each fleet runs at only one period.
- ***Small* buses running alone at the off-peak, *large* buses complementing at the peak**.

Problem: at the peak, different buses will have different time at stops. We will use a *holding* strategy.

The holding strategy

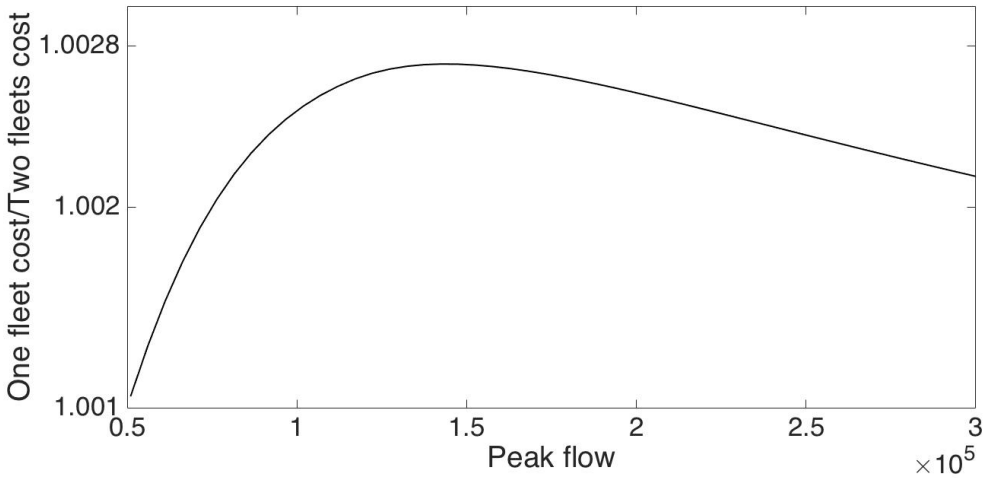
$$H = (K_L - K_S)\frac{L}{l_P} \cdot t$$

The equations that determine the system are found, and the system can be optimized (numerically).

Extension to two periods

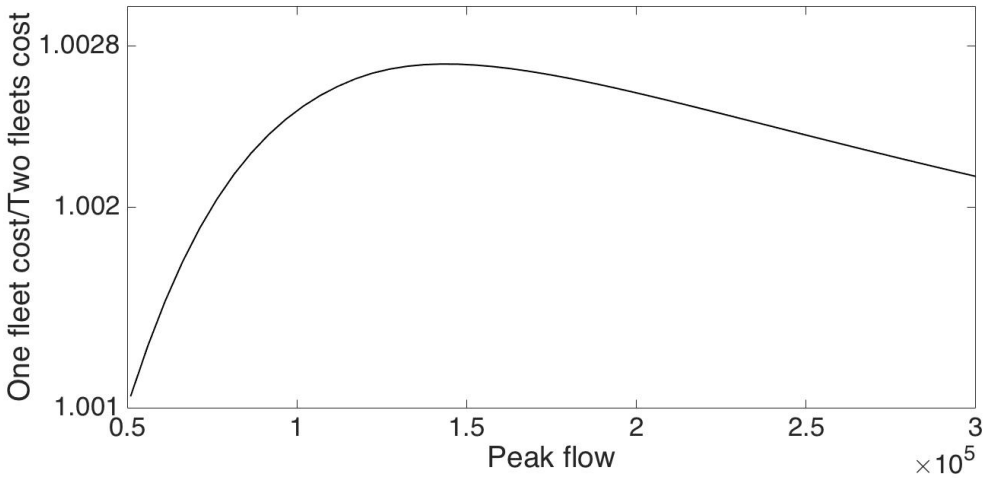
Comparison of these systems

General comparison



They have results that are too similar. Let us look into these results in more depth.

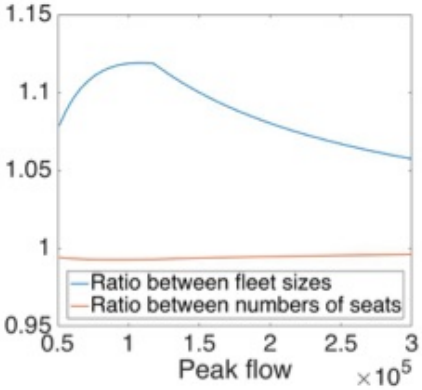
General comparison

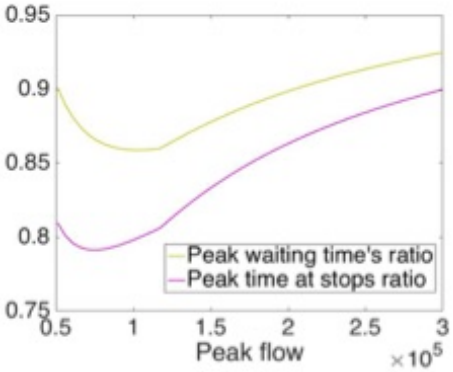


They have results that are too similar. Let us look into these results in more depth.

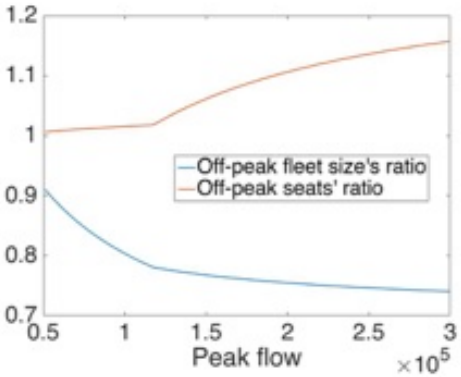
Results per period

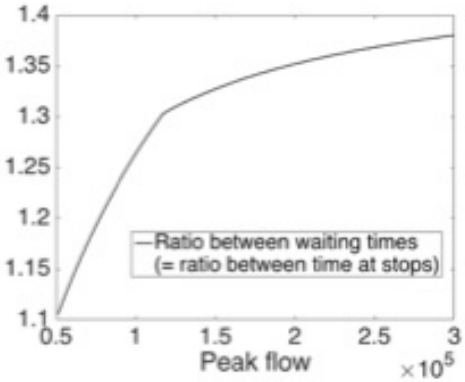
Peak:





Off-peak:





Peak users are favored by the one-fleet system, off-peak users by the two-fleets system.

Extension to a network

Comparing basic lines structures

Complexity of this extension

The extension towards a network is way more complex:

Complexity of this extension

The extension towards a network is way more complex:

- We need to find the optimal *lines structure*, i.e., the set of routes of the transit lines. Optimizing the lines structure, misconsidering frequencies, is already an NP-Hard problem.

Complexity of this extension

The extension towards a network is way more complex:

- We need to find the optimal *lines structure*, i.e., the set of routes of the transit lines. Optimizing the lines structure, misconsidering frequencies, is already an NP-Hard problem.
- Passengers can **choose routes**. These routes depend on the frequencies, which in turn need to be adjusted to carry every passenger.

Complexity of this extension

The extension towards a network is way more complex:

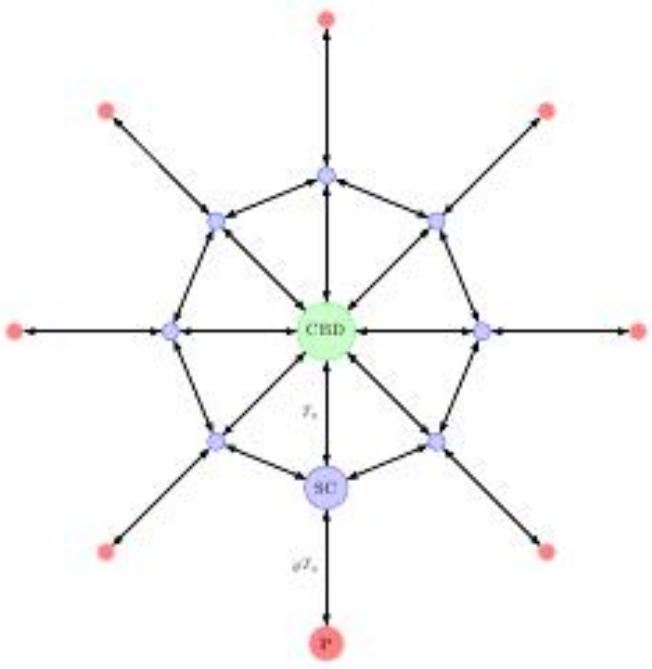
- We need to find the optimal *lines structure*, i.e., the set of routes of the transit lines. Optimizing the lines structure, misconsidering frequencies, is already an NP-Hard problem.
- Passengers can **choose routes**. These routes depend on the frequencies, which in turn need to be adjusted to carry every passenger.

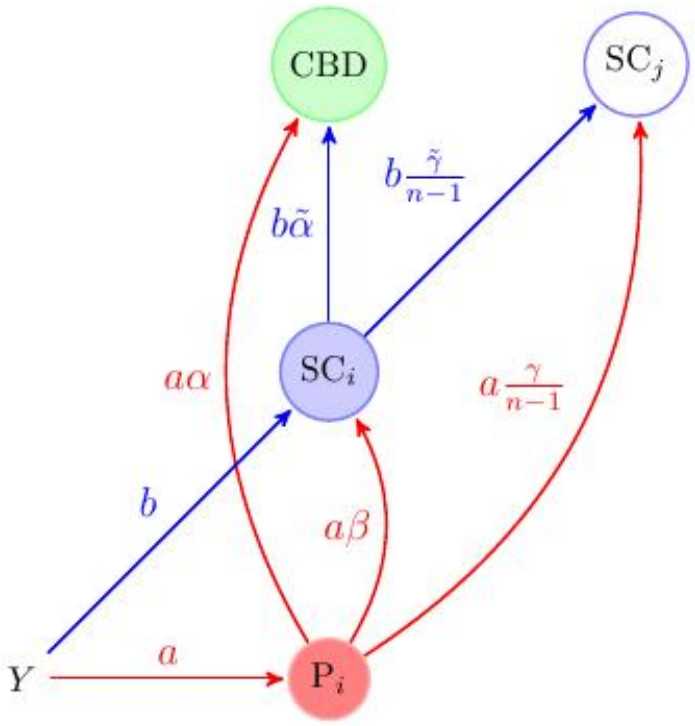
We are not solving the problem, but understanding it better. Fleet and bus capacity are now defined per line, and two spatial elements appear in the users' costs: access time and number of transfers.

$$VRC = \sum_L B_L(c_B + c_K K_L) + Y(p_v \bar{t_v} + p_w \bar{t_w} + p_a \bar{t_a} + p_T \bar{T})$$

The parametric city

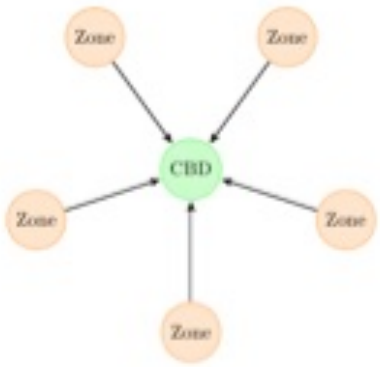
We analyze several aspects of the problem over this city model

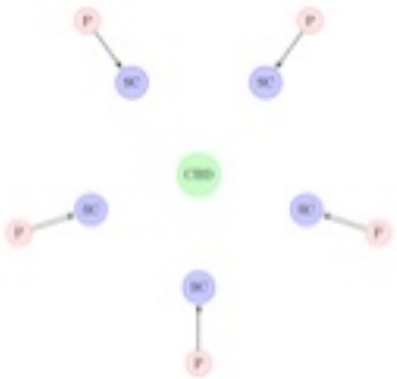


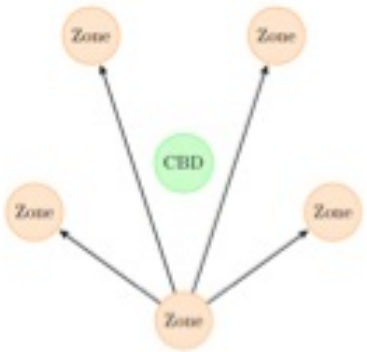


SC=Subcenter, P=Periphery.

Role of the parameters





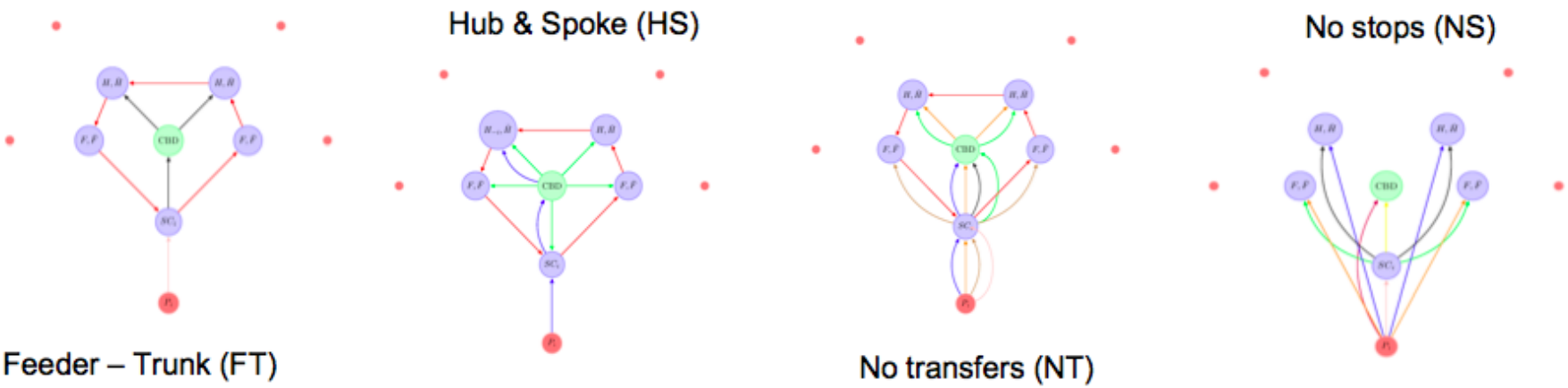


$\alpha \to 1$ Monocentric, $\beta \to 1$ Polycentric, $\gamma \to 1$ Dispersed

$$\alpha + \beta + \gamma = 1$$

Basic lines structures

Four lines structures over this city model.



Results of the basic structures

For each lines structure, and for each α, β, γ, Y:

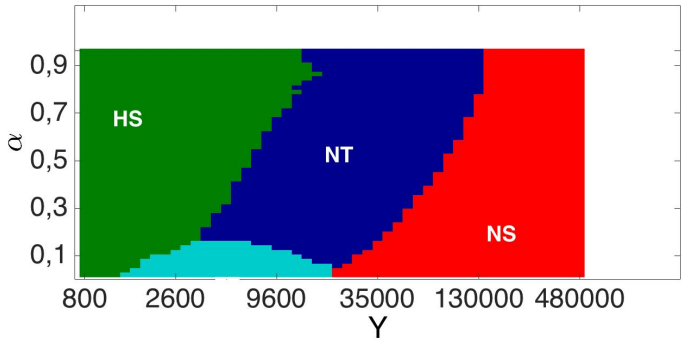
1. *VRC* is expressed as a function of the vector of frequencies.
2. The frequencies are optimized.
3. The resulting minimum *VRC* is obtained.

Results of the basic structures

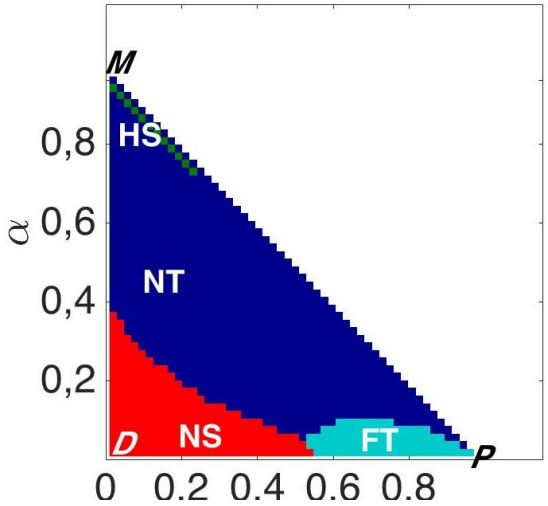
For each lines structure, and for each α, β, γ, Y:

1. *VRC* is expressed as a function of the vector of frequencies.
2. The frequencies are optimized.
3. The resulting minimum *VRC* is obtained.

We can identify the *best* lines structure for each OD pattern:



$\alpha - Y$ space ($\beta = \gamma$)



$\alpha - \beta$ space ($Y = 24000$)

Extension to a network

Scale economies and directness

The concept of directness

$$DSE = \frac{\text{Average costs}}{\text{Marginal costs}}$$

There are scale economies iff $DSE > 1$. DSE increases each time a lines structure changes... why?

The concept of directness

$$DSE = \frac{\text{Average costs}}{\text{Marginal costs}}$$

There are scale economies iff $DSE > 1$. DSE increases each time a lines structure changes... why?

Previous results have shown that, as the number of passengers increases:

- Passengers follow shorter routes.

The concept of directness

$$DSE = \frac{\text{Average costs}}{\text{Marginal costs}}$$

There are scale economies iff $DSE > 1$. DSE increases each time a lines structure changes... why?

Previous results have shown that, as the number of passengers increases:

- Passengers follow shorter routes.
- The number of transfers decreases.

The concept of directness

$$DSE = \frac{\text{Average costs}}{\text{Marginal costs}}$$

There are scale economies iff $DSE > 1$. DSE increases each time a lines structure changes... why?

Previous results have shown that, as the number of passengers increases:

- Passengers follow shorter routes.
- The number of transfers decreases.
- The number of stops decreases.

The concept of directness

$$DSE = \frac{\text{Average costs}}{\text{Marginal costs}}$$

There are scale economies iff $DSE > 1$. DSE increases each time a lines structure changes... why?

Previous results have shown that, as the number of passengers increases:

- Passengers follow shorter routes.
- The number of transfers decreases.
- The number of stops decreases.

This suggests the definition of the *Directness* of a lines structure, encompassing these three characteristics.

Directness over the parametric city

Structure	FT	HS	NT	NS
Transfers per trip	0.47	0.35	0	0
Stops per trip	3.06	3.06	3.06	2
Distance traveled/Min distance	1	1	1	1

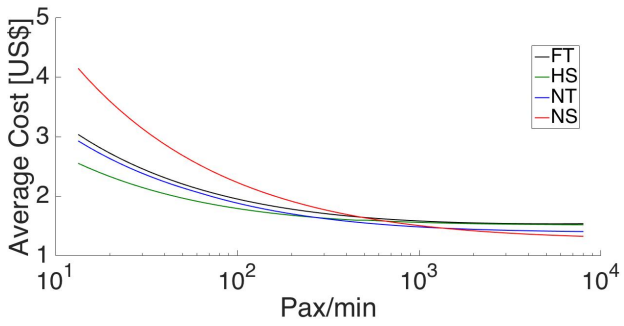
Directness over the parametric city

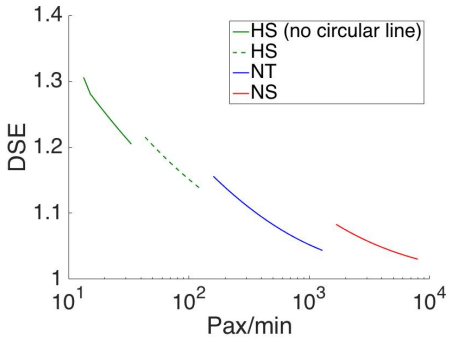
Structure	FT	HS	NT	NS
Transfers per trip	0.47	0.35	0	0
Stops per trip	3.06	3.06	3.06	2
Distance traveled/Min distance	1	1	1	1

These are indices before assigning passengers. Directness increases: FT $\rightarrow$ HS $\rightarrow$ NT $\rightarrow$ NS

DSE in the parametric city

$(\alpha = 0{,}5, \beta = \gamma = 0{,}25)$

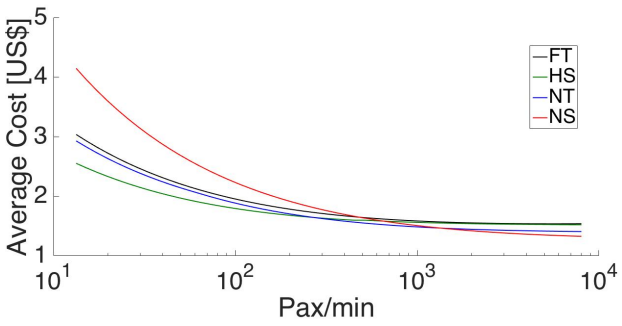


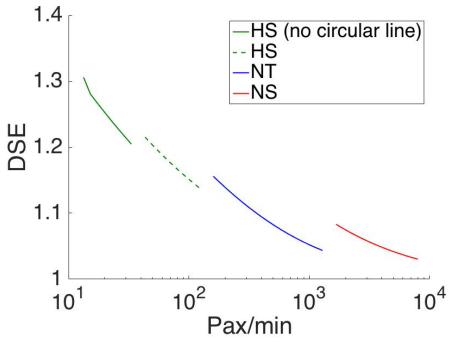


Best structure: HS→NT→NS: Directness increases!.

DSE in the parametric city

$(\alpha = 0{,}5, \beta = \gamma = 0{,}25)$

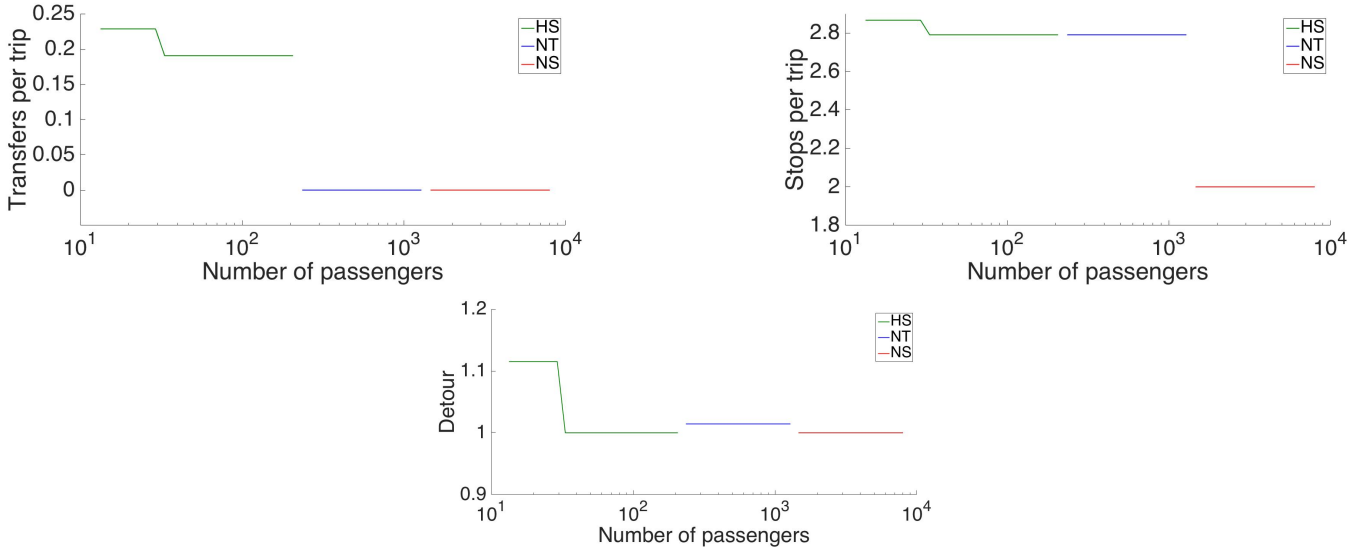




Best structure: HS→NT→NS: Directness increases!.

DSE curve: always larger than 1, it decreases to 1, jumping at each change in lines structure.

Evolution of the indices



Indices decrease (grossly speaking). There are trade-offs between the components of directness. In addition, each time the lines structure changes:

- Waiting time increases (diseconomies of scale).
- The number of seats decreases (idle capacity is reduced; economies of scale).

Conclusions

1. Considering temporal and spatial heterogeneity adds complexity to public transport models. Relevant qualitative effects can be identified.
2. Different strategies are possible to face the two-periods problem. In this research, we analyzed and compared two strategies: one fleet or two fleets.
3. In the one-fleet system, buses always run full at the peak. Vehicles' capacity lies in between the ones obtained by optimizing each period independently.
4. In the two-fleets system, a holding strategy is needed.
5. Both strategies yield similar total costs. The two-fleets system is better for off-peak users and worse for peak users.

Conclusions

6. Spatial exact optimization cannot be achieved; alternative approaches are needed.
7. We identified the best basic lines structure for different city and demand patterns.
8. We identified a novel source of economies of scale: the level of directness of the lines structure. As the number of passengers increases, routes become shorter, with less stops and less transfers. Scale economies get eventually exhausted.
9. *Spatial density is another source of scale economies. Total waiting costs = total walking costs. Increasing spatial density reduces the directness.*

References

- Jara-Díaz, S., Fielbaum, A., & Gschwender, A. (2017). Optimal fleet size, frequencies and vehicle capacities considering peak and off-peak periods in public transport. *Transportation Research Part A: Policy and Practice, 106*, 65-74.
- Jara-Díaz, S., Fielbaum, A., & Gschwender, A. (2020). Strategies for transit fleet design considering peak and off-peak periods using the single-line model. *Transportation Research Part B: Methodological, 142,* 1-18.
- Fielbaum, A., Jara-Diaz, S., & Gschwender, A. (2017). A parametric description of cities for the normative analysis of transport systems. *Networks and Spatial Economics, 17*(2), 343-365.
- Fielbaum, A., Jara-Diaz, S., & Gschwender, A. (2016). Optimal public transport networks for a general urban structure. *Transportation Research Part B: Methodological, 94*, 298-313.
- Fielbaum, A., Jara-Diaz, S., & Gschwender, A. (2020). Beyond the Mohring effect: Scale economies induced by transit lines structures design. *Economics of Transportation, 22,* 100163.
- Fielbaum, A., Jara-Diaz, S., & Gschwender, A. (2020). Lines spacing and scale economies in the strategic design of transit systems in a parametric city. *Research in Transportation Economics*, 100991.

Thank you for your attention!