DAC Migration Model

Hanno Schülldorf, Frederik Schäfer

Revision July 2024 [for discussion]

This is a technical paper intended for people working on implementing the DAC Migration DSS. The reader should be familiar with the concepts of DAC Migration which are NOT explained here. Except for chapter 1 some knowledge about mathematical modelling is necessary as well.

1 The DAC Migration Problem

1.1 User Perspective: Capabilities & Benefits

As of today the goal of migration is that 500 K wagons all over Europe will be migrated from screw couples (SC) to the digital automatic coupler (DAC) within 6 years.
This models keeps track of major resources and supplies and checks if the strategic requirements of migration can be met. Besides checking the feasibility of migration scenarios it helps setting up a more detailed retrofitting plan, reveals resource and supply bottlenecks as well as hard to meet strategic requirements. It even can be used to optimize migration in relation to various objectives.

The model is flexible and open to different strategies. Weather the intended migration process takes 6 month, 6 years or 26 years, the model will answer if it is feasible and how an optimal migration plan would look like. On the other hand the user is required to have some idea about the strategic requirements and the consequences of not archiving them. In the simplest case the requirement is that all wagons are going to be retrofitted by the end of migration. However the more preferences can be stated the more the plan computed by the model will match the actual needs of transportation industry.

1.2 Modelling Approach

In terms of modelling that means that there is a block of equations for enforcing the resource and supply restrictions.
The second block of equations ensures intertemporal consistency of vehicles and supplies. They can be thought of as bookkeeping making sure vehicles don't vanish or appear between time periods.

A third block enforces the strategic requirements, which means it controls how many vehicles have to be retrofitted until a certain point in time.
Concerning resources as well as strategic requirements the models is capable of soft limits. This means e.g. that instead of setting an exact number of vehicles which have to be retrofitted an optimal range as well as an larger acceptable range can be set.

1.3 Context

The model is intended to serve as core for the open-source Decision Support System "DAC Migration DSS". The DSS and hence the model are adopted to the questions to be answered by the DACFIT-Project workpackage 7 "Retrofit Planning". However several use cases beyond DACFIT have been identified.
The outline for a (larger) DAC Decision DSS was first described in the whitepaper "Decision Support System for DAC Migration" as result of a joint working group of DB Cargo and DB Analytics in 2023.

2 Linear Model Overview

2.1 Notation

Sets

Notation	Meaning
I	Set of vehicle groups
J	Set of areas (country, region, NUTS zone, rail yard, etc.)
T	Set of consecutive time periods $t_1 \dots t_n$
A	Set of possible retrofitting actions
U	Set of statuses
R	Set of resources
S	Set of supplies
E	Set of tiers

Variables

Notation	Meaning
$x_{i,j}^{t,a}$	Number of vehicles of group i on which action a is performed in area j within time period t (main decision variable)
$y_{i,t,u}$	Number of vehicles of group i that are in a certain status u at the beginning of time period t (bookkeeping)
$v_{t,s}$	Stock of supply s at the beginning of time period t (bookkeeping)
$w_{j,t}^{r,e}$	Extras capacity of a resource r used in area j in time period t from tier e
$d_{i,t,u,e}^{+/-}$	deviation (excess / missing) from target number of vehicle group i in status u for time period t by tier e

Properties

Notation	Meaning
$c(x)$	Basic cost for performing action a on one vehicle of vehicle group i in area j at time period t
$c(w)$	Cost of extra capacity for resource r in area j in time period t from tier e
$c(d)$	Cost of deviation from target number of vehicle group i in status u in time period t by tier e
$u_{start}(a)$	Transformation from status u by action a
$u_{end}(a)$	Transformation to status u by action a
$r(i,a,r)$	Consumption of resource r by action a done on a vehicle of group i
$s(i,a,s)$	Consumption of supply s by action a done on a vehicle of group i
$\phi_{i,t,u}$	Number of vehicles of group i entering or leaving the model at the beginning of time period t in status u
$\sigma_{t,s}$	Additional supply s delivered at the beginning of time period t
$P_{j,t,r}$	Basic amount of resource r available in area j in time period t
$\rho_{j,t,r,e}$	Extra amount of resource r available in area j in time period t by tier e
$\Delta_{i,t,u}$	Target number set for vehicles of vehicle group i at the beginning of time period t in status u
$\delta^{+/-}_{i,t,u,e}$	Admissible deviation (excess /missing) from target number

2.2 Constraints

Capacity & Supply Constraints

Resources:

$$\sum_{a\in A, i\in I} r(i,a,r)\cdot x^{t,a}_{i,j} \leq P_{j,t,r} + \sum_{e\in E} w^{r,e}_{j,t} \quad \forall j\in J, t\in T, r\in R \tag{1}$$

Extra Resource Limits:

$$w^{r,e}_{j,t} \leq \rho_{j,t,r,e} \quad \forall j\in J, t\in T, r\in R, e\in E \tag{2}$$

Supply:

$$\sum_{a\in A, i\in I, j\in J} s(i,a,s)\cdot x^{t,a}_{i,j} \leq v_{t,s} \quad \forall t\in T, s\in S \tag{3}$$

Bookkeeping

Vehicles:

$$y_{i,t,u} - \sum_{\substack{j\in J\\ a\in A^{out,u}}} x_{i,j}^{t,a} + \sum_{\substack{j\in J\\ a\in A^{in,u}}} x_{i,j}^{t,a} + \phi_{i,t,u} = y_{i,t+1,u} \quad \forall i \in I, t \in T, u \in U \tag{4}$$

Where $A^{out,u} = \{a \in A | u_{start(a)} = u\}$ and $A^{in,u} = \{a \in A | u_{end(a)} = u\}$

Supplies:

$$v_{t-1,s} - \sum_{a\in A, i\in I, j\in J} \cdot s(i,a,s) \cdot x_{i,j}^{t-1,a} + \sigma_{t,s} = v_{t,s} \quad \forall t \in T, s \in S \tag{5}$$

Strategic Requirements

Vehicles:

$$y_{i,t,u} = \Delta_{i,t,u} - \sum_{e\in E} d_{i,t,u,e}^{-} + \sum_{e\in E} d_{i,t,u,e}^{+} \quad \forall i \in I, t \in T, u \in U \tag{6}$$

Maximal Allowed Deviation:

$$d_{i,t,u,e}^{+/-} \leq \delta_{i,t,u,e}^{+/-} \quad \forall i \in I, t \in T, u \in U, e \in E, \{+,-\} \tag{7}$$

Non-Negativity

$$x, y, v, w, d \in R_0^+ \tag{8}$$

2.3 Objective Function

$$min \;\; z = \sum_{\substack{i\in I, j\in J\\ t\in T, a\in A}} c(x_{i,j}^{t,a}) \cdot x_{i,j}^{t,a} + \sum_{\substack{i\in I, t\in T,\\ u\in U, e\in E,\\ \{+,-\}}} c(d_{i,t,u,e}^{+,-}) \cdot d_{i,t,u,e}^{+/-} + \sum_{\substack{j\in J, t\in T\\ r\in R, a\in a}} c(w_{j,t}^{r,e}) \cdot w_{j,t}^{r,e} \tag{9}$$

3 Concepts of The Model

3.1 Indices and Variable x

The decision to be taken by migration planing is which vehicle to retrofit when and where. Subsequently the most important decision variable is $x_{i,j}^{t,a}$ which expresses how many vehicles of a vehicle group i in an area j are retrofitted by perfoming action a within a time period t.

3.1.1 Vehicle Group

We consider groups of vehicles instead of particular vehicles. The level of detail is flexible. For a first approach it might be sufficient to have very rough-cut vehicle groups:

$$I = \{locomotives, wagons, maintenance\_vehicles\}$$

As a general idea there should be as few vehicle groups as possible. This reduces the size of the model and makes it easier for the user to control input data and interpret output data. Furthermore data and business secret protection rules calls for a certain level of aggregation. However from a modelling point of view the following factors require to form separate groups:

- The retrofitting process differs, thus possible actions differ, resource consumption differs significantly or different supplies are needed. E.g. the intermodal wagon T-3000 is known to need some extra effort to be retrofitted since the installation space is lacking where as for other intermodal wagons this is is not a problem.
- Distinct strategic requirements are set.
- The vehicles can be retrofitted only in specific areas or the cost for bringing them to an other area would differ. E.g. they only operate in Greece and bringing them to Sweden for retrofit could cause higher costs compared to retrofitting them in Greece.

Just because two similar wagons are of different type of construction there is no reason to introduce separate groups as long as they are commercially fungible and technically equally retrofitable. The same applies to ownership which only matters if there are different strategic requirements, costs or retrofitting options and efforts.

3.1.2 Areas

An area is an aggregate of workshops and other locations where retrofitting actions are performed. Working on individual workshop level would not only blast the model, it also would reveal sensitive data of specific workshops. The reason for working with areas at all is that cost levels differ across Europe and wagons often operate only in specific areas thus bringing them to another area would cause transportation cost. For a first approach it seams pretty reasonable to define areas on basis of countries (NUTS-0-Level).

Certain attention require vehicles which are not homologated for public infrastructure. An specific area like "Industrial Railroad Duisburg" or an virtual area like "Mobile Workshops Germany" could be defined.

3.1.3 Time Periods

For a first approach time periods like 12 month or 6 month seem to be appropriate. Periods don't need to be equally long. E.g. it is possible to have an own period of 3 weeks for big bang.

3.2 Actions, Statuses and Variable y [eq. 4]

The fourth index a of variable x indicates which action is performed on the vehicle. In this context it is important to distinguish between "migration strategy", "action" and "status". An strategy requires certain actions to be taken at a certain time. Each vehicle has a status. An action a transforms a vehicle from it's initial status $u_{start}(a)$ to another status $u_{end}(a)$ consuming resources and supplies.
As of now for wagons four strategies are foreseen:

- Direct Retrofit: On wagons with status Screw Couplers (SC) the action "Full Retrofit" is performed. At the end of the action they are fully DAC equipped, thus their status is DAC.

- DAC Ready Strategy: This is a two step retrofit. First the wagons with status Screw Coupler (SC) are prepared for the DAC. At the end of this action they are in DAC Ready status (DR). This status is only technically relevant and has no commercial impact.
  The second action is to change the couplers which will lead to status DAC.

- Wagon Pairs: This strategy includes multiple actions. First only one side of the wagons is fully retrofitted ($FullRetrofit1/2$). The result is that one side of the wagon is equipped with screw couplers, the other with DAC (S/D). Send back to operations two wagons of this type remain together with screw couplers on their outside and DACs between them. At a certain point this is changed by shunting the wagons so the DAC is outside. Then for both wagons the remaining screw couplers are replaced by a full retrofit.
  From a modelling point this is a two action approach since the shunting happens outside the workshops and doesn't fit well with the resource model. For the model the states are limited to retrofit relevant states such as (S/D) and the number of wagons in this state.

- Swap Strategy: Wagons are fully retrofitted and used for certain commercial operations while others stay in their original status doing other commercial jobs. At a certain point in time the wagons are swapped between their commercial use. From a "action" and "status" modelling perspective this is not different to a "Direct Retrofit".

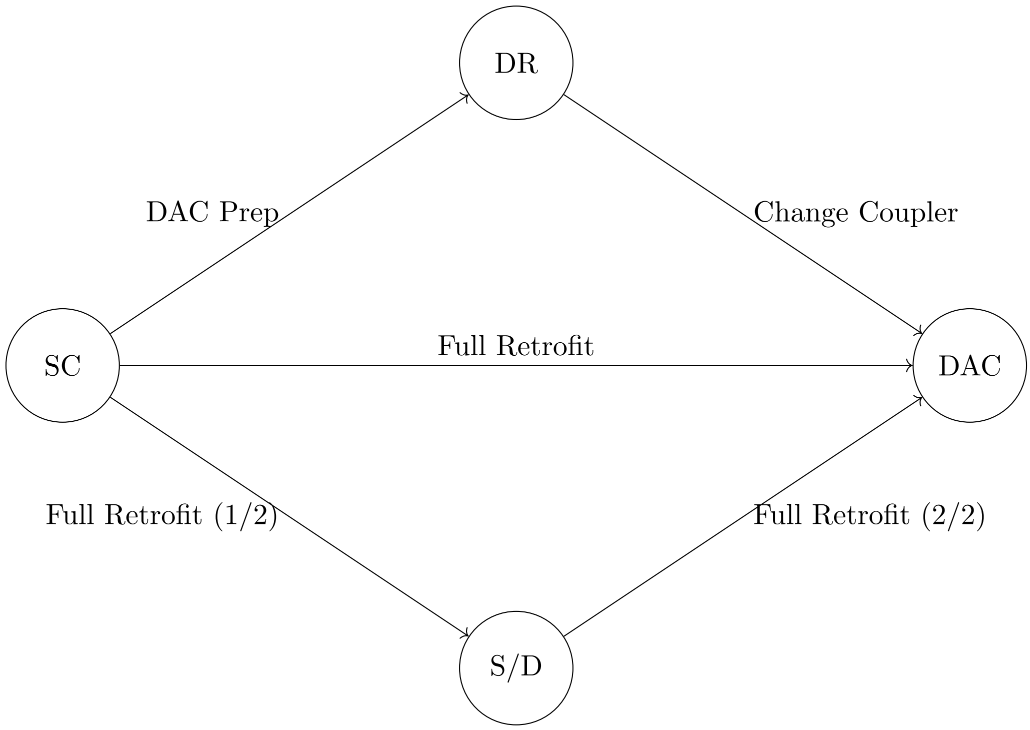

Figure 1: Retrofitting actions (edges) and statuses (nodes)

The described strategies are depicted in figure 1. Please note, that the abilities of the model are not limited to the aforementioned strategies.
The variable x correspondents with actions. Where as the number of vehicles of a certain group i in a certain status u at the beginning of a certain time period t is tracked by variable y.
Eq. 4 does the bookkeeping, it is explained in depth in table 1.

Please note besides the set of time periods T a period 0 to set initial values for the beginning of time period 1 and a time period $t+1$ to evaluated the state at the end of time period t are defined. Eq.4 has to hold for every period t, every vehicle type i and every status u.

3.3 Ressources [eq. 1, 2]

3.3.1 Basic Ressources

In the model it is distinguished between resources (Set R) and supplies (Set S). Resources can't be stored or transferred between areas and time periods. Therefore the available resources are index by time and area. The resource consumption of resource r to perform action a on one vehicle of group i is represented by the coefficient $r(i,a,r)$. A resource might be the capacity to perform a certain action on one vehicle. This is

	$y_{i,t,u}$	Number of vehicles in group i in status u at the start of time period t
-	$\sum_{\substack{j \in J \\ a \in A^{out,u}}} x_{i,j}^{t,a}$	Number of vehicles leaving status u due to an action starting at this status. The set $A^{out,u}$ denotes the actions a where $u_{start}(a) = u$
+	$\sum_{\substack{j \in J \\ a \in A^{in,u}}} x_{i,j}^{t,a}$	Vehicles entering this status as a result of an action are added
+	$\phi_{i,t,u}$	Vehicles entering or leaving ($\phi_{i,t,u} < 0$) the model (usefull for setting start values, deliveries and withdrawings decided outside the model)
=	$y_{i,t+1,u}$	Number of vehicles in this state this state at the end of time period respective at the beginning of the next time period $t+1$.

Table 1: Eq. 4 in detail

simple though limiting. In this case $r(i, a, r)$ would be 1 for all vehicles regardless of their group for one resource.
A more sophisticated approach would distinguish between available workstation hours, fitter work hours, electrician work hours, helper hours e.t.c. And for each vehicle group individual consumption of the resources can be modelled. This way also different actions can consume the same resources.

3.3.2 Extra Resources

There are two reasons for making resource limits not to hard:

- If limits are strict, the model will quickly return an infeasible solution giving no clue which resource actually was lacking. The option of a penalized overuse is suitable to detect the bottlenecks.
- In reality there is not an exact capacity of workshops in an area. E.g. extra shifts could be planed, the workshop could reject other jobs or additional workshops could be set up. Each of this measures causes over-proportional costs.

The approach is that several "tiers" of extra capacity are available in addition to the basic capacity of the resource $P_{j,t,r}$ on a per area, time period and resource level (see eq.1). Each tier's extra capacity is limited by $\rho_{j,t,r,e}$ (see eq. 2), its actual usage is tracked by $w_{j,t}^{r,e}$. The costs of the basic capacity use should be included in $c(x)$ while coefficient $c(w_{j,t}^{r,e})$ tracks the per unit extra costs that occur when using the extra capacity. The user should make sure that the per unit are increasing for each tier. Otherwise the model picks the cheapest extra capacity first regardless of tier numbering.

3.4 Supplies [eq. 3, 5]

Supplies can be transferred between time periods and areas. Therefore the variable v indicating the available stock is only indexed with supply s and time period t. Eq. 3 controls that never more supplies than in stock are consumed. The coefficient $s(i, a, s)$ indicates the amount of supply s needed to perform the action a. In 2 an example is shown. Different vehicle groups i may need different supplies.

Example: The supplies newly delivered each month are denoted by $\sigma_{t,s}$. Eq. 5 keeps

r("all wagons",a,s)	s_1:Draft Gear	s_2: Couplers
a_1: DAC Ready	2	0
a_2: Change couplers	0	2
a_3: Full Retrofit	2	2
a_4: Full Retrofit (1/2)	1	1
...	..	..

Table 2: Example how resources can be modelled

track of the stock of supplies which can be calculated by subtracting the consumption of the previous period from the stock at the beginning of the previous period and adding the newly delivered supplies.

3.5 Strategic Requirements and Associated Costs [eg. 6,7 and 9]

A Strategic requirement is a number or a range set for vehicles of group i which have to be in a certain status u at the beginning of time period t. The goal is set by $\Delta_{i,t,u}$ and the allowed excess (+) or underachievement ($-$) which may be defined in several tiers e is expressed by coefficient $\delta^{+/-}_{i,t,u,e}$. The usage of the allowed excess or underachievement is tracked by the variable by $d^{+/-}_{i,t,u,e}$ and may be free of or associated with costs $c(d)$.
The easiest goal is to require that all vehicles have to be retrofitted by the end of migration $(t + 1)$. In this case the model has the highest degree of freedom. A strict number can be set ($\delta_{i,t,u,e} = 0$ ...), but it has to be kept in mind that resources and supplies may be not sufficient. The result of this is that the model runs end with an infeasible solution. Therefore it is recommended to set a sufficiently large, highly penalized range instead.
As an idea different cases could be modelled as follows:

- Continuous retrofit strategy, 100 wagons of group "*bulk*" are used in a traffic which would benefit from DAC whenever possible. Due to status DAC 123€ could be saved per time period and wagon.

$$\Delta_{bulk,t,DAC} = 100; \delta^{-}_{bulk,t,DAC,1} = 100, c(d^{-}_{bulk,t,DAC,1}) = 123 \quad \forall t \in T$$

Ideal would be to have all wagons retrofitted at the beginning of period 1 which is of cause not realistic. Therefore a generous range is set associated with a penalty which should equal the lost benefits.

- Core network, 100 wagons of group "*cov*" to be retrofitted in big bang at the end of $t = 3$. Before big bang wagons with DAC can't be used causing a penalty of 4.321€ per wagon and period. After big bang it is other way round and wagons without DAC can't be used causing the same penalty.

$$\Delta_{cov,t,DAC} = 0; \delta^{+}_{cov,t,DAC,1} = 100, c(d^{+}_{cov,t,DAC,1}) = 4.321 \quad \forall t = \{1, 2, 3\}$$
$$\Delta_{cov,t,DAC} = 100; \delta^{-}_{cov,t,DAC,1} = 100, c(d^{-}_{cov,t,DAC,1}) = 4.321 \quad \forall t = \{4, ..., t\}$$

- Wagons in swap strategy, 50 wagons of group *lumb* are used in a traffic which can be retrofitted anytime causing no loses nor gains. Another 50 wagons are used in the core network. Those are not to be DAC equipped before big-bang at the end of $t = 3$. At big-bang time the wagons within the group *lumb* are swapped. After big-bang the wagons in the core network have to be DAC equipped. Failure to meet the goals will cause a penalty of 4.321€ per time period. At the end of migration all wagons have to be retrofitted.
  Before big bang:

$$\begin{aligned} \Delta_{lumb,t,DAC} = 50; \\ \delta^{+}_{lumb,t,DAC,1} = 50, c(d^{+}_{lumb,t,DAC,1}) = 4.321; \\ \delta^{-}_{lumb,t,DAC,1} = 50, c(d^{-}_{lumb,t,DAC,1}) = 0 \\ \forall t = \{1, 2, 3\} \end{aligned}$$

  After big bang:

$$\begin{aligned} \Delta_{lumb,t,DAC} = 50; \\ \delta^{+}_{lumb,t,DAC,1} = 50, c(d^{+}_{lumb,t,DAC,1}) = 0, \\ \delta^{-}_{lumb,t,DAC,1} = 50, c(d^{-}_{lumb,t,DAC,1}) = 4.321 \\ \forall t = \{4, ..., t\} \end{aligned}$$

  At the end of migration:

$$\Delta_{lumb,t,DAC} = 100; \delta^{-}_{lumb,t,DAC,1} = 100, c(d^{-}_{lumb,t,DAC,1}) = 4.321, \quad \forall t = \{t + 1\}$$

3.6 Costs c(x) [Eq. 9]

The objective of the model is to minimize the sum of costs of retrofit c(x), the cost of deviations from strategic requirements c(d) and extra capacity c(w). The model looks for a trade-off between retrofit costs, deviation costs and costs for extra capacity. So far the costs c(d) and c(w) have been explained. Now it is time to go more into detail and

$+$	Basic costs for retrofit action a in area j	
$+$	Average costs for bringing vehicle to workshop in area j	
$+$	Average costs for vehicle of group i to be out of service for the duration of retrofit and transportation to workshop in area j.	
$+$	political costs	
$=$	Costs $x_{ij}^{t,a}$	

Table 3: Example cost calculation

have also a look at the costs c(x).
It is important to note that the term cost has to be seen in a wider sense. The span ranges from costs directly paid to penalties introduced solely for modelling purposes. As far a costs are "real" it is important to ask to whom they are born. E.g. a grant given by some governmental institution will reduce the rail undertakings costs where as seen from a global perspective the costs remain the same - assuming that there is no impact on the decisions taken. As a hint, the model can also be used to some extend for game theory studies.
Back to the costs c(x). As pointed out there is no "right" way to model costs.

In table 3 possible elements of $c(x_{ij}^{t,a})$ are listed. The basic costs are the retrofit costs for a specific area which may vary between time periods t. They contain the costs that occur when workshops operate within their normal workload for one vehicle of group i. This means the resource and supply costs have to be included outside the model. Only the extra costs which are discussed in subsection 3.3.2 are included in c(w). Please not that the model can't handle fixed costs.
So far only workshops where attached to an area. The spacial component of vehicles is modelled through the costs for bringing vehicles to the workshop. In figures 3 there are several examples. E.g. traffic A is operating only in area 1 and traffic B operating in area 2 with identical vehicles. Despite they are identical they should be modelled as two groups e.g. $bulk_a$ and $bulk_b$. Since bringing an wagon of traffic A to area 2 is more expensive than bringing it to a workshop in area 1 the costs differ. Furthermore it will take longer and the wagon will be for longer time out of service. This could lead to costs as tabled in table 4.
As traffic C operates in both areas it might be possible to flexibly chose in which area the wagon is retrofitted thus the costs are equally low in both cases.
For D and E it depends on the circumstances. If the wagons from traffic E can be flexibly dispatched to traffic D and therefore can reach easily area 1 they could be modelled as one group with low costs in both areas. However it could be found to not be possible to dispatch the wagons flexibly. Than it is necessary to define separate vehicle groups with different costs.
In general the computation of costs can be very sophisticated. This should be done outside the model environment since it is a separate task.

	Area 1	Area 2
Traffic A	500 €	1.000 €
Traffic B	1.000 €	500 €
Traffic C	500€	500 €
Traffic D and E	600€	500 €
or alternatively		
Traffic D	500€	500 €
Traffic E	500€	1.000 €

Table 4: Example spacial costs

It is possible to define political costs. E.g. some vehicle owner doesn't want its vehicles to be retrofitted in another area than the preferred one. In this case all other areas could be penalized.
If the retrofit is not possible at all, then there are two options: To not define the decision variable $x_{i,j}^{t,a}$ at all or to define very high costs which are prohibitive.

3.7 Continuous Variables

It is recommended that the model is used with continuous variables (see eq. 9). This requires sufficiently large groups of vehicles in order to consider them as approximately continuous. The expected advantages of continuous over integer variables are the following:

- Lower computational efforts needed.
- Shadow prices may be interpreted and sensitivity analysis are easy to conduct.
- No fake precision, there is a lot of uncertainty and it would be an illusion anyway to have very realistic looking results.
- Having larger groups of vehicles well fits the political need to have sufficiently aggregate and anonymous data.

However the model is perfectly compatible with an integer requirement. It is easy to define the $x_{i,j}^{t,a}$ variables as integer. The extreme would be to allow only one vehicle group and make $x_{i,j}^{t,a}$ binary.

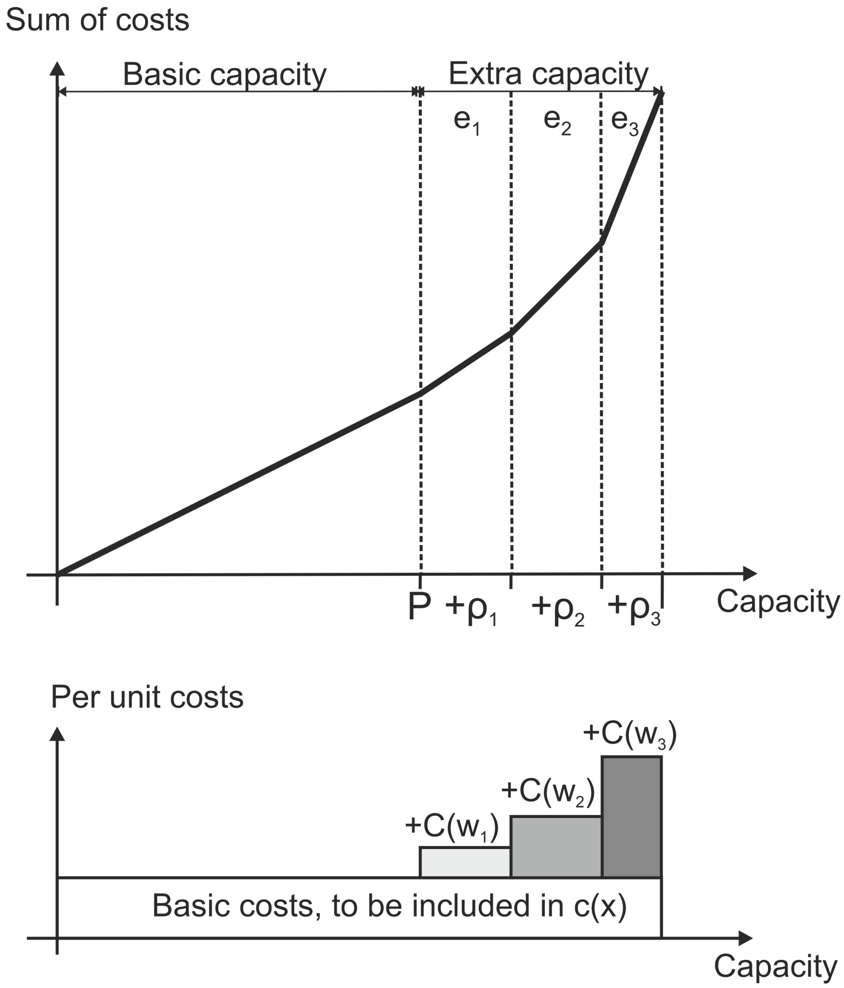

Figure 2: Extra capacity

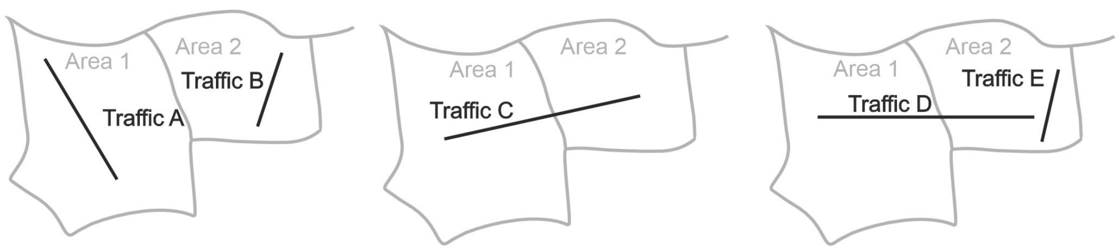

Figure 3: Spacial component of costs