# The Rolling Stock Recovery Problem Julie Jespersen Groth<sup> $\dagger^{*\alpha}$ </sup>, Jesper Larsen<sup> $\dagger^{\beta}$ </sup> and Jens Clausen<sup> $\dagger^{*\gamma}$ </sup>

*† DTU Management Engineering, The Technical University of Denmark, Produktionstorvet, DTU – Building 424, 2800 Kgs. Lyngby* 

\* DSB S-tog a/s, Produktionsplanlægning, Kalvebod Brygge, 32, 5., 1560 Copenhagen V.  $\alpha$ : Industrial Ph.D. Student,  $\beta$ : Associate Professor,  $\gamma$ : Professor and Chief Analysist

## Introduction

DSB S-tog (S-tog) operates on the double tracked, suburban network surrounding Copenhagen, Denmark. S-tog is the sole operator on the network. The network is owned and controlled by the infrastructure manager BaneDanmark.

During the last years there has been an increased focus on developing tools to aid the planning process in railway transportation. The tools are computer software, which can fully or partly automate some part of the planning process. As in other industries the initial focus has been on strategic, tactical and operational planning. Only lately focus has turned to the area of short term and real time planning. This paper concentrates on the area of rolling stock real time planning.

In practice rolling stock dispatchers monitor the operation of the rolling stock plan and the depot plans. When the rolling stock plan is disrupted, the rolling stock dispatcher makes real time decisions on the re-assignments of train units to train tasks. This process is called recovery. An automated tool will improve the recovery process, help supplying sufficient seat capacity for passengers and reduce the operating cost.

## Literature review

The research within the area of rolling stock schedule optimization has up to recently mainly focused on the planning phases prior to the day of operation. No emphasis has been on the area of real time rolling stock recovery. Most recent surveys on rail operation models are given by Cordeau et al. in 1998, [5], Törnquist in 2006, [12], and Huisman et al. in 2005, [8].

The problem of planning rolling stock can be divided into two sub problems: firstly, finding the compositions for each train task in the network and secondly, finding the paths for each virtual train unit ensuring depot feasibility and regular maintenance checks. The compositions indicate the type, number and order of train units assigned to a train task. The paths ensure that the train units are routed by the workshop at regular intervals. The first problem of finding compositions is widely explored. There is a distinction between the problems of allocating rolling stock when the fleet is composed by train units compared to when it is composed by train carriages and train locomotives.

Papers concerning the locomotive scheduling problem are e.g. [4], [9] and [3]. Papers concerning the problem where train units are self-propelled are [11], [2], [1], [10] and [6].

This paper addresses the area of real time rolling stock recovery. No prior research is available on this subject. We introduce a decomposed approach for the problem.

In the first section we give an introduction to the rolling stock plan. After that we define the disorder. In the next three sections we introduce the Train position model, the Train sequence model and the Train unit routing model respectively. Finally, in the two last sections we present some computational results and give a conclusion.

## The rolling stock plan

Train operations run according to a timetable consisting of terminal departures with predefined stopping patterns. Terminal departures are assembled in *Trains*, which is a virtual concept. Each train is represented by a set of *Train tasks* forming a *Train sequence*. The train tasks of a train sequence form a predefined work plan for the train in which each task except for the first and the last have a known predecessor and successor. This means that for two subsequent tasks  $t_1$  and  $t_2$ ,  $ArrivalTime_{t1} < DepartureTime_{t2}$  and  $ArrivalDepot_{t1} = DepartureDepot_{t2}$ . We will in the models, presented later in this paper, exploit the predecessor/successor relation between the tasks.

For each train task it is known which specific train units will cover it i.e. plans are pre-made. Given a disruption to the plan, we seek to return to the original plan as soon as possible.

A rolling stock plan consists of *Train unit routes* where each route covers a path of train tasks. These train tasks may or may not be consistent with a train sequence. The train path is feasible w.r.t. departure and arrival times of each of the tasks so that no two train tasks running simultaneously are covered by the same train unit. The train unit must be present at a task departure station to be able to cover that task.

Each train unit must submit to maintenance requirements. The maintenance constraint on driven distance between maintenance checks is maximum 22,000 km or 60 days.

When a train unit leaves or is added to a train sequence it is said to be decoupled from or coupled to the train i.e. a composition change occurs. There are two different train unit types. These can be coupled in all possible combinations limited by a maximum length of the train.

At S-tog (de)coupling always occurs at only one end of the train depending on the depot at which the (de)coupling occurs i.e. the train is only open for (de)coupling in one end. To decouple a train unit from a train, the train unit must be in the open end of the composition. When coupling a train unit to a train, the train unit must also be assigned to the open end of the train. The open versus the closed end of a composition at the terminal is illustrated in Fig. 1.





# Defining a disruption

In real time incidents occur that disturb the timetable. Some of these incidents are of such a size that also the rolling stock plan is disrupted. For a more detailed description of the effect disruptions have on the S-tog timetable see [7].

To minimize the impact of an incident, train route dispatchers monitoring the operation reroute trains and change the stopping patterns of trains to get operation back to normal as quickly as possible. A rolling stock plan is disrupted when train units are not able to cover the tasks they were expected to cover.

#### The disruption

It is likely that several delays demanding recovery occur during the day. Also, in the real-time situation, time is a critical factor and recovery decisions must be made fast. We will therefore

for each recovery instance solve within a specified time window e.g. two hours and include a limited set of train units. The start and end of the time window is the considered start and end time of the disruption. The tasks,  $t \in T$ , considered are those left uncovered, those which are insufficiently covered w.r.t. demand and those for which the assigned train units have been included in the disorder. For each task the start and end time,  $\tau_{d(t)}$  and  $\tau_{a(t)}$ , and start and end depot,  $\delta_{d[t]}$  and  $\delta_{a(t)}$ , are known. Each task is associated with a length in kilometers,  $Km_t$ , and duration measured in seconds,  $Time_t$ . The set of tasks having no predecessors constitutes  $T_0$ . The tasks having no successors likewise constitute  $T_1$ . The successor of task t is denoted v(t). Each task has a seat demand,  $Demand_t$ .

The set of depots involved in the disorder, D, is defined by the routes of the train lines included. For each depot,  $d \in D$ , included in the disorder the start capacity of each type of train unit *m* is given by  $DepotCap_{d,m}$ .

Typically, all train units,  $k \in K$ , assigned to the train lines of the affected train units are included plus train units located on the affected depots at the start time of disruption.

For each train unit a kilometer limit,  $KmLimit_k$ , is given. It indicates the maximum sum of kilometers of train tasks that a train unit can be assigned without violating maintenance constraints. Each train unit has a seat capacity matching its train type. For each train unit the start depot, *StartDepot\_k*, and a preferred end depot, *EndDepot\_k*, are given.

At all times two rolling stock types,  $m \in M$ , are available. These are short and long train units named SE and SA respectively.

We have for the models assumed that the maximum length of the composition assigned to a train is equivalent to the length of two SA train units. There are two positions,  $p \in P$  in a given order on each composition.

## Train unit position model

In this section we introduce the variables, objective and constraints of the Train Unit Position model (TUP). The main variables of the model are the integer *X*-variables assigning train unit types to the positions of each train task.

$$X_{tp}^{m} = \begin{cases} 1 & \text{If a train unit of type } m \text{ is assigned to task } t \text{ in position } p \\ 0 & \text{Otherwise} \end{cases}$$

From the X-variables the L-variables are derived. The  $L_t^m$  variables are inventory variables indicating the number of train units of type *m* saved at the departure depot of *t* immediately before the departure of *t*.

Finally,  $O_t^m$  and  $N_t^m$  are variables indicating whether respectively coupling and decoupling is carried out between the tasks *t* and *v*(*t*). Both sets of variables are binary.

 $L^0$  are the start inventory parameters.  $L_{dm}^0$  indicates the number of train units of type *m* located at depot *d* in the beginning of the disorder.  $L_{dm}^1$  are the end capacity variables indicating the number of train units of type *m* present at depot *d* in the end of the considered recovery period. A desired end depot capacity is given by the parameter  $E[cap]_d^m$ . The variables  $E_d^m$  indicate the lack of train units of type *m* in depot *d* in the end of the recovery period.

As the *D*-variables are calculated entirely from the *X*-variables they are always integer. Therefore,  $D_t^m \in R$  for all  $t \in T$ ,  $m \in M$ .

Also,  $L_t^m$  are found from  $L_{dm}^0$  and  $D_t^m$ . As both will be integer (*D*-variables given indirectly from the *X*-variables), the *L*-variables will automatically be integer. Therefore,  $L_t^m \in R_+$  for all  $t \in T$ ,  $m \in M$ .

The relevant aspects to include in the objective of the positioning model are: minimizing seat shortage, minimizing number of composition changes and minimizing the cost of covering

train tasks with train units. The expected seat shortage for each train task is the difference between the sum of the seats allocated to the train task and the demand of the task. The objective is to minimize the sum of the expected seat shortage over all tasks, the driven kilometers, the couplings and the number of train units not ending at the right depot, see Eq. 1.

$$OBJ = C^{1} \cdot \sum_{t \in T} \{Demand_{t} - \sum_{m \in M, p \in P} Seats^{m} \cdot X_{tp}^{m}\} + C^{2} \cdot \sum_{t \in T, p \in P, m \in M} Km_{t} \cdot X_{tp}^{m} + C^{3} \cdot \sum_{t \in T, m \in M} O_{t}^{m} + C^{5} \cdot \sum_{d \in D, m \in M} E_{d}^{m}$$
(1)

As a train has a maximum length each train task can not be covered by more than the maximum number of train units per train. This is guaranteed by Eq. 2.

$$MinTrainLength \leq \sum_{m \in M, p \in P} Length_m \cdot X_{tp}^m \leq MaxTrainLength, \quad \forall t \in T$$
 (2)

Physically only one train unit can be assigned to each position of a train task. Eq. 3 ensures this.

$$\sum_{m \in M} X_{tp}^m \le 1$$
,  $\forall t \in T, p \in P$  (3)

We control the incoming and outgoing flow of depots by three sets of inventory constraints, see Eq. 4 to 6. The first set of constraints controls that the initial inventory level is not violated. This means that for each depot *d* the tasks departing before the first arriving task must not use more capacity than what is present initially given by  $L_{dm}^0$ . The set of departing tasks before the first arrival task on depot *d* is given by *FirstTasks<sub>d</sub>* for all  $d \in D$ . See Eq. 4

$$\sum_{p \in P, t \in FirstTasks_d} X_{tp}^m \le L_{dm}^0, \quad \forall \quad d \in D, m \in M$$
(4)

The inventory in a depot of train unit type m immediately after the arrival of a task t is given by the start capacity on the depot minus the sum of every train unit of type m coupled to tasks at that depot before and including task t and plus the sum of every train unit decoupled from tasks at that depot before and including task t. This is handled by Eq. 5.

$$L_{t}^{m} = L_{a(t)m}^{0} - \sum_{\substack{p \in P, t' \in T \\ \tau_{d(t')} = \delta_{a(t)}}} X_{t'p}^{m} + \sum_{\substack{p \in P, t' \in T \\ \tau_{a(t')} \leq \tau_{a(t)}}} X_{t'p}^{m}, \quad \forall \quad t \in T, m \in M$$

$$(5)$$

The last set of inventory constraints concerns the end capacity. The end capacity,  $L_{dm}^{l}$ , of train unit type *m* in depot *d* is given by  $L_{t}^{m}$  for which task *t* is the last task arriving on *d*, *FinalTask<sub>d</sub>*. See Eq. 6.

$$L^1_{dm} = L^m_{FinalTask_2}, \quad \forall \quad d \in D, m \in M$$
 (6)

We wish to control the end depot balance by minimizing in the objective function the deficit of train units defined by variables  $E_d^m$ . These are defined in Eq. 7

$$E_d^m \ge E[cap]_d^m - L_{dm}^1, \qquad \forall \quad d \in D, m \in M$$

$$\tag{7}$$

Each depot has an individual upper capacity on the number of units that can be stored at that depot. The upper capacity is estimated by controlling the length of the rolling stock stored at each depot relative to the length of the depot tracks, *DepotCap<sub>d</sub>*. Eq. 8 controls the capacity of each depot right after the departure of each task, that is,  $\delta(t)$  is the departing depot of t.

$$0 \le \sum_{m} Length_m \cdot L_t^m \le DepotCap_{\delta(t)}, \quad \forall t \in T$$
(8)

The coupling and decoupling variables are found in Eq. 9 and 10.

$$M \cdot O_t^m \ge L_{\nu(t)}^m - L_t^m, \qquad \forall \quad t \in T \backslash T^1, m \in M$$
(9)

$$M \cdot N_t^m \ge L_t^m - L_{\nu(t)}^m$$
,  $\forall t \in T \setminus T^1, m \in M$  (10)

To ensure that no train unit is decoupled from a train if it is positioned in the closed end of the train composition, we include Eq. 11. In the equation we use a parameter  $After_{tp}$  to indicate whether a position is closed for decoupling after train task t and thereby if the position is still occupied in the successor task.

$$ChangePosition_{t} = \begin{cases} 1 & \text{If closed position of task } t \text{ is different than} \\ & \text{closed position of successor } \nu(t) \\ 0 & \text{Otherwise} \end{cases}$$

If 
$$ChangePosition_t = 1$$
  

$$X_{tp}^m \leq \sum_{p' \in P, p'! = p} X_{\nu(t)p'}^m + W_{\nu(t)}$$

$$X_{tp}^m \leq \sum_{p' \in P, p' = p} X_{\nu(t)p'}^m + V_{\nu(t)}$$
else
$$(11)$$

$$X_{tp}^m \le X_{\nu(t)p}^m$$

 $V_t$  and  $W_t$  are length indicator variables.  $V_t$  is one if one train unit is assigned to task t and zero otherwise for all  $t \in T$ .  $W_t$  is one if two train units are assigned to task t and zero otherwise for all *t C T*. They are found in Eq. 12 and 13.

$$W_t \le \sum_{p \in P, m \in M} X_{tp}^m - 1 \le 2 \cdot W_t, \qquad \forall \quad t \in T$$
(12)

$$V_t + W_t = 1$$
,  $\forall t \in T$  (13)

#### Solution strategy for Train unit position model

The model is implemented in C# using Concert Technology from ILOG to solve it in Cplex 10.0. Given the size of the problem we expect solutions to be achieved within acceptable computation times. The TUP model is the first step in an integrated process of finding a recovery plan. The joint computation time must be less than approx. 5 minutes. We therefore wish for the TUP model to find solutions within approx. 15-30 seconds.

## Train sequence model

When a train unit's path consists of one train sequence it is certain that the train unit is not decoupled or coupled at any time. That is, coupling and decoupling refer to train unit flows to and from the train sequence. The conducting of (de)coupling is time demanding and in a periodic timetable there will not necessarily be ample time for performing these. It is assumed that if the number of (de)couplings is decreased the robustness of the rolling stock plan is increased. That is, the rolling stock plan will be less sensitive towards minor interferences in the operation.

This section shortly describes the Train sequence model. The model tries to assign a single, physical train unit to each train sequence in the disruption, such that the train unit can feasibly cover the entire train sequence. In this way the model copies qualities that are always present in a good solution in practice.

Note that we in the Train sequence model handle the physical train units, that is, the Train sequence model is a multi commodity flow model where the commodities are individual train units which are each identified with individual maintenance requirements, time and place origins and destinations.

The consequence of only covering the set of train sequences with one train unit each is that for a set of train tasks the demand will not be sufficiently covered. For some of the train tasks one train unit will be assigned but the demand exceeds the seat capacity of the train unit. For some units no train unit will be assigned and the demand not covered at all. The set of train tasks not covered sufficiently will be addressed in a third model. There is a preference of which train unit type to assign in the process of assigning train units to train sequences. The most preferable train unit type is chosen given the results of the Train unit position model. Recall that this model gives information regarding number and type of train unit types assigned to each train task. For each train sequence the train unit type chosen as the most preferable coverage is the type being present on composition of the train tasks of the train sequence. If no train type exists present on all the tasks, no preference is awarded.

The great advantage of the train sequence model is both that it reduces the number of variables considerably for the third model (the routing model) and that it enforces some good, practical qualities in the solution.

Again the model is not of considerable size and we solve it using Cplex 10.0 and Concert Technology where the model is implemented in C#.

#### Train Routing model

As mentioned in the previous section the Train sequence model will only cover some of the train tasks according to their respective demands. Some will either be left uncovered or covered insufficiently according to demand. These must be covered by valid train task paths using the train units not yet assigned to a train path. This is done by the Train Routing Model.

The main variables of the Train Routing model are,  $q_t^k$ . These variables assign train units to train tasks.

$$q_t^k = \begin{cases} 1 & \text{If train unit } k \text{ is assigned to train task } t \\ 0 & \text{Otherwise} \end{cases}$$

To control the solutions of the model a second set of variables is included,  $\rho_t^k$ . The  $\rho_t^k$  variables are used to control the number of (de)couplings in the solution.

$$\rho_t^k = \begin{cases}
1 & \text{If train unit } k \text{ is assigned to train task } t \text{ and the successor of } t, \nu(t) \\
0 & \text{Otherwise}
\end{cases}$$

The objective function maximizes the total sum of covered demand and the sum of couples of consecutive tasks covered by the same train unit. The use of physical train units is included in the objective with some weight  $W_2$ , which will be less than zero. See Eq. 14.

$$Maximize \quad W_1 \cdot \sum_{t \in T_{tasks}, k \in K} q_t^k + W_2 \cdot \sum_{t \in T_{sources}, k \in K} q_t^k + W_3 \cdot \sum_{t \in T_{tasks}, k \in K} \rho_t^k$$
(14)

Each train task must be covered at most according to the number of each train unit type assigned to the task in the TUP model, see Eq. 15 and 16.

$$\sum_{k \in K, type_k=m} cars_{m(k)} \cdot q_t^k \le \sum_{p \in P, m \in M} cars_m \cdot X_t^{m,p}, \qquad \forall \quad t \in T_{tasks}$$
(15)

$$\sum_{k \in K, type_k=m} units_{m(k)} \cdot q_t^k \le \sum_{p \in P, m \in M} units_m \cdot X_t^{m,p}, \quad \forall \quad t \in T_{tasks}$$
(16)

The  $\rho_t^k$  variables are defined in Eq. 17.

$$2 \cdot \rho_t^k \le q_t^k + q_{\nu(t)}^k, \qquad \forall \quad k \in K, t \in T_{tasks} \setminus T^1$$
(17)

The train tasks assigned to a train unit must form a valid train route i.e. a path through the network, which is feasible with respect to time and place of each adjacent pair of train tasks on the route. Also, the train route for each individual train unit must be valid with respect to any required start and end depots of the train unit. We add a set of virtual nodes to the network one set representing the source nodes,  $N_{so}$ , of each individual train unit and one set representing the sink nodes,  $N_{si}$ , of each individual train unit. For each train unit there is a sink node for each depot i.e. there are |Depots| |*Train units*| sinks in total.

The flow constraints ensuring valid paths are in Eq. 18 to 22. Eq. 18 ensures that if the source

of a train unit is not covered, the train unit is not covering any of the tasks. Eq. 19 ensures that if the source is covered for a train unit, then so is exactly one of the sinks of the train unit. Eq. 20 and 21 are equivalent to flow constraints. They ensure that if train unit k is covering task tthen at least one of the predecessors, pred(t), respectively successors, succ(t) are covered. Finally Eq. 22 ensure that if train unit k is assigned to task t then it can cover none of the train units parallel in time to t. A task is time parallel to t when its dismantling time intersects the dismantling time of t. The parameter n in Eq. 22 indicates the maximum number of tasks present within the time interval of task t on any other sequence in the relevant problem instance.

$$q_t^k \le q_{t'}^k, \quad \forall \quad k \in K, t' \in T_{source(k)}, t \in T_{tasks} \setminus T_{source(k)}$$

$$(18)$$

$$\sum_{t \in T_{sinks(k)}} q_t^k - q_{t'}^k = 0, \qquad \forall \quad k \in K, t' \in T_{source(k)}$$

$$\tag{19}$$

$$\sum_{t' \in T_{pred(t)}} q_{t'}^k \ge q_t^k, \qquad \forall \quad k \in K, t \in T_{tasks}$$

$$\tag{20}$$

$$\sum_{t' \in T_{succ(t)}} q_{t'}^k \ge q_t^k, \qquad \forall \quad k \in K, t \in T_{tasks}$$

$$\tag{21}$$

$$\sum_{t' \in T_{parallel(t)}} q_{t'}^k \le n - (n-1) \cdot q_t^k, \qquad \forall \quad k \in K, t \in T_{tasks}$$

$$\tag{22}$$

In the present form the train routing model does not carry information of the position in each train task composition. This means that if two train units are assigned to a train task and do not carry information of their respective position in the composition, the model will not know which train unit can be decoupled after the task.

Evidently this will only be a problem if more than one unit is assigned to the train task.

If two train units of the same type are assigned to the train task it will not make a difference either. The two train units can simply be assigned their respective positions in the composition after the train paths for each train unit have been found.

A problem may arise when two train units of different types are assigned to the train task. Successor train tasks represent relationships from one train task to the next. The information of open positions are related to the composition of the task and not to the individual train unit. To be able for each train unit to determine if it is in the open end of the composition and therefore can be decoupled and assigned to another train sequence we must transfer the information of open end of the train unit to the train types. That is, we must identify the type closed for decoupling. This can be done using the information from the Train unit position model. It leads to a parameter  $After_t^m$ .

 $ClosedAfter_t^m = \begin{cases} 1 & \text{If train unit of type } m \text{ is closed for decoupling after arrival of train task } t \\ 0 & \text{Otherwise} \end{cases}$ 

$$ClosedAfter_t^{m_k} \le q_{\nu(t)}^k, \quad \forall \quad t \in T_{tasks} \setminus T_1, k \in K$$
 (23)

Notice that the Train position model and the Train ID model can function without the Train Sequence model. When the train sequence model is included in the solution process the constraints in Eq. 24 is included in the train routing model.

$$\Phi_s^k = 1 \Rightarrow q_t^k = 1, \qquad \forall \qquad k \in K, t \in T_{tasks}$$
(24)

First we will try to implement the model using Concert Technology and solve the model with Cplex. There are however potentially more than 75,000 variables and solving the model with Cplex is expected to be too time inefficient. The large number of variables comeS mainly from the  $q_t^k$  variables, which account for 60,500 of the total. The rest being the auxiliary variables.

#### Initial computational results

We have tested the decomposed approach for the Rolling Stock Recovery Problem on scenarios where the recovery window is 1, 2 and 3 hours all starting at 7:00 o'clock which is the time of day with the most departures. Also, we test scenarios with 1, 2 and 4 train lines. See Appendix A.

As expected computation time increases as the size of the test scenarios increase. It appears that increasing the length of the recovery window has a larger effect on computation time than increasing the number of involved train lines. See Appendix B

We have tested the effect of leaving out the sequence model. The preliminary results show that solution time increases when the sequence model is left out entirely. Also, we are able to cover far more tasks when it is included. See Appendix C.

#### Conclusion

In this paper we addressed the rolling stock recovery problem. We have formulated a decomposed solution approach of three iterative models. Initial computational results indicate that the models are a valid approach for minor problems. However, as the problem size increases so does the computation time considerably. Also, the computation time and solution quality decreases considerably if we leave out the sequence model.

#### References

- A. Alfieri, R. Groot, L. Kroon, and A. Schrijver. Efficient circulation of railway rolling stock. *Transportation Science*, 40(3):378–391, Aug. 2006.
- [2] J. K. C. Q. W. S. Ben-Khedher, N. Schedule optimization at sncf: From conception to day of departure. *Interfaces*, 28:6–23, 1998.
- [3] P. Brucker, J. Hurink, and T. Rolfes. Routing of railway carriages. Journal of Global Optimization, 27(2-3):313-332, Nov. 2004.
- [4] J.-F. Cordeau, G. Desaulniers, N. Lingaya, F. Soumis, and J. Desrosiers. Simultaneously locomotive and car assignment at via rail canada. *Transportation Science*, 35(8):767–787, Sept. 2001.
- [5] J.-F. Cordeau, P. Toth, and D. Vigo. A survey of optimization models for train routing and scheduling. *Transportation Science*, 32:380–404, 1998.
- [6] P.-J. Fioole, L. G. Kroon, G. Maróti, and A. Schrijver. A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research*, 174:1281–1297, 2006.
- [7] M. Hofman, L. F. Madsen, J. J. Groth, J. Clausen, and J. Larsen. Robustness and recovery in train scheduling - a simulation study from dsb s-tog a/s. In *ATMOS 2006*. Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2005. IMM-Technical Report-2006-12.
- [8] D. Huisman, L. G. Kroon, R. M. Lentink, and M. J. C. M. Vromans. Operations Research in passenger railway transportation. *Statistica Neerlandica*, 59(4):467–497, 2005.
- [9] N. Lingaya, J.-F. Cordeau, G. Desaulniers, J. Desrosiers, and F. Soumis. Operational car assignment at via rail canada. *Transportation Science*, 36(9):755–778, Nov. 2002.
- [10] M. Peeters and L. Kroon. Circulation of railway rolling stock: A branch-and-price approach. Technical report.
- [11] A. Schrijver. Minimum circulation of railway stock. CWI Quart, 6:205-217, 1993.
- [12] J. Törnquist. Computer-based decision support for railway traffic scheduling and dispatching: a review of models and algorithms. In Presented at ATMOS2005, Palma de Mallorca, Spain, October 2005 (proceedings to be published on the Dagstuhl Research Online Publication Server (DROPS) during 2006, 2006.

#### A Appendix: Test scenarios

| Time span | Lines        | Hours        | Sequences | Units | Tasks | Variables | Constraints |
|-----------|--------------|--------------|-----------|-------|-------|-----------|-------------|
| 1 Hour    | С            | 7:00 - 8:00  | 6         | 36    | 12    | 192       | 224         |
| 2 Hours   | С            | 7:00 - 9:00  | 6         | 36    | 24    | 384       | 442         |
| 3 Hours   | С            | 7:00 - 10:00 | 6         | 36    | 36    | 576       | 660         |
| 1 Hour    | A, A $+$     | 7:00 - 8:00  | 16        | 43    | 36    | 576       | 660         |
| 2 Hours   | A, A $+$     | 7:00 - 9:00  | 16        | 43    | 72    | 1152      | 1314        |
| 3 Hours   | A, $A+$      | 7:00 - 10:00 | 16        | 43    | 108   | 1728      | 1968        |
| 1 Hour    | A+, C, H, H+ | 7:00 - 8:00  | 34        | 67    | 60    | 960       | 1086        |
| 2 Hours   | A+, C, H, H+ | 7:00 - 9:00  | 34        | 67    | 120   | 1920      | 2166        |
| 3 Hours   | A+, C, H, H+ | 7:00 - 10:00 | 34        | 67    | 180   | 2880      | 3252        |

#### **B** Appendix: Comparing effect of size of problems

| Scenario      | Ideal KM    | Act. KM     | Av. S. | Comp. | Time      | $Gap_{pos}$ | $Gap_{rout}$ |
|---------------|-------------|-------------|--------|-------|-----------|-------------|--------------|
| 1 Line, 1 H.  | 200,67      | 242,77      | 2      | 5     | $1,\!58$  | 0,00        | 0,00         |
| 1 Line, 2 H.  | 374,08      | 391,42      | 9      | 5     | $1,\!69$  | 0,00        | 0,00         |
| 1 Line, 3 H.  | 485,55      | 567, 33     | 4      | 9     | $3,\!88$  | 0,00        | 0,00         |
| 2 Lines, 1 H. | 986,71      | 561, 13     | 124    | 7     | 2,56      | 0,00        | 0,00         |
| 2 Lines, 2 H. | $1911,\!88$ | 1059,73     | 118    | 15    | 16,33     | 0,00        | 0,00         |
| 2 Lines, 3 H. | 2490,78     | 1088,54     | 146    | 50    | 620, 36   | 6,29        | 0,00         |
| 4 Lines, 1 H. | $1642,\!85$ | 1360, 80    | 53     | 11    | $^{7,19}$ | 0,00        | 0,00         |
| 4 Lines, 2 H. | 3068, 34    | $2601,\!64$ | 45     | 25    | 642, 25   | 1,02        | 0,00         |
| 4 Lines, 3 H. | $3925,\!40$ | 3833, 37    | 48     | 35    | 1209, 19  | 7,64        | 0,22         |

## C Appendix: Testing effect of including sequence model

| Lines | Scenario        | Ideal KM | Act. KM | Av. S. | Comp. | Time         | $Gap_{pos}$ | $Gap_{rout}$ |
|-------|-----------------|----------|---------|--------|-------|--------------|-------------|--------------|
| 1     | Incl. Seq.      | 374,08   | 391     | 9      | 5     | 1,88         | 0,00        | 0            |
| 2     | Incl. Seq.      | 1911,88  | 1.060   | 118    | 20    | 16,41        | 0,00        | 0            |
| 4     | Incl. Seq.      | 3068, 34 | 2.602   | 45     | 27    | 642, 14      | 1,01        | 0            |
| 1     | Pref. from Seq. | 374,08   | 399     | 9      | 5     | 1,81         | 0,00        | 0            |
| 2     | Pref. from Seq. | 1911,88  | 913     | 151    | 13    | 19,92        | 0,00        | 0            |
| 4     | Pref. from Seq. | 3068, 34 | 2.426   | 57     | 13    | 624,00       | 1,01        | 0            |
| 1     | No Seq.         | 374,08   | 332     | 32     | 22    | $^{5,72}$    | 0,00        | 0            |
| 2     | No Seq.         | 1911,88  | 638     | 202    | 28    | 615, 63      | 0,00        | $^{0,14}$    |
| 4     | No Seq.         | 3068, 34 | 2.351   | 61     | 96    | $1.208,\!00$ | 1,01        | 17,60        |