DYNAMIC MODEL OF THE PASSENGER FLOW ON RAIL BALTICA

Juri Tolujew
Irina Yatskiv
Ilya Jackson

Tobias Reggelin

Department of Mathematical Methods and Modelling
Transport and Telecommunication Institute
Lomonosova Str. 1
LV-1019 Riga, LATVIA

Department of Logistics and Material Handling Systems
Otto von Guericke University Magdeburg
Universitätsplatz 2
39106 Magdeburg, GERMANY

ABSTRACT

This paper describes the design and application of a model simulating the passenger flows on the high-speed rail line Rail Baltica. The proposed model allows one to define passenger flow service indicators, according to data on the behavior of potential passengers that take the train in different localities along Rail Baltica. If corresponding baseline data is provided and the duration of the longest trip will not exceed one day, the model can be used to study processes on other railways with a similar structure. The developed model simulates movements and accumulation of passengers at stations and in trains on four sections connecting five cities along Rail Baltica. Besides that, the model allows for studying particular scenarios related to large sports or other events that will take place near one or several cities through which the railway passes.

1 INTRODUCTION

Rail Baltica is a project to construct a high-speed railway that will link up the Baltic states Lithuania, Latvia, and Estonia with Central European countries. As the result of the “Rail Baltica I” stage, a standard-gauge railway with a track gauge of 1,435 mm will be built. The railway will connect Tallinn to the Polish border passing through Kaunas and Riga. The endpoint of the railway in Tallinn will be connected to Finland by a ferry. Until approximately the year 2030, the route between Warsaw and Berlin, as well as the route between Warsaw and the Polish-Lithuanian border will be modernized. In such a case, the length of the railway between Tallinn and Warsaw will be at least 950 km (Railway Gazette 2015).

It is known that until recently all three Baltic states utilized only railways with a track gauge of 1,524 mm, which corresponds to the modern Russian standard. Such a track gauge was developed by Russian engineers in the middle of 19th century for the Saint Petersburg-Moscow railway construction. This decision was influenced by the experience of the southern states of the USA, because American engineers took part in the Russian railway construction. The most famous of them is George Washington Whistler (Gasparini et al. 2015).

Rail Baltica is one of the priority technical projects of the EU (TEN-V 27). The project is a part of the Trans-European Networks (TEN) concept, which aims to reinforce economic and social connections among European countries. An additional extension of Rail Baltica is the connection of Berlin to northern Italy and Venice by a high-speed railway.

Even though all pivotal decisions concerning Rail Baltica have already been made, lots of issues related to the demand for passenger and freight transportation are still waiting to be resolved. The project developers are currently relying on a forecast, which says that 5 million passengers and 13 million tons of cargo will be carried in 2030. However, such figures must be better-founded and related to actual dynamics.
This paper describes a model to calculate dozens of logistical and financial indicators taking into account the forecast data. The mentioned data reports on the behavior of the various groups of potential passengers that start their trips in the specific cities connected by the railway. Since the model utilizes the time counting method with the constant step “delta t” and dynamical group behavior is displayed instead of the individual behavior, the simulation paradigm is close to System Dynamics. The proposed model is universal in the sense that it may be applied with different numerical data to analyze the dynamics of any railway passing through several cities.

2 RELATED WORK

In the field of railway transport simulation modeling is usually applied in order to assess the route capacity and optimize train schedules. Models of this type are commonly microscopic, because they display a traffic of particular objects (locomotives or entire trains) on a particular railway section. Software implementation of these models utilizes the discrete event simulation paradigm (Banks et al., 2014). Far less often, simulation modeling is used to solve problems of long-term railroad and adjoin infrastructure planning. These models belong to the macroscopic class and require the System Dynamics paradigm (Sterman 2000).

2.1 Discrete Event Simulation

Simulation-driven methods allow one to assess the total delay time for all trains in the schedule. There are two modeling approaches: asynchronous and synchronous (Pachl 2014). The asynchronous modeling is particularly applicable for studying train schedules under boundary conditions. Such modeling can be used as intellectual support for scheduling software. Synchronized modeling, on the other hand, displays a specific railway operation with a high level of detail. This modeling approach is applicable for short-term processes’ forecasting and, for example, to assist the dispatcher in selecting one of several possible options for skipping competing trains.

Several simulation methods are represented in commercial software products. These tools are supported with GUIs and also imitate railway traffic and processing operations based on the real objects of the railway infrastructure. Software packages for railway simulation are reviewed in the paper by Kontaxi and Ricci (2011).

For instance, in Germany the LUKS system is well-known. The system contains the module LUKS-S for medium-term analysis and planning of the processes on the railway sections and intersections by the simulation (Janecek and Weymann 2010; LUKS 2011). The simulation of timetabling provides information on whether a predefined operating schedule is viable, i.e., whether it allows a conflict-free timetable to be compiled that contains the envisaged train movements. Simulation provides the ability to deal with unforeseen disruptions in a timetable already compiled. Conflicts are detected and asynchronously resolved by means of mapping at operation control centers. Train movements, by contrast, are synchronously simulated so as to be able to map any unforeseen disruptions.

2.2 System Dynamics

The System Dynamics models described by Forrester (Forrester 1972) belong to the class of continuous models with the constant time-step “delta t”. Such models do not display the traffic of the physical objects (vehicles, cargos, passengers, etc.), but only calculate the number of these objects that must be located in a certain point in space at a certain time.

The study by Homer et al. (1999) describes the application of System Dynamics modeling for strategic planning of a large railway station operating in the USA. The cause-effect relationship between the modeled processes is described in graphic form. The comparison of mean delay time obtained from the model and factual delay time is given as an example. The model was approved by the top management of CSX Transportation and was used for budgeting. It was applied to analyze the influence of various demand scenarios and investment policies on the performance indicators.
Mannaerts et al. (2013) describe in details the model of the railway transport in the Netherlands. The model allows users to assess three of the most well-known indicators on which the travel performance ratio is based, such as price-performance ratio, subjective assessment of travel comfort and travel time reliability. At the same time, the model displays direct travel expenses and additional expenses related to the reliability, and the travel time is also calculated. The expenses are calculated for all stages of travel, for instance, the parking charge or the cost of a ride to the station and back are taken into account. Chao et al. (2013) report on the evolution model of the urban transport in Shanghai. The model is designed according to the System Dynamics approach based on the observations from 2004 to 2009. The described system also includes railway transport. Right after a quantitative analysis of the passenger transportation structure, the model will be applied to analyze trends in urban transport development as part of the long-term strategic planning. The paper by Li et al. (2013) parses the System Dynamics model that was applied to study passenger flow at the railway transit station Beijing Dongdan. The indicators of the station capacity are defined under various conditions. The decisions on service of the required station capacity and passenger flow safety are made based on the modeling results.

3 CONCEPTUAL MODEL

The object of the modeling is the passenger flow that may be carried by trains along Rail Baltica. The unit of time in the model is a day, that is to say, the pivotal logical part of the model must provide a way to calculate the number of passengers that may take the train or leave it at all stations during a day. Since the demand for passenger transport basically depends on the week number, a year is chosen as the experiment run time, i.e., 365 days. The passenger flows are not attached to particular trains, on contrary, they are related to particular days during the year. Interpreting the modeling results, we make the final decision on a required number of trains with a known capacity. The passenger transportation processes may be modeled in both directions by Rail Baltica. Information on the passenger flows is an initial modeling result. As secondary result, the model must calculate the indicators related to financial flows, i.e., monetary transactions that refer to ticket selling under the assigned policy.

3.1 Passenger Flow Modeling

The passenger flows between starting and ending stations Tallinn and Berlin are the core subject of the modeling. Riga, Kaunas and Warsaw are chosen as the intermediate stations. The model does not take into account the distance between stations, the speed of trains and the travel time, because all input parameters and output indicators are attached to the chosen unit of time – a day. Thereby, the model parses the passenger flows in two directions:

Tallinn – Riga – Kaunas – Warsaw – Berlin
and

It is assumed that there are the forecast data on the number of passengers, who wish to travel by the railway during a day. Trips of this passengers’ groups are denoted by the term “there”. It is also assumed that each passenger after a certain time will make a trip in the opposite direction, hereinafter “back”. The passengers taking the train just “there” constitute the insignificant share. It is assumed that groups of passengers, who will travel “back” in 2, 4, 6, 8 or 15 days right after the trip “there”. The days of stay determined in the model correspond to such types of trips as a trip for a weekend, a business trip, a tourist trip or a fortnight visit. It is not really necessary to have extremely accurate data on such a distribution of passengers by the time they stay at the destination, since all of them will return in not later than 15 days. Generally, the model may also take into account passengers, who will stay longer at the destination, if such passengers make up a significant share. Figure 1 shows the exemplary distribution of passengers taking the
train in Riga, according to the directions and the time of their stay at the destination. The flows of passengers starting their trips from the other four stations may be simulated by the same scheme.

![Figure 1: The principle of the group formation and distribution (Riga station).](image)

The average daily demand for transportation is set for all five starting stations of the route for all 365 days of the year. Data on the distribution of passengers by the directions and the time of stay are also declared as values of the corresponding probabilities for all five starting stations.

In the model there are five possible spans for buying a ticket in accordance with the day of the “there” trip:

- from 0 to 10 days,
- from 11 to 20 days,
- from 21 to 30 days,
- from 31 to 60 days,
- from 61 to 90 days.

The percentage of the basic ticket price, as well as the proportion of passengers buying a ticket can be set for each of these spans.

3.2 The Structure of the Initial data and the Model Indicators

The initial data of the model include the following:

- daily demand for transportation from each of the five starting points of the route for 365 days;
- distribution of passengers by the length of stay at the destination for each of the five starting stations;
- table with basic ticket prices for all directions of the trip;
- price table (in percentage) in relation to the basic price for five possible periods of ticket purchase time;
- table with shares of passengers by the time of ticket purchase.
The performance indicators of the model related to the passenger flows service are the following:

- the number of passengers who travel “there” on the current day from each of the five starting stations, taking into account four possible destinations (20 indicators for each day of the model's operating time);
- the number of passengers who travel “back” on the current day to each of the five starting stations, taking into account four possible destinations (20 indicators for each day of the model's operating time);
- the number of passengers who boarded a train on the current day in each of the five starting stations for both directions of the train traffic (8 indicators for each day of the model's operating time);
- the number of passengers who get off the train on the current day in each of the five starting stations for both directions of the train traffic (8 indicators for each day of the model's operating time);
- the number of passengers traveling by the train on each of the four route sections in the directions Tallinn-Berlin and Berlin-Tallinn (8 indicators for each day of the model's operating time).

Thus, 64 indicators related to the passenger flow service are calculated for each day of the model. The performance indicators of the model related to the income dynamics, i.e., monetary flows arising from the ticket selling at various time spans before the trip commencement, are the following:

- daily income;
- cumulative income for the year.

In order to obtain accurate annual revenue indicators from ticket sales, only those passengers who have started their trip in the considered year are taken into account. Since ticket sales can begin up to 90 days before the trip, an extended span of 90 + 365 days is considered to calculate and display revenue indicators.

4 SIMULATION MODEL

Since a large size of both initial data and data with simulation results is the distinguishing feature of the model, it was decided to develop a special API based on MS Excel. Developing the simulation model, it was necessary to envisage the possibility of exchanging data with this API. Since the ExtendSim 7 software provides this capability (Krahl 2007), it was chosen for the software implementation of the model.

The structure of the simulation model completely corresponds to the structure of the conceptual model (Figure 1). Figure 2 represents one-fifth of the whole model. The formation and motion of the passenger flows that go to Tallinn are displayed in the model. Similar fragments were designed for modeling flows to Riga, Kaunas, Warsaw, and Berlin.

“Data Import Export” blocks are used in the model in order to provide the data exchange with Excel. We import three tables with initial data and export two tables with a modeling result. “Read” blocks are utilized to read data from three tables with initial data. “Write” blocks are needed to write down data in two tables with a modeling result. If the “Read” or the “Write” block is connected to the “Time” block, the corresponding operations are performed at each of the 365 steps during the simulation time.

The feasible way to verify the model is to compare the indicators passengers from Riga “there” and passengers from Riga “back”. These indicators are calculated for 380 days of the simulation (Figure 3). The first 15 days refer to the initial simulation phase: the early-bird passengers departed from Riga begin to return, after that, the simulation turns into the normal phase. According to Figure 3, the charts are almost the same, which proves that nearly all passengers departed from Riga returned back no later than in 15 days. The first passengers begin their trips in day number 1. Since the shortest trip lasts 2 days, the first passengers in the “back” flow appear only in day number 3. Such a backlog (red line) of the “back” flow is observed throughout the year, which proves the error-free operation of the model.
5 EXPERIMENTS AND RESULTS

This section describes two experiments with the model. In the first series we study only passenger flows based on demand forecast data, i.e., a number of passengers along the railway (at the stations and on the trains) during the year. The second series studies the influence of pricing strategy for tickets on the annual income.

The experiments described in this paper utilize the initial data based on the current flow of passengers that use other modes of transport to travel between cities through which the Rail Baltica will pass. Such data are very approximate and serve only to verify the performance of the model. The key positive property of the initial data is their completeness. This means that by conducting experiments with more accurate prediction data, users of the model will prepare the same Excel tables that were used in the experiments described below.
5.1 Passenger Flows Modeling

The total demand for transportation at each station is initially set in the form of a deterministic trend. After that, the deterministic trend is randomized in order to produce a factual demand (Figure 4a). Figure 4b demonstrates the probability distribution on the direction of travel for passengers starting their trip in Riga. For passenger flows at each initial station, we set the probability distribution of passengers by the length of their stay at the destination (Figures 1 and 2). However, these probability distributions are not functions of time.

![Figure 4a: Demand from Riga (passengers per day)](image)

![Figure 4b: Distribution by directions](image)

Figure 4: Demand for trips from Riga.

Running the simulation with the mentioned initial data, we obtained the number of passengers traveling round-trip (Table 1). As the result, 586,325 passengers made their trips in 365 days. Taking into account that each passenger has to take the train twice, the number of the factually transported passengers is 1,172,650.

<table>
<thead>
<tr>
<th>to Tallinn</th>
<th>to Riga</th>
<th>to Kaunas</th>
<th>to Warsaw</th>
<th>to Berlin</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>from Tallinn</td>
<td>0</td>
<td>14594</td>
<td>4866</td>
<td>29167</td>
<td>48618</td>
</tr>
<tr>
<td>from Riga</td>
<td>15781</td>
<td>0</td>
<td>5024</td>
<td>30146</td>
<td>50245</td>
</tr>
<tr>
<td>from Kaunas</td>
<td>4163</td>
<td>7123</td>
<td>0</td>
<td>21388</td>
<td>35632</td>
</tr>
<tr>
<td>from Warsaw</td>
<td>67987</td>
<td>40445</td>
<td>7785</td>
<td>0</td>
<td>46665</td>
</tr>
<tr>
<td>from Berlin</td>
<td>24471</td>
<td>77775</td>
<td>7785</td>
<td>46665</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>112402</td>
<td>139937</td>
<td>25460</td>
<td>127366</td>
<td>181160</td>
</tr>
</tbody>
</table>

Figure 5 contains detailed data on the passengers’ position in the space along the Tallinn-Berlin direction. The maximum values of the indicators are of a particular interest, since the corresponding number of passengers per day will use the station's infrastructure. For instance, up to 800 passengers a day can board a train at the Tallinn and Riga stations. These flows take into account both groups of passengers: those, who start their trip in Tallinn or Riga and those, who return from Tallinn and Riga to all other cities. Figure 5b shows that at the end of the summer, up to 1,200 passengers arrive in Berlin every day.

Figure 5c represents key route capacity indicators of railway sections for Tallinn-Berlin direction. These indicators are the most significant for logistics. It is easy to explain logically the fact that the largest number of passengers on the train may occur on the inner sections of the railway, and as soon as they approach the terminal station (in this case, Berlin), their number decreases. For instance, in the Kaunas-Warsaw section in the summer months it will be necessary to transport up to 1600 passengers a day. If the capacity of one train is, for example, 800 passengers, then, at this time of year, two trains may be completely filled. In February, the maximum number of passengers on this site hardly exceeds 900. In this regard, the carriers will have to make a decision: either apply one “elongated” composition or two “shortened”. 

3102
The simulation results of the transportation process in the Berlin-Tallinn direction are interpreted in the same way (these diagrams are not presented in this article). Up to 1,300 passengers a day can be observed getting off in Berlin. More than 800 passengers a day can leave the train at the terminal station in Tallinn. The busiest section is, once again, Warsaw-Kaunas, where in the summer months up to 1,700 passengers per day will be transported.

![Diagram](image1)

**Figure 5:** The modeling result on Tallinn-Berlin.

### 5.2 Monetary Flows Analysis

The estimated basic price for all directions of the trip is shown in Table 2. Table 3 describes examples of pre-sale strategies. The data shown in the columns have the following meaning:

- time spans for the preliminary sale of tickets;
- prices (in percentage) in relation to the basic price for each time span;
- the proportion of passengers buying tickets in each time span.
Table 2: The ticket price (Euros) for all directions of the trip “there and back”.

<table>
<thead>
<tr>
<th></th>
<th>to Tallinn</th>
<th>to Riga</th>
<th>to Kaunas</th>
<th>to Warsaw</th>
<th>to Berlin</th>
</tr>
</thead>
<tbody>
<tr>
<td>from Tallinn</td>
<td>0</td>
<td>60</td>
<td>120</td>
<td>170</td>
<td>200</td>
</tr>
<tr>
<td>from Riga</td>
<td>60</td>
<td>0</td>
<td>60</td>
<td>120</td>
<td>170</td>
</tr>
<tr>
<td>from Kaunas</td>
<td>120</td>
<td>60</td>
<td>0</td>
<td>60</td>
<td>120</td>
</tr>
<tr>
<td>from Warsaw</td>
<td>170</td>
<td>120</td>
<td>60</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>from Berlin</td>
<td>200</td>
<td>170</td>
<td>120</td>
<td>60</td>
<td>0</td>
</tr>
</tbody>
</table>

Time spans in such a model are constant, however, two sale strategy parameters may be altered. Table 3 shows the modeling-based annual income by each of the three strategies. If strategy 1 is considered as the initial one, it can be seen that strategy 2 increases the income, and strategy 3 reduces it. The income dynamics from ticket sales during the year in case of strategy 1 is illustrated in details (Figure 6). Besides the monetary flow, the diagrams show the cumulative passenger flow.

Figure 6: The income dynamics from ticket sales.

Since the preliminary sale begins 90 days prior to the start of the first trip, the charts with a minus sign represent the number of days before the commencement of this trip, i.e., before the first day of the simulated process. The chart “Income per day” demonstrates the decrease, since ticket sales are not taken into account for passengers who will begin their trips next year. Thus, the model calculates the income from the sale of tickets to passengers who begin their trips within one year taking the advantage of pre-sale at the same time.
Table 3: Examples of pre-sale strategies.

<table>
<thead>
<tr>
<th>Pre-sale</th>
<th>Strategy 1</th>
<th>Strategy 2</th>
<th>Strategy 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prices in relation to the basic in percentage</td>
<td>Ticket buyers in percentage</td>
<td>Prices in relation to the basic in percentage</td>
</tr>
<tr>
<td>for 61-90 days</td>
<td>50</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>for 31-60 days</td>
<td>100</td>
<td>40</td>
<td>90</td>
</tr>
<tr>
<td>for 21-30 days</td>
<td>125</td>
<td>10</td>
<td>110</td>
</tr>
<tr>
<td>for 11-20 days</td>
<td>150</td>
<td>10</td>
<td>130</td>
</tr>
<tr>
<td>for 0-10 days</td>
<td>100</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Annual income (K€)</td>
<td>71,746</td>
<td>73,490</td>
<td>67,480</td>
</tr>
</tbody>
</table>

6 CONCLUSIONS

The developed model shows the complete “anatomy” of movements and accumulation of passengers at stations and in trains on four sections connecting five cities along Rail Baltica. The pivotal research outcome is the model itself, but not the results of the experiments, which were conducted in order to verify the model. At the same time, the key modeling result, namely 1.17 million passengers a year, seems to be quite realistic, according to the demand for transportation, which was very roughly estimated for the end of 2016. The research demonstrated to the Rail Baltica stakeholders all variabilities of the initial data. Based on that, the Rail Baltica stakeholders may design and conduct working experiments with the model. All the data to forecast the demand for transportation from five European cities for the period until 2030 should be obtained within the framework of a different project that is not directly related to the simulation.

If the duration of the longest trip will not exceed one day, the model can be used to study processes on similar railways. Such a condition arises, since the model is macroscopic and the unit of time is one day. Any attempt to take into account passengers who will spend more than 24 hours in a trip will require to add train timetables into the model. Such an extension will lead to a radical change in the nature of the model.

Besides that, the model allows for studying particular scenarios related to large sports or other events that will take place near one or several cities through which the railway passes. The simulation result will be the expected number of arriving and departing passengers on each day of the event for each city along the railway. In such a case, the task of specialists in the field of logistics will be train scheduling taking into account the capacity of other infrastructure facilities, i.e. railway stations and urban public transport systems.

ACKNOWLEDGEMENTS

This work has been supported by the ALLIANCE project (http://alliance-project.eu/) and has been funded within the European Commission’s H2020 Programme under contract number 692426. This paper expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this paper.

REFERENCES


AUTHOR BIOGRAPHIES

JURI TOLUJEW is a professor of Mathematical Methods and Modelling at the Transport and Telecommunication Institute (Riga, Latvia) and a project manager at the Fraunhofer Institute for Factory Operation and Automation IFF in Magdeburg, Germany. He received a doctoral degree in automation engineering from the Technical University of Riga. He also received a habil. degree in computer science from Otto von Guericke University Magdeburg. His research interests include the simulation-based analysis of production and logistics systems, protocol-based methods for analyzing processes in real and simulated system as well as mesoscopic approaches in the area of simulation. He is an active member in the ASIM, the German organization of simulation. His email address is juri.tolujew@iff.fraunhofer.de.

IRINA YATSKIV is a professor of Mathematical Methods and Modelling and Vice-rector of Research and Development, Transport and Telecommunication Institute (Riga, Latvia). She has received a doctoral degree in engineering from the Riga Aviation University. Her research interests include data analysis and mathematical modelling and simulation with application in transportation. She is a member of the Latvian Simulation Society, Classification Society of North America, etc. She has participated in more than 20 European and Latvian research projects and more than 10 national transportation studies and was coordinator in 8 of them. She has been chair or member of the programme and organizing committee of more than 25 International Conferences on Applied Statistics, Transportation Research, Complex Systems, etc.; reviewer in transportation journals and conference proceedings. Her email address is Jackiva.I@tsi.lv.
ITYA JACKSON is a PhD student of Telematics and Logistics at the Transport and Telecommunication Institute, Riga. His research interests include but are not limited to stochastic combinatorial optimization, metaheuristics and artificial intelligence. His email address is jcksnl93@gmail.com.

TOBIAS REGGELIN is a project manager at the Fraunhofer Institute for Factory Operation and Automation IFF and Otto von Guericke University Magdeburg. His main research and work interests are logistics system modeling and simulation and the development and conduction of logistics management games. Tobias Reggelin received a doctoral degree in engineering from Otto von Guericke University Magdeburg. Furthermore, he holds a diploma degree in industrial engineering in logistics from Otto von Guericke University Magdeburg and a master’s degree in Engineering Management from the Rose-Hulman Institute of Technology in Terre Haute, IN. His email address is tobias.reggelin@ovgu.de.