



TRANSPORTATION NETWORK MODELING & ANALYSIS

## Applying Traffic Assignment to a railway network



*Fletcher Collis  
Luca Liuzzi*

# Contents

|          |                                     |           |
|----------|-------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                 | <b>3</b>  |
| <b>2</b> | <b>Methodology</b>                  | <b>3</b>  |
| 2.1      | Raw Data . . . . .                  | 3         |
| 2.1.1    | Passenger Data . . . . .            | 3         |
| 2.1.2    | Network Data . . . . .              | 4         |
| 2.2      | OD Generation . . . . .             | 5         |
| 2.3      | Network Definition . . . . .        | 6         |
| 2.4      | Model Specification . . . . .       | 8         |
| 2.4.1    | Link Cost Function . . . . .        | 8         |
| 2.4.2    | Static Traffic Assignment . . . . . | 9         |
| <b>3</b> | <b>Findings</b>                     | <b>10</b> |
| 3.1      | Results . . . . .                   | 10        |
| 3.1.1    | Baseline Scenario . . . . .         | 10        |
| 3.1.2    | Case Study . . . . .                | 10        |
| 3.1.3    | Hypothetical Scenario . . . . .     | 11        |
| 3.2      | Limitations & Next steps . . . . .  | 12        |
| <b>4</b> | <b>Conclusion</b>                   | <b>12</b> |
| <b>5</b> | <b>References</b>                   | <b>13</b> |

# 1 Introduction

Traffic assignment is a problem primarily used to analyze and optimize networks and vehicle traffic. However, this type of problem can also be applied to networks and systems very different from the classic road scenario. In this project, the algorithms and model for solving static traffic assignment were adapted and applied to the SBB intercity train network. The main goal of this project is not so much to analyze the system, but rather to establish a general framework for solving this type of problem in any railway network, define the key characteristics to consider, and understand the potential outputs and their applications.

The project is structured into three parts. The first part focuses on processing the available data to derive the OD matrix and the railway network for IC trains, including the various nodes (stations) and links (tracks). The second part is dedicated to model specification, defining the cost function for the links, the algorithms for solving static traffic assignment, and the equilibrium conditions. Finally, the third part presents the results and explores the potential uses of the model and its outputs.

From our perspective, this project is highly interesting because it allowed us to work with real-world data, which made the project more concrete but also more challenging due to the need to manipulate the available data and handle missing information. Moreover, it enabled us to explore a system different from those typically studied and identify the unique characteristics of railway networks that are not present in road networks.

The motivations and objectives behind this project can be summarized with these four words:

- **Understanding:** Explore the various challenges of a railway network and learn to model one.
- **Comparison:** Uncover the fundamental differences in road vs. rail networks.
- **Feasibility:** See if it is possible to create a framework to easily apply traffic assignment to any real railway network.
- **Flexibility:** Make the approach very flexible and modular to be able to use it for different use cases.

# 2 Methodology

## 2.1 Raw Data

All raw data files used in this project were obtained from the official SBB data portal. In particular, we relied on two main sources: passenger count data and railway network geometry. Both datasets required significant preprocessing and filtering before they could be used in a traffic assignment framework.

### 2.1.1 Passenger Data

The passenger data was obtained from the publicly available dataset *anzahl-sbb-bahnhofbenutzer.csv*, which contains average daily counts of people Swiss railway stations. The data covered approximately 30 of the most frequented stations in Switzerland. However, the dataset only provided marginal totals per station and did not include any information about passenger origins and destinations. Furthermore, there was no disaggregation by time of day or by train line. This lack of OD (origin-destination) specificity meant that a synthetic OD matrix had to be generated later in the process, based on reasonable assumptions.

| Bahnhof_Gare_Stazione | Unité         | Jahr | Anzahl Bahnhofbenutzer |
|-----------------------|---------------|------|------------------------|
| Bern                  | DP/jour ouvré | 2016 | 318000                 |
| Bern                  | DP/jour ouvré | 2020 | 227600                 |
| Bern                  | DP/jour ouvré | 2021 | 228800                 |
| Bern                  | DP/jour ouvré | 2023 | 296500                 |
| Bern                  | DP/jour       | 2013 |                        |
| Bern                  | DP/jour       | 2015 |                        |
| Bern                  | DP/jour       | 2020 | 211000                 |
| Bern                  | DP/jour       | 2022 | 261500                 |
| Bern                  | DP/jour       | 2024 | 281900                 |
| Basel SBB             | DP/jour ouvré | 2013 | 134000                 |
| Basel SBB             | DP/jour ouvré | 2014 | 135000                 |
| Basel SBB             | DP/jour ouvré | 2021 | 95700                  |
| Basel SBB             | DP/jour       | 2015 | 126400                 |
| Basel SBB             | DP/jour       | 2016 | 126400                 |
| Basel SBB             | DP/jour       | 2019 | 128400                 |
| Lausanne              | DP/jour ouvré | 2014 | 140000                 |
| Lausanne              | DP/jour ouvré | 2018 | 144700                 |
| Lausanne              | DP/jour ouvré | 2020 | 96600                  |
| Lausanne              | DP/jour       | 2017 | 135600                 |

Figure 1: Extract from the anzahl-sbb-bahnhofbenutzer.csv

From the full list of available stations, we retained only those served by the InterCity (IC) network. This filtering step left us with the following selection of major stations: ['Bern', 'Basel SBB', 'Lausanne', 'Luzern', 'St. Gallen', 'Winterthur', 'Zug', 'Aarau', 'Baden', 'Biel/Bienne', 'Chur', 'Fribourg/Freiburg', 'Genève-Aéroport', 'Genève', 'Neuchâtel', 'Olten', 'Thun', 'Bellinzona', 'Lugano', 'Zürich Oerlikon', 'Zürich HB'].

### 2.1.2 Network Data

The geometry of the railway network was obtained from the shapefile dataset *linie-mit-polygon.csv*. This file contains polylines representing every segment of railway track in Switzerland, including InterCity (IC) lines, regional services, freight corridors, and tram routes. While the dataset was spatially accurate and highly detailed, it lacked crucial semantic information. In particular, there was no explicit indication of which train lines operated on each track segment, nor were there attributes directly linking the geometry to specific IC services.

Figure 2: Extract from the linie-mit-polygon.csv file

Figure 2 shows a sample extract of the raw data. Each track segment includes attributes such as its length and a line identifier under the **Linie** column. This line identifier proved essential, as it allowed us to determine which track segments belonged to the InterCity network.

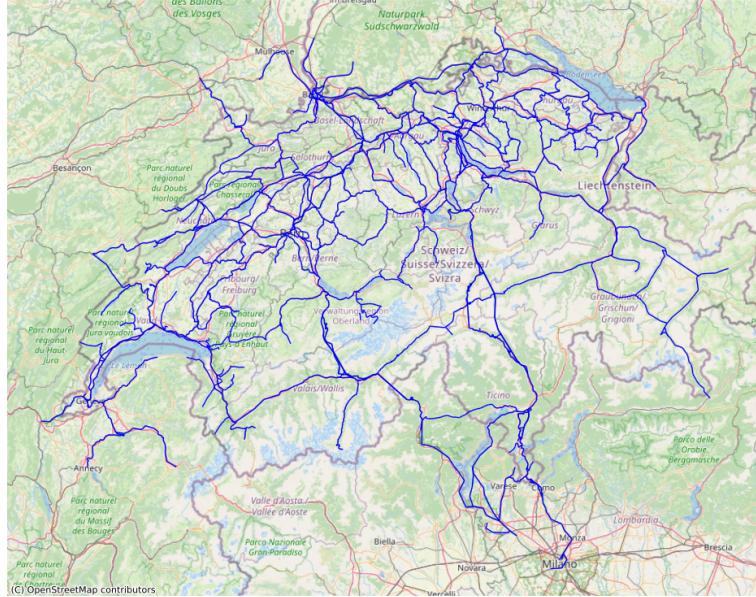


Figure 3: Visualization of the full railway network

Due to the raw and unstructured nature of the dataset, substantial effort was required to clean and prepare it for use in our model. The most significant challenge involved isolating only the relevant InterCity (IC) tracks from the full Swiss railway network. This filtering was performed manually by cross-referencing the line identifiers in the dataset with publicly available route maps and schedules from the SBB website. Through this process, we were able to accurately identify and retain only the segments used by IC services, discarding all unrelated infrastructure.

Once the IC network was extracted, we identified which of our selected passenger data cities were directly connected within it. This cleaned network also allowed us to calculate the real distances between these connected cities, enabling us to represent the railway system as a classical graph, with nodes (stations), edges (connections), and edge weights (distances). The resulting cleaned and structured IC network is shown in Figure 4.

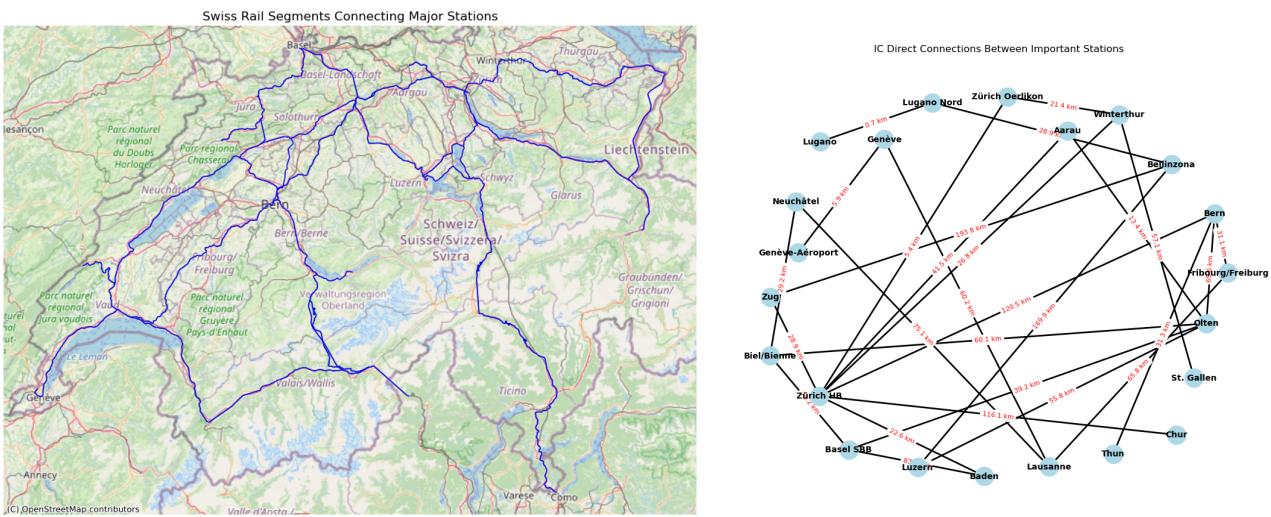


Figure 4: Final network representation

## 2.2 OD Generation

The available data did not include information about OD pairs, but only data on the total number of people in the stations. This data was considered as the production of each station. To obtain the OD pairs between the cities in the intercity train network, it was necessary to generate data for the attraction of each station and then use a doubly constrained gravity model to derive the OD matrix.

## Synthetic Attraction

Considering the production of stations and the population of cities, it was possible to calculate weights,  $\omega$ , using Formula X. These weights were then used to redistribute the production and obtain the attraction of each station, as shown in Formula X.

$$\omega_i = \frac{\text{population}_i + \text{production}_i}{\text{total population} + \text{total production}} \quad (1)$$

$$\text{Attraction}_i = \omega_i \times \text{total production} \quad (2)$$

## Gravity Model

With the production and attraction values for each station, a gravity model could be applied. The model comes from the EPFL course CIVIL-557, "Decision-aid methodologies in transportation," and is a doubly constrained model. The cost function was defined considering only the actual distance between two stations, determined using SBB data. The project exclusively considers trips between two different cities, thus the demand for internal trips (within the same city) was set to 0.

### Consideration about the OD generation

It must be acknowledged that this method is not highly accurate, both in how the attraction was generated and in the cost function for the gravity model. This is because the weight calculation did not account for city attraction characteristics such as the number of jobs or available services. A similar limitation applies to the gravity model's cost function. Despite its simplicity, this method was sufficient to obtain an OD matrix suitable for achieving the project's objectives. Figure 5 shows stations with higher production and attraction values, which are nonetheless consistent with the Swiss reality.

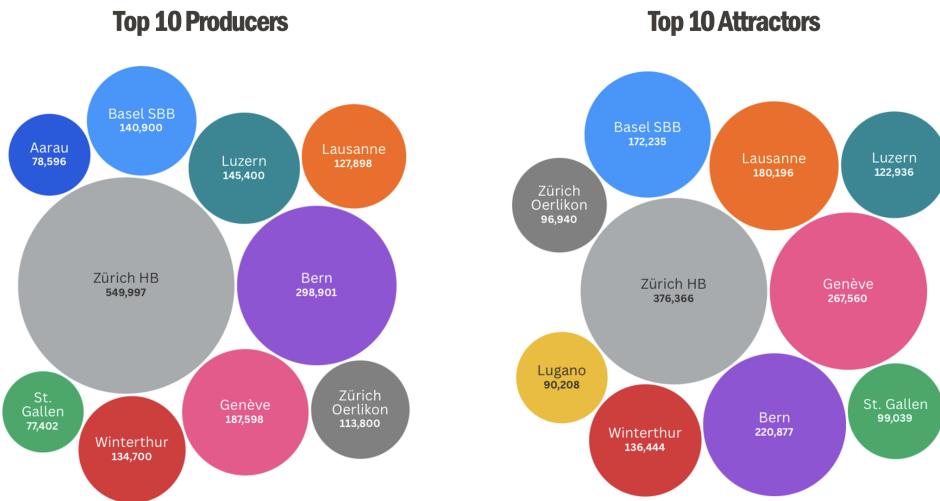


Figure 5: Top 10 Producers and Attractors

## 2.3 Network Definition

The cleaned network data initially lacked explicit information about the train lines operating on each track segment. To address this, we manually mapped the correspondence between InterCity (IC) lines and stations using publicly available schedules and route information from the official SBB website. This allowed us to enumerate which IC lines pass through each edge in our network, resulting in a dataframe that links stations and edges to their respective train lines, as shown in Figure 6.

|    | start_node        | end_node        | distance [m] | IC_lines             |
|----|-------------------|-----------------|--------------|----------------------|
| 0  | Fribourg/Freiburg | Bern            | 31109.895    | IC1                  |
| 1  | Fribourg/Freiburg | Lausanne        | 65778.350    | IC1                  |
| 2  | Chur              | Zürich HB       | 116075.402   | IC3                  |
| 3  | St. Gallen        | Winterthur      | 57139.042    | IC1, IC5             |
| 4  | Neuchâtel         | Lausanne        | 75087.957    | IC5                  |
| 5  | Neuchâtel         | Biel/Bienne     | 29200.120    | IC5                  |
| 6  | Lugano            | Lugano Nord     | 659.934      | IC2, IC21            |
| 7  | Lugano Nord       | Bellinzona      | 28895.062    | IC2, IC21            |
| 8  | Basel SBB         | Baden           | 83614.363    | IC3                  |
| 9  | Basel SBB         | Olten           | 39172.397    | IC21, IC6, IC61      |
| 10 | Basel SBB         | Biel/Bienne     | 73186.658    | IC51                 |
| 11 | Winterthur        | Zürich Oerlikon | 21418.950    | IC1                  |
| 12 | Winterthur        | Zürich HB       | 26787.367    | IC5, IC81            |
| 13 | Bellinzona        | Luzern          | 169866.354   | IC21                 |
| 14 | Bellinzona        | Zug             | 193790.225   | IC2                  |
| 15 | Zürich Oerlikon   | Zürich HB       | 5368.417     | IC1                  |
| 16 | Aarau             | Olten           | 13373.404    | IC5                  |
| 17 | Aarau             | Zürich HB       | 41492.905    | IC5                  |
| 18 | Genève            | Genève-Aéroport | 5918.188     | IC1                  |
| 19 | Genève            | Lausanne        | 60226.630    | IC1                  |
| 20 | Baden             | Zürich HB       | 22587.175    | IC3                  |
| 21 | Luzern            | Olten           | 55768.093    | IC21                 |
| 22 | Olten             | Bern            | 65626.280    | IC6, IC61            |
| 23 | Olten             | Biel/Bienne     | 60070.608    | IC5                  |
| 24 | Thun              | Bern            | 31264.012    | IC6, IC61, IC8, IC81 |
| 25 | Bern              | Zürich HB       | 120492.589   | IC1, IC8, IC81       |
| 26 | Zug               | Zürich HB       | 28929.698    | IC2                  |

Figure 6: Mapping of stations and track segments to IC train lines.

To accurately represent train movements and passenger flows within the network, we implemented a multi-layer graph structure based on the following modeling approach:

- 1. Node-Line Duplication:** Instead of modeling each station as a single node, we created multiple nodes per station, one for each IC line that stops there. This approach allows us to distinguish the different routes trains take even at shared stations, avoiding artificial shortcuts between lines that do not directly connect.
- 2. Transfer Modeling:** Within each station, all duplicated line-specific nodes were interconnected by transfer arcs to model passenger transfers. These transfer links represent the physical movement passengers must make when switching lines at a station.
- 3. Transfer Penalties:** Transfer arcs were assigned a fixed penalty time reflecting realistic transfer conditions. This penalty consists of a 5-minute base transfer time plus an additional half of the average headway for the connecting line, thereby capturing waiting times and transfer inconvenience. Incorporating this penalty makes transfers explicitly penalizable in shortest path and traffic assignment calculations, reflecting passengers' natural preference to minimize transfers.

We again used information available on the official SBB website to determine the headway for each line. The headway, defined as the scheduled interval between consecutive trains on the same line, is an important parameter for modeling transfer penalties and passenger waiting times. Based on the available data, the lines were categorized into two groups with distinct headways:

- **30-minute headway lines:** IC3, IC5, IC6, IC61, IC81
- **60-minute headway lines:** IC1, IC2, IC8, IC21, IC51

Finally, we obtained a fully processed and modular representation of the railway network that can be directly used with classical algorithms commonly applied in road traffic assignment. This model allows for flexible adjustments of key parameters such as train headways, link speeds, and transfer penalties, enabling scenario analysis and calibration. The final network structure is illustrated in Figure 7.

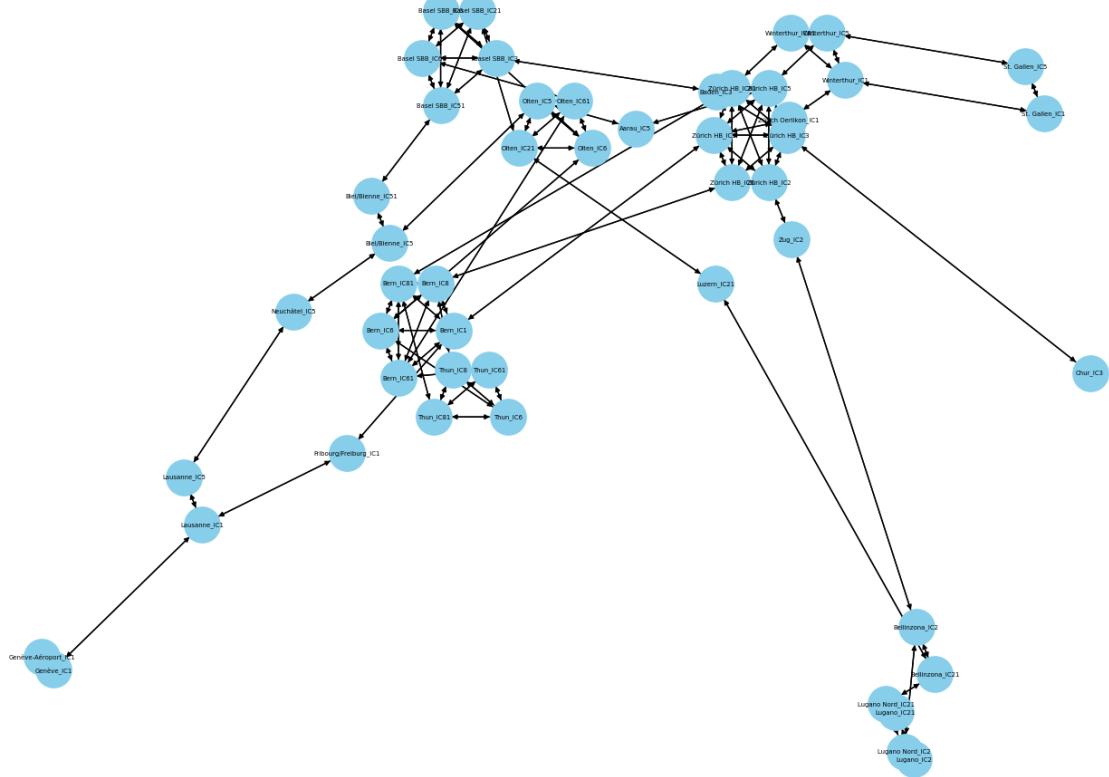


Figure 7: Final representation of the railway network model suitable for traffic assignment algorithms.

## 2.4 Model Specification

### 2.4.1 Link Cost Function

The cost of each link is characterized by three elements: travel time, a penalty due to line transfer, and discomfort caused by train crowding. The equation 3 is suitable for all types of links (travel links and transfer links), made possible by how each link has been modelled in the network.

Table 1 shows the characteristics of the two link types present in the network and helps understand how the function adapts to the considered link. The maximum capacity is defined considering 500 seats per train, then multiplied by the number of daily trains (based on headway). In reality, capacity can vary significantly both between different lines and within the same IC line. This variation occurs because SBB has a fleet of different train types that can be coupled to increase maximum capacity based on demand. The transfer penalty includes both the walking time to the new platform (5 minutes) and waiting time. The travel time is calculated using actual link distances and an average speed of 80 km/h. The crowding discomfort is defined by the Bureau of Public Roads (BPR) function and acts as an amplifier of perceived travel time. In this case, the parameters alpha and beta have standard values but can be modified to give more or less importance to crowding.

$$t_a = (TT + \tau) \cdot \left[ 1 + \alpha \left( \frac{x}{\text{cap}} \right)^\beta \right] \quad (3)$$

where:

$TT$ : travel time of the link [min]

$\tau$ : penalty due to transfer [min]

$x$ : link flows [pax]

$\text{cap}$ : maximum train capacity [pax]

$\alpha = 0.15, \beta = 4$ : Parameters of the BPR function

This function considers only some elements that could be included when defining link costs, but has the advantage of being simple to define and modular based on the link type. This allows for easy addition of other elements. The alpha and beta parameters could also be modified according to the line or train type to create an increasingly realistic model.

|               | Travel Time $TT$                       | Transfer penalty $\tau$        | Capacity                     |
|---------------|--|--------------------------------|------------------------------|
| Train Link    | $\frac{\text{distance}}{\text{speed}}$ | 0                              | $500 \cdot \# \text{trains}$ |
| Transfer Link | 0                                      | $5 + \frac{\text{headway}}{2}$ | infinity                     |

Table 1: Visualisation of how links are defined in the network

#### 2.4.2 Static Traffic Assignment

To solve the static traffic assignment problem, the following standard algorithms were used:

- **Frank-Wolfe Algorithm:** The main algorithm for solving the problem was used with 100 iterations and a max gap threshold of  $1^{-4}$ .
- **All-or-Nothing Assignment:** Using the label correction algorithm to obtain the shortest path and identify the route to assign to each OD pair's demand
- **Label Correction Algorithm:** An efficient method for finding shortest paths in networks by systematically updating node labels until optimal routes are identified
- **Bisection Search:** This algorithm helps find the optimal solution, with the max gap threshold set to  $1^{-2}$

The problem was solved considering user equilibrium conditions, with the objective function defined according to Beckmann's Formulation.

$$\int_0^{x_a} t_a(u) du = x(TT + \tau) \left[ 1 + \alpha \left( \frac{x}{\text{cap}} \right)^\beta \right] \quad (4)$$

Figure 8 shows the value of the objective function during the iterations. It can be observed that the result converges very quickly toward the optimal solution.

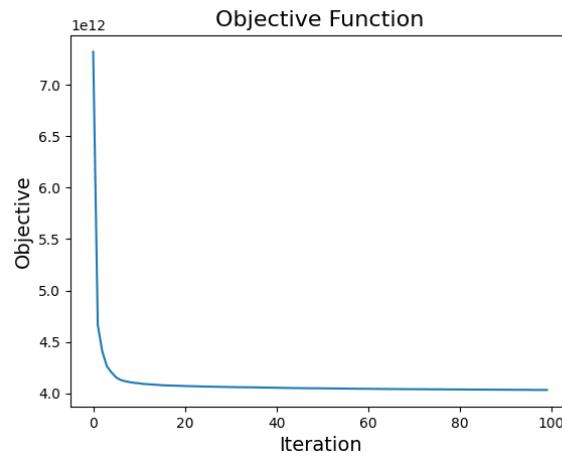


Figure 8: Convergence of the objective function

### 3 Findings

#### 3.1 Results

This section presents and discusses the results obtained with the previously specified model. It should be noted that the main objective of this project is not so much to concretely analyze the Intercity network, but rather to understand what outputs can be obtained and how they could be used. Furthermore, this model could be used to consider future scenarios. Initially, the baseline scenario described in previous sections was analyzed, followed by an evaluation of a potential new scenario.

##### 3.1.1 Baseline Scenario

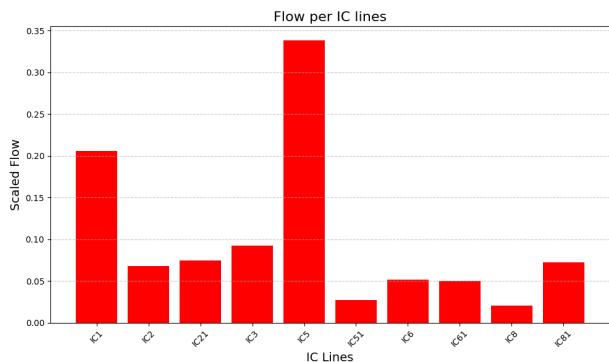
For this scenario, three different types of information were obtained, allowing analysis of both links and stations in the network:

- **Passengers per line:** This information identifies which lines are most frequently used. This data could be used to adapt the maximum capacity of trains operating on each line to avoid overcrowding or empty carriages. It could also inform frequency adjustments.
- **Transfer passenger flow per station:** This data consists of knowing how many people are at stations and need to change trains to reach their destination. This is important information as it enables station design and adaptation based on passenger flow, serving both safety and commercial purposes.
- **Passenger flow per link:** This data can be obtained in two different forms - one considering individual segments of a line and another aggregating the flow of all lines using the link. This data can be used to better understand line usage patterns, identify busiest sections, and adapt capacity or add interregional trains to enhance service on specific routes.

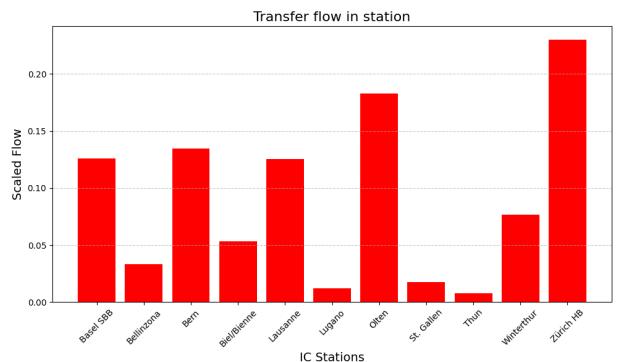
These outputs could also be used to inform passengers about train and station crowding levels.

##### 3.1.2 Case Study

Figures 9a and 9b show the number of passengers per line and transfer passenger flow per station respectively. Values have been normalized to clearly identify the most utilized lines. The analysis reveals that the most used lines are IC5 and IC1, while stations in major cities and central network hubs (Zürich, Bern, Olten, Lausanne) show the highest transfer passenger volumes.



(a) Passengers per line (normalized)



(b) Transfer passengers per station

Figure 9: Analysis of passenger distribution

Figure 10 shows the aggregated link flow for all lines, again demonstrating that links connecting major cities are the most utilized.

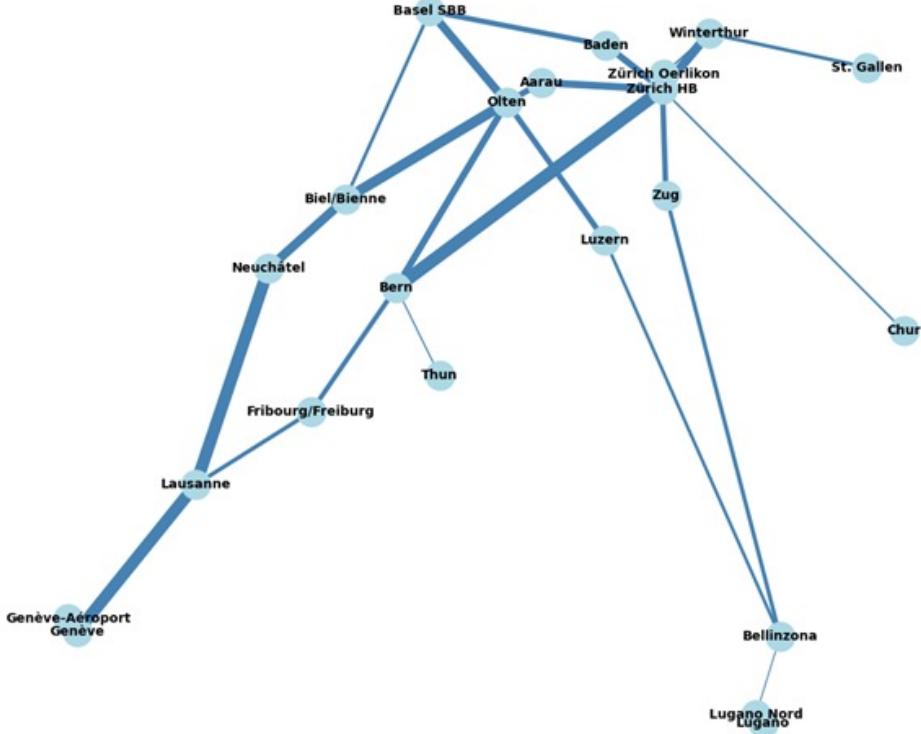


Figure 10: Aggregated link flow for all lines

### 3.1.3 Hypothetical Scenario

Another objective of this project was to test whether the model could be used to evaluate future projects and scenarios such as:

- Considering new OD pairs
- New major projects (e.g., Alptransit, Léman 2030)
- New train types (speed, capacity, comfort)
- New stations in the Intercity network
- New alternative lines (e.g., Swissmetro)

To evaluate these possibilities, a simple example inspired by the SwissMetro project was created. A high-speed link between Lausanne and Zürich stations was added to the network (Table ??). Defining its differentiating characteristics was straightforward thanks to the modular way links were defined in the network specification.

| Origin   | Destination | Speed [km/hr] | TT | Headway    | $\tau$ | Capacity                     |
|----------|-------------|---------------|----|------------|--------|------------------------------|
| Lausanne | Zürich HB   | 300           | 65 | 1 train/hr | 0      | $500 \cdot \# \text{trains}$ |

Table 2: High-speed link specifications

Using the same traffic assignment model as the baseline scenario, the results shown in Figure 11 were obtained. The figure compares the two scenarios, showing how adding a high-speed link slightly changes passenger flow distribution across different lines. This effect is particularly noticeable for line IC5, which serves the Lausanne-Zürich connection. Approximating total travel time with formula ?? shows approximately 14% reduction between the baseline and high-speed link scenarios.

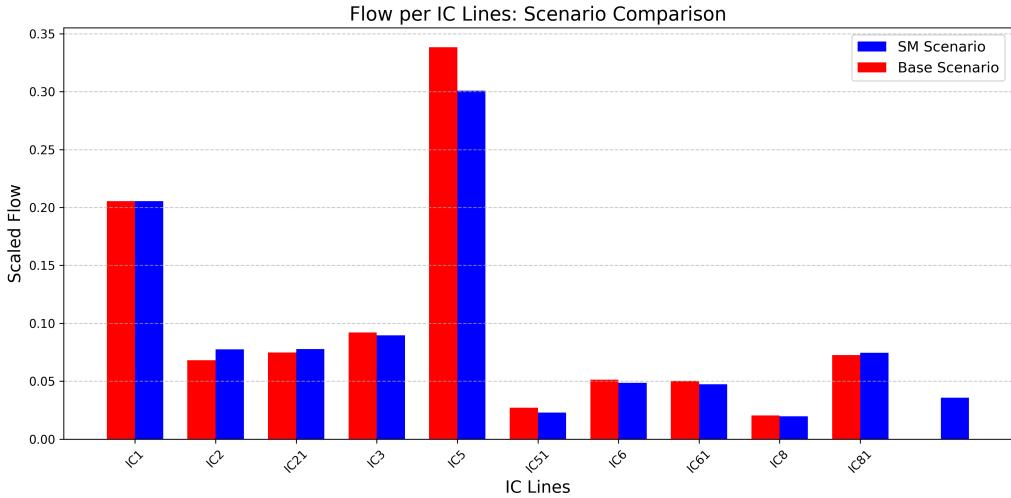


Figure 11: Comparison between baseline and high-speed scenarios

$$TT = \sum_a x_a \cdot t_a \quad (5)$$

### 3.2 Limitations & Next steps

This study faced several limitations primarily due to the availability and granularity of the data. The network data was limited to InterCity (IC) lines and included only a subset of stations, excluding many regional or local services that could influence passenger flows. Additionally, the passenger demand data consisted of fixed average daily counts without detailed origin-destination matrices or temporal disaggregation, such as peak and off-peak variations. The modeling approach also involved simplifying assumptions: travel links were treated uniformly with constant speeds, ignoring variability from operational conditions or train types. The generalized cost function used in the traffic assignment omitted important factors like ticket pricing, time of day, and trip purpose, which are known to affect passenger route choice. Future work should aim to incorporate richer data sources, including more comprehensive network coverage, dynamic demand patterns, and more detailed cost components, to improve the model's accuracy and applicability.

## 4 Conclusion

In conclusion, this project developed a flexible and modular framework for applying traffic assignment techniques to a railway network. By explicitly modeling line-specific paths and transfer penalties, the approach captures key aspects of passenger routing behavior within the InterCity network. The model's modular design allows easy modification of assumptions and parameters such as headways, transfer times, and link speeds, making it adaptable to different use cases and data availabilities. Overall, the project successfully achieved its goal of representing the railway network in a form compatible with classical traffic assignment algorithms, laying the foundation for further enhancements as additional data becomes available.

## 5 References

We acknowledge the use of ChatGPT for enhancing the clarity and precision of the text in this paper and for translation support. It was also used to help with code debugging. However, ChatGPT was not involved in data analysis or in deriving the project's conclusions.

Passenger count data: <https://data.sbb.ch/explore/dataset/anzahl-sbb-bahnhofbenutzer/>

Railway network geometry: <https://data.sbb.ch/explore/dataset/linie-mit-polygon/>

Population of cities: <https://www.bfs.admin.ch/bfs/en/home/statistics/population.html>

Official SBB website (timetables, routes, and IC line information): <https://www.sbb.ch>