Modelling Pedestrian Traffic in and Around Swiss Railways Stations

13th Symposium of the European Association for Research in Transportation

Alexander Erath[1], Michael van Eggermond[2]

[1] Professor for Transport and Mobility, University of Applied Sciences of Northwestern Switzerland, Switzerland

[2] Senior Researcher and Lecturer, University of Applied Sciences of Northwestern Switzerland, Switzerland

SHORT SUMMARY

Railway stations are evolving into multifunctional hubs, offering shopping, leisure, and services alongside transport. This shift calls for pedestrian infrastructure planning that goes beyond optimizing transfers to consider the station's role as a destination. Pedestrian traffic is typically modeled by examining trip generation and link-based pedestrian volumes, using factors like population density, land use diversity, and network centrality.

This study analyzed pedestrian traffic at Lucerne and Uster stations in Switzerland. Lucerne serves a mix of commuters and tourists, while Uster primarily caters to regional travelers. Models built with public data explained 60% of the pedestrian traffic variance in Lucerne and 54% in Uster. However, differences in parameter estimates suggest limited transferability between stations.

The findings underscore the importance of detailed, location-specific data to improve pedestrian traffic models. Additionally, the models show potential for evaluating how improvements in pedestrian infrastructure can better integrate stations into their surroundings, as well as assessing the impact of new urban development on pedestrian traffic and its potential to activate ground floors for commercial use.

Keywords: Cycling and walking behaviour and design, direct demand model, pedestrian flow modelling, transport network modelling, railway stations, urban development.

1. INTRODUCTION

Railway stations are no longer merely points for transferring between different modes of transport; they are increasingly evolving into mobility hubs and destinations for diverse activities such as shopping, personal services, and leisure. Consequently, the planning of pedestrian infrastructure in and around stations should go beyond optimizing passenger transfers and also account for the station's role as a multifunctional destination.

Pedestrian traffic is typically modelled using statistical regression in the form of direct demand models, which incorporate two interrelated dimensions:

- **Trip generation and attraction:** This refers to the volume of foot traffic generated or attracted by individual buildings, retail stores, public transport stops, and other points of interest (POIs) or activity hubs.
- **Link-based pedestrian volume:** This represents the pedestrian volume that traverse a specific link of the pedestrian network within a defined time period.

Cooper et al. (2019), Lerman et al. (2014) und Van Eggermond et al., (2022) identify a comprehensive set of variables commonly used in such modelling approaches. These include population and employment density (potentially differentiated by age or employment type), absolute population or employment figures, land use diversity, characteristics and locations of individual walking routes, vehicle traffic factors (e.g., speed and composition), availability of transportation options, and proximity to points of interest.

Network topological indicators, which describe the structural properties of each link of a network, can provide valuable insights into pedestrian traffic volumes. The most used measure is betweenness centrality, also referred to simply as centrality. This metric calculates, for each connection within a network, the number of shortest paths between all points in the network that pass through that specific connection.

When computing betweenness centrality, various distance measures can be applied to determine the shortest route. These measures may involve metric distances or angular distances, the latter referring to the cumulative changes in direction along a route. Angular distance is frequently used in Space Syntax applications (Bafna, 2003) and is typically expressed either as the number of directional changes or as the total of degree values for those changes. Recent studies suggest that that turns, as well as other environmental qualities of a route, should be considered in addition to, not in lieu of, distance (Sevtsuk & Basu, 2022).

The most effective modelling approaches for describing and predicting pedestrian volumes in urban contexts integrate data on boarding and alighting passengers with information on the attraction potential and visibility of commercial establishments, alongside detailed land use data. (Lerman et al., 2014; Porta et al., 2009; Sevtsuk, 2021). Pedestrian networks and destinations relevant to walking connect multiple building levels within stations and surrounding areas. It is essential to account not only for horizontal and vertical distances but also for the angular variations of the different route options between all origins and destinations. (Cooper et al., 2019; Van Eggermond et al., 2022).

Models for describing link-based pedestrian traffic volumes support the planning of pedestrian infrastructure by assessing the impact of various design and development scenarios. They allow for prioritising design interventions and provide a quantitative framework for evaluating station accessibility today and for future scenarios (Sevtsuk et al., 2024). Furthermore, these models help in assessing the suitability of ground floor uses as well as evaluating the suitability of public spaces for recreational use by analysing pedestrian volume and pedestrian trip purposes.

So far, most studies on pedestrian traffic modelling have focused on urban quarters (e.g., (Sevtsuk, 2021; Sevtsuk & Kalvo, 2024)) or even smaller cities Lerman et al. (2014). Fewer studies have explored smaller spatial units such as urban mobility hubs (e.g. Van Eggermond et al., 2022) or university campuses (e.g Zhang et al., 2024), which utilise geographic data with finer spatial resolution. This includes pedestrian networks (e.g., sidewalks on both sides of streets, pedestrian crossings, 3D networks), as well as trip generators and attractors (e.g., individual building entries and shops within larger structures such as railway stations). To our knowledge, there has yet to be an attempt to model pedestrian flow at such a high spatial detail in and around railway stations in Europe.

The research presented in this paper aimed to answer the following research questions:

- To what level of accuracy and detail (e.g., by trip purpose or age group) can pedestrian traffic in and around Swiss railroad stations be described using current modelling approaches and available data sources?
- How can existing approaches to model pedestrian traffic be improved for application in railroad stations and their surroundings?

- Can a model to pedestrian traffic be meaningfully transferred to describe pedestrian traffic of a different railways station?

2. METHODOLOGY

Study area

The study was conducted at two different types of railway stations to evaluate the method's applicability across varying spatial settings and to assess the transferability of the model results. With approximately 144,000 daily users, Lucerne is Switzerland's fifth most frequented railway station (SBB, 2024). It caters to a diverse range of users, including commuters, national and international tourists, and serves as a regional hub for nearby leisure activities and cultural institutions. Adjacent to the railway station, a bus hub serves local and regional bus lines, complemented by a boat landing stage. As illustrated in Figure 1, the surrounding urban fabric is characterised by a mixed-use quarter to the west, featuring a grid-like street network and buildings dating back to the early 19th century. To the north lies the historic old town, separated by the river Reuss, while to the east is an urban quarter undergoing redevelopment, with a mix of educational institutions, office buildings, and residential housing. The station features 15 rail tracks and approximately 6,000 m² of commercial space, about 70% of which is located on a basement level. This level provides direct connections to the city and the bus stop to the northwest, eliminating the need to cross roads at ground level.

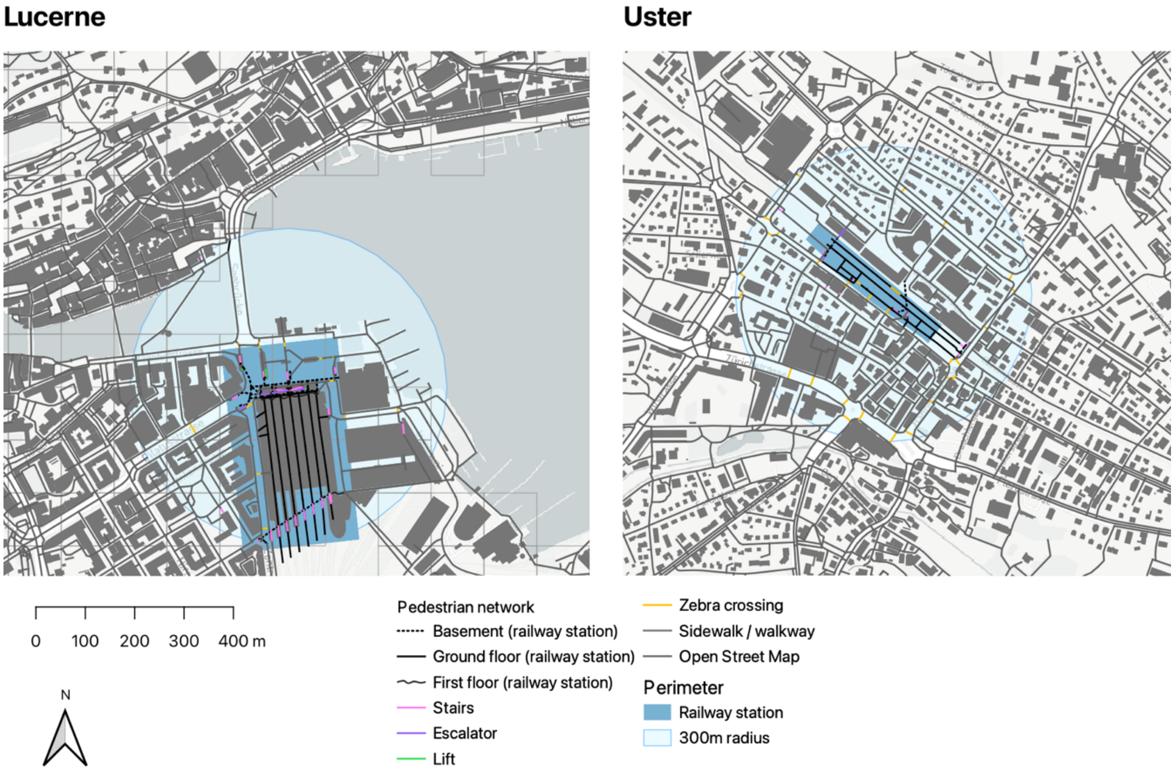
With 36’000 inhabitants, Uster is the third largest town of the canton of Zurich. It’s railway station connects four S-Bahn lines with the local and regional bus network but is not served by any regional or national rail lines. The railway station attracts approximately 27,500 daily users, around five times fewer than Lucerne’s station. It primarily serves regional travellers, including commuters working in other cities, students attending regional educational institutions in Uster, and residents of Uster along with its predominantly regional visitors. South of the railway station lies Uster's commercial centre, featuring two major shopping centres and Poststrasse, a main street lined with numerous shops and public-facing ground-floor uses. To the north, a residential urban quarter predominates, while to the west, two large areas are slated for redevelopment to support higher density and mixed-use developments.

Pedestrian network

For both stations, the pedestrian networks were developed using a combination of a newly designed pedestrian network for the area within the railway station and enriched OpenStreetMap (OSM) data (see Figure 1). Inside the railway station, multiple floor levels were differentiated, with three types of vertical connections—escalators, stairs, and lifts. Each vertical connection was assigned a distinct generalized cost based on previous research (e.g. Olszewski & Wibowo, 2005).

Within a 300-metre radius around the railway station, the pedestrian network was redrawn using OSM data, incorporating pavements on both sides of major roads and pedestrian street crossings. Beyond this range, up to a radius of 2 kilometres, the pedestrian network was directly derived from OSM data. This includes walkways through parks but generally does not feature separate pavements. Instead, a centreline approach was used for these areas.

Figure 1: Railway stations and pedestrian network



Trip generation and attraction

The data used to describe pedestrian trip generation and attraction for a typical weekday stems from three key data sources:

- Outside the railway station: National spatial statistics on the distribution of the resident population and workplaces by workplace type.
- Within the railway station sales transaction data from stores.
- To account for boarding, alighting and transferring public transport passenger: An activity- and agent-based national transport model.

Outside the railway station, data from the Swiss national statistics office on residential population and workplaces, provided at a spatial resolution of a 100-metre raster, is used to determine the number of trips generated and attracted within each raster cell. To achieve this, pedestrian trip generation and attraction factors, along with the mode share of pedestrian trips (including public transport access and egress stages), are derived from the national household travel survey. The number $n_{i,k}$ of pedestrian trips generated or attracted by a workplace or a person at residence i for trip purpose k is calculated using the following formula:

$$n_{i,k} = w_i \cdot a_{i,k} \cdot p_{F,k}$$

where w_i represents the number of trips generated per workplace or per person at the place of residence, $a_{i,k}$ is the share of trip purpose k, and $p_{F,k}$ is the share of pedestrian traffic for stages starting or ending at the location.

Workplaces also generate and attract walking trips by visitors. To describe the volume of pedestrian traffic V_i by these visitors, volume rates differentiated by workplace type are used according to the formula below.

$$V_i = FTE_{i,k} \cdot w_k$$

Where $FTE_{i,k}$ describes the number of full-time equivalents at location i of workplace type k, and w_k is the generation rate of pedestrian trips by visitors of workplace type k. The values used for w_k were derived from common-sense assumptions to represent the number of clients or people a typical worker would serve in a day. These values range from 10 (e.g., travel agency, hairdresser) to 40 (e.g., education), with the highest being 80 for retail stores.

Within the railway station, the number of generated and attracted pedestrian trips is derived from sales transaction data. This data, available to the Swiss Federal Railways, is collected from commercial areas rented out within the station. To estimate pedestrian trip generation and attraction from sales transactions, an intercept survey was conducted, yielding 234 valid responses from Lucerne and 204 from Uster. The survey aimed to determine the proportion of visitors who browse stores without making a purchase. Additionally, the number of visitors was manually counted at approximately five stores and compared to their transaction data. This comparison allowed for the derivation of different visitor-to-transaction ratios for various store types, as follows: clothing stores (4.6); health and wellness (1.8); accessories (1.7); restaurants, medical clinics, bookshops (1.3); drinks, pharmacies, fast food outlets, electronics stores (1.2); supermarkets, convenience stores and kiosks (1.1).

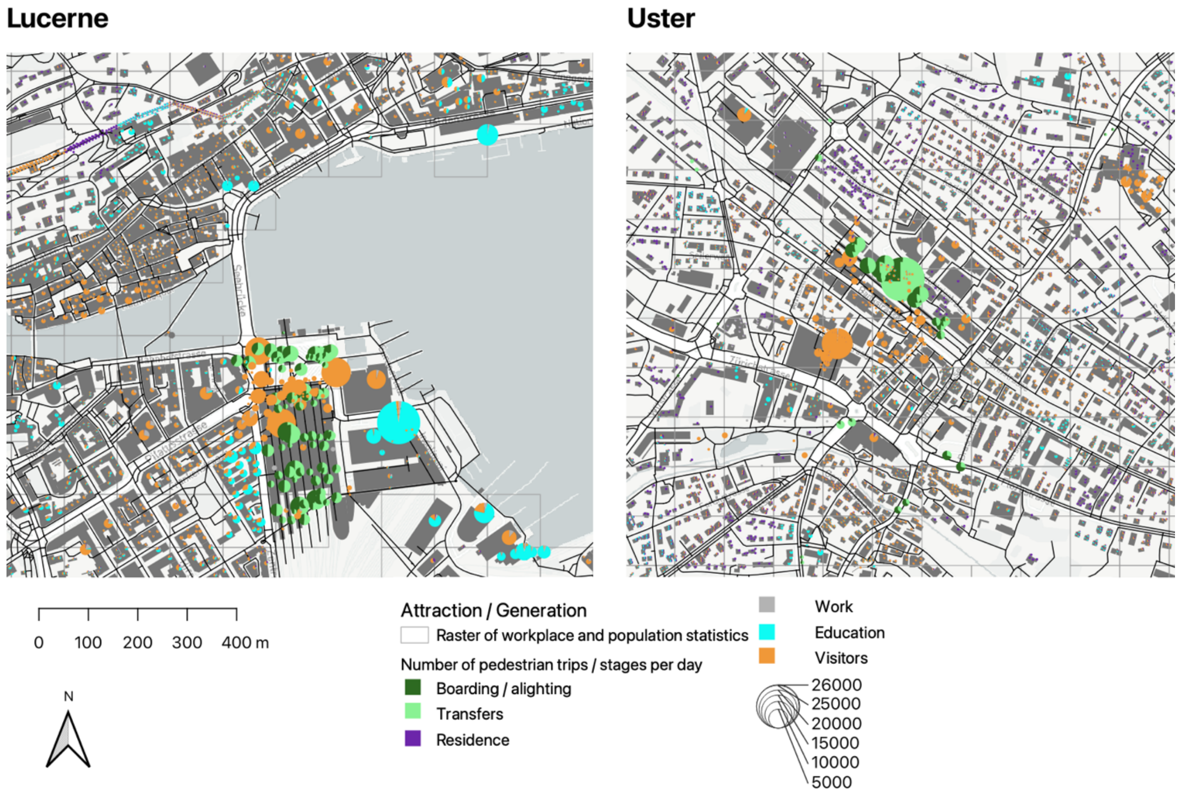
However, not all store operators agreed to share their data with the research team. As a result, estimates for pedestrian trip generation and attraction were also defined for other stores and additional activity generators within the station, such as public toilets, car parks, a police station, and the local public transport operator's information booth.

Data on the number of boarding and alighting as well as transferring passengers are derived from SIMBA MOBi, an activity- and agent-based traffic model developed by SBB for whole Switzerland (Scherr et al., 2020). The stage-based traffic simulation enables the derivation of information on the origins and destinations of all trips made on foot in and around a station.

Using the 2020 model, datasets were created for the Uster and Lucerne stations, encompassing all walking and cycling stages that start or end at each respective station. Stages at the beginning or end of a train journey are assigned to the platform specified in the timetable and are distributed evenly along the platform at 50-metre intervals. All other public transport stages are mapped to the level of individual public transport stops.

Figure 2 illustrates the number of trips generated or attracted for each considered point of interest (POI), categorized by type of pedestrian activity.

Figure 2: Trip generation / attraction



Pedestrian counts

The pedestrian count data is derived from two key sources:

- Data from 26 automatic counters installed at Lucerne and 10 at Uster railway stations. These counters are located at various station entrances, selected vertical connections within the station, and specific stores.
- Manual pedestrian counts conducted specifically for this research.

Data on person frequencies from automatic counting points are stored in one-minute intervals. For this research, the data was aggregated into 60-minute intervals, including averages and standard deviations for a typical week in May 2023, as well as an annual average for the entire year 2022. Using this data, manually collected person frequencies were validated and scaled. The derived yearly average weekday value serves as the basis for modeling route-related pedestrian traffic volumes.

To collect pedestrian traffic volumes at as many cross-sections as possible with limited personnel resources, counts were conducted at individual cross-sections for 10 to 15 minutes at various times, based on experience from an earlier research project on pedestrian counts (Pestalozzi et al., 2022). In Uster, 51 counting cross-sections were counted, in Lucerne 50 counting cross-sections. The counting cross-sections were selected based on the network indicators (betweenness and integration), as well as the location of permanent counting points. Attention was also given to ensuring that pedestrian flows, both concentric to the station and at various distances from the station, were adequately covered.

Network indicators

Various network indicators are calculated, which differ in the distance metric used for routing and the weights applied to the segments. A distinction was made between the following distance metrics

- Metric distance: The length of a link is counted as the distance metric.
- Angular distance: Each angle change is counted as a distance.
- Combination of angular and metric distance: In this case, metric and angular distances are combined, with both weighted at 50 percent.
- Random perception of distance: Link costs are slightly varied to allow for slightly different routes. This method assumes that pedestrians have varying preferences and may not possess complete knowledge of the network. It is particularly useful when multiple similar alternatives exist between an origin and a destination.

Different weighting approaches are also distinguished:

- Link-based: Each link is counted as 1.
- Aggregated Trip Generation Rates: Trips generated across all purposes are assumed to represent pedestrian flow. The advantage of this approach is that fewer network indicators need to be calculated. However, the disadvantage is that differences between trip purposes cannot be disaggregated in the model.
- Disaggregated Trip Generation Rates: Trip generation rates are differentiated by "public transport boarding or alighting", "public transport transfers", "visitors", "workers", and "residents." The advantage of this method is that differences between trip purposes can be disaggregated in the model. The disadvantage is that additional network indicators must be calculated, and more parameters need to be estimated. The differentiation between workers and education was abandoned during the model development phase, as no significant differences in related parameters were identified.

All network indicators are calculated for different distance categories (up to 200 m, 200–400 m, 400–800 m, 800–1200 m). This means that for each link, network indicators are calculated as distinct independent variables, representing the number of trips within the distance classes up to 200 meters, 200 to 400 meters, and 400 to 800 meters passing through the link. The assumption is that longer trips are less important for predicting pedestrian flows because they occur less frequently. By differentiating network indicators by distance classes, this effect can be represented in the statistical model.

Modelling pedestrian demand

To predict pedestrian demand based on the calculated network measures we assume that the following relationship holds:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \varepsilon$$

where variable Y (pedestrian demand) can be described as function of the parameters β and the independent variables X.

Commonly ordinary least squares (OLS) is used to solve the equation above. One issue that can occur with OLS is multicollinearity, which is likely to exist in this specific case. First, there will exist a correlation between the network measures generated for the different distance bins. For instance, it might be that links with a high accessibility to retail within 200 meters will also have a high accessibility to retail between 200 and 400 meters. Second, there will exist a correlation between network measures for different land-use pairs. With ridge regression, an extra parameter λ is added to the function that is multiplied with the β parameters to be estimated:

$$L_{ridge}(\hat{\beta}) = \sum_{i=1}^{n}(y_i - x_i^j\hat{\beta})^2 + \sum_{j=1}^{m}\hat{\beta}_j^{\,2} = \|y - X\hat{\beta}\|^2 + \lambda\|\hat{\beta}\|^2$$

The conducted pedestrian counts are characterized by several locations with high counts and many counts with much fewer counts. Locations with high counts generally occur much less than locations with low counts. To this end, it is necessary to weight the pedestrian counts. This weight is passed to the estimation of the parameter estimates, such that locations with high counts are considered relatively less important than locations with low counts. This weight is calculated with the following formula (Cooper, 2018):

$$weight(y, \lambda) = \frac{y^{\lambda}}{y}$$

3. RESULTS AND DISCUSSION

Models were estimated for Lucerne and Uster using different sets of variables, initially using metric distance as routing metric for both cities.

The first model specification assumed a uniform weight of 1 to all links. For Lucerne, this model shows that link density, both within and outside the station, is a good proxy for pedestrian volumes, suggesting that the density of connections between nodes significantly influences pedestrian traffic. In contrast, for Uster, link density does not adequately describe pedestrian volumes, likely due to the lower network density compared to Lucerne.

The second model described pedestrian volumes based on public transport passengers and points of interest (POIs). In Lucerne, the model indicates that public transport transfers and POIs within 200–400 meters and 400–800 meters are crucial for determining pedestrian volumes. For Uster, the model highlights the importance of paths between POIs and public transport passengers within 200 meters and 200–400 meters, with public transport transfers also playing a significant role.

The third model focused on betweenness indicators between public transport passengers and various POI types, without including betweenness between different POIs. In Lucerne, this model shows that public transport transfers and paths within 200–400 meters and 400–800 meters are the most important variables. In Uster, similar to Lucerne, the most significant factors are public transport transfers and paths within 200 meters and 200–400 meters, with the strongest explanatory power coming from the betweenness between public transport passengers and residential areas or workplaces.

The fourth model included all betweenness indicators and differentiates them by the types of traffic volumes. For Lucerne, the inclusion of these additional betweenness indicators results in a higher model fit, improving predictive accuracy. In Uster, the inclusion of these additional indicators does not significantly improve the model's explanatory power compared to Model 3, although it does enhance the fit compared to Model 2.

Subsequently, for both cities different distance metrics were evaluated. For Lucerne metric distance yielded the highest model goodness-of-fit. For Uster, randomisation of link costs, combined with a combination of angular and metric distance yielded the highest model fit. More specifically, angular change was set to 30% of the costs and metric distance to 70% of the costs of a route. Link costs were varied in 30 runs using a normal distribution with mean 1 and standard deviation of 1, and truncated to fall between 0.1 and 10. Note that this randomization is applied to links and not to routes.

Figure 3 shows the standardized model coefficients for Lucerne, highlighting the most important variables influencing pedestrian footfall. The most significant factors are public transport passenger volumes to all links within 200 and 400 meters, as well as public transport transfers to links within the same distances. Following these, public transport passenger volumes to shopping areas emerge as the next most influential variable.

Figure 3: Standardized Model Coefficients for Lucerne Model

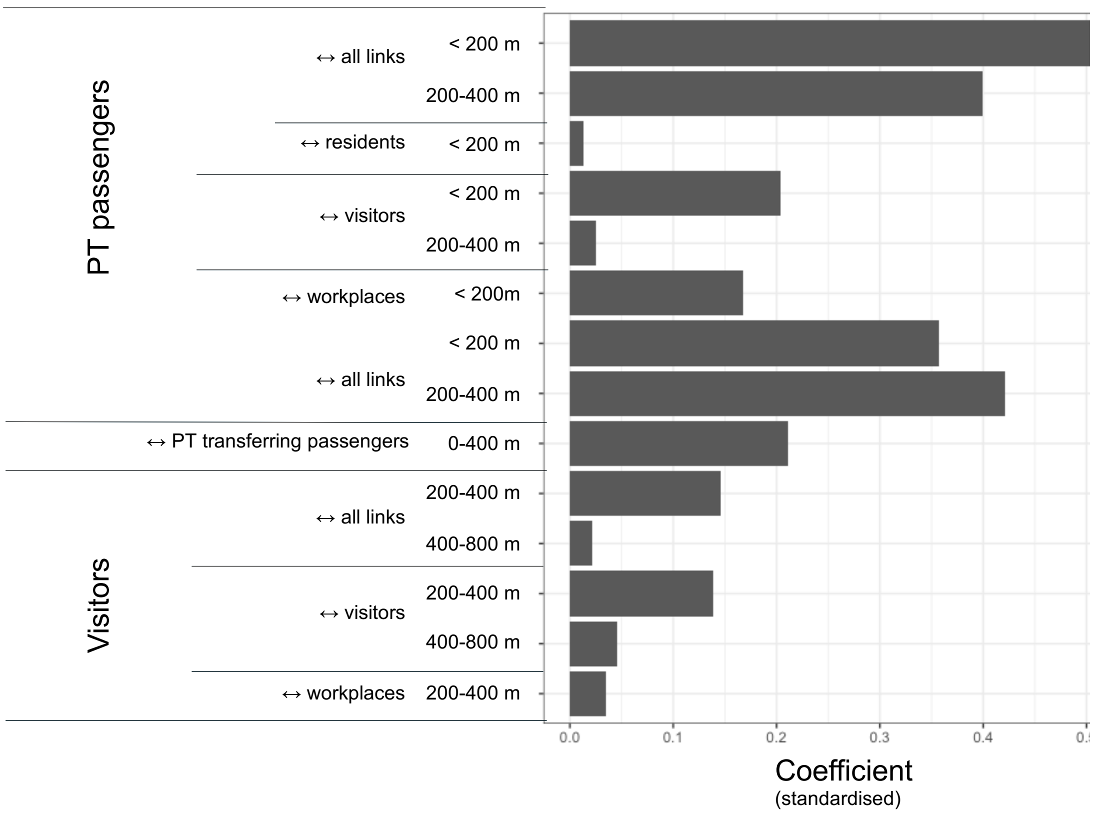


Figure 5 presents the normalized model coefficients for Uster. Compared to Lucerne, pedestrian footfall in Uster is primarily driven by the proximity of residential areas and workplaces to the train station. The next most important factors include the number of visitors to the train station itself. In addition to trips generated and attracted by the station, footfall is also influenced by visitors to various destinations around the station, highlighting the broader impact of surrounding attractions on pedestrian traffic.

Scatterplots showing the counts versus the predicted volumes are shown in Figure 5

Figure 4: Standardised Model Coefficients for Uster Model

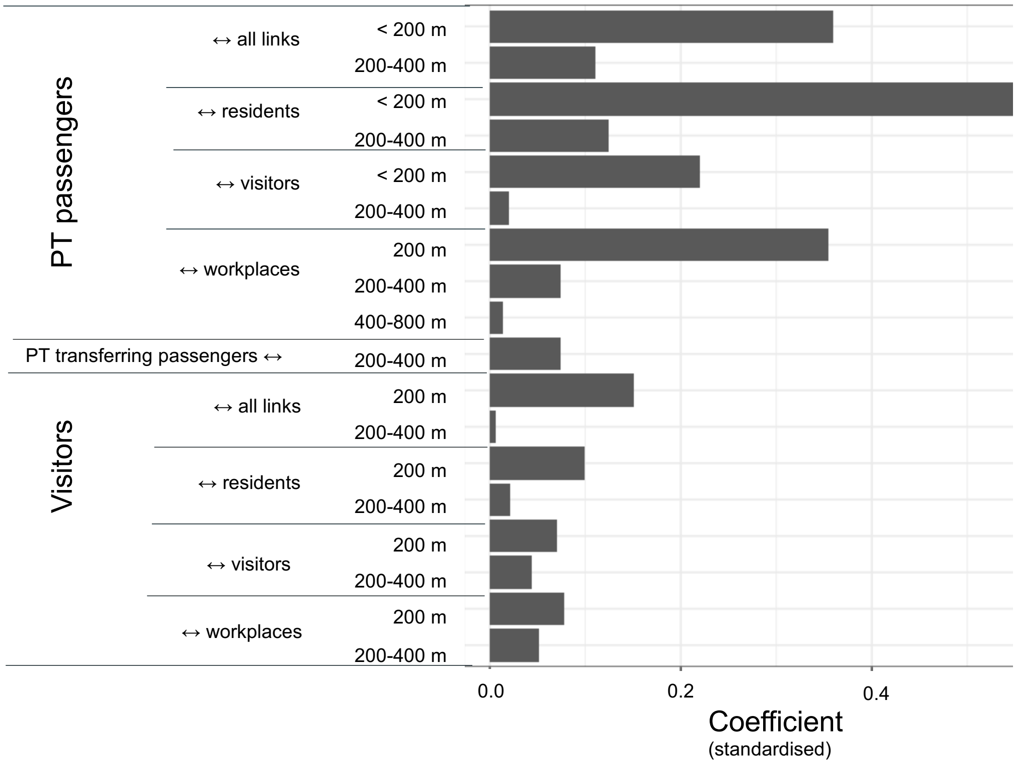
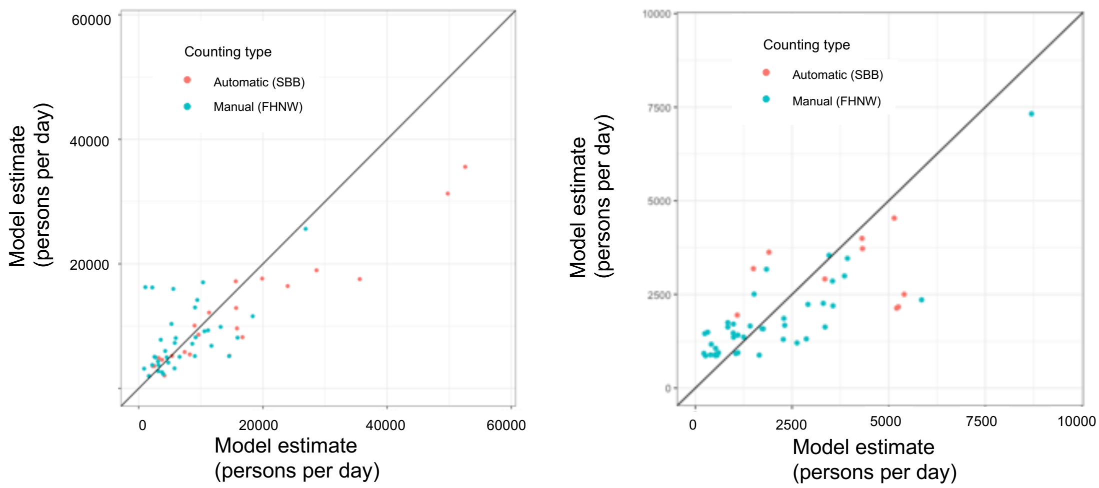


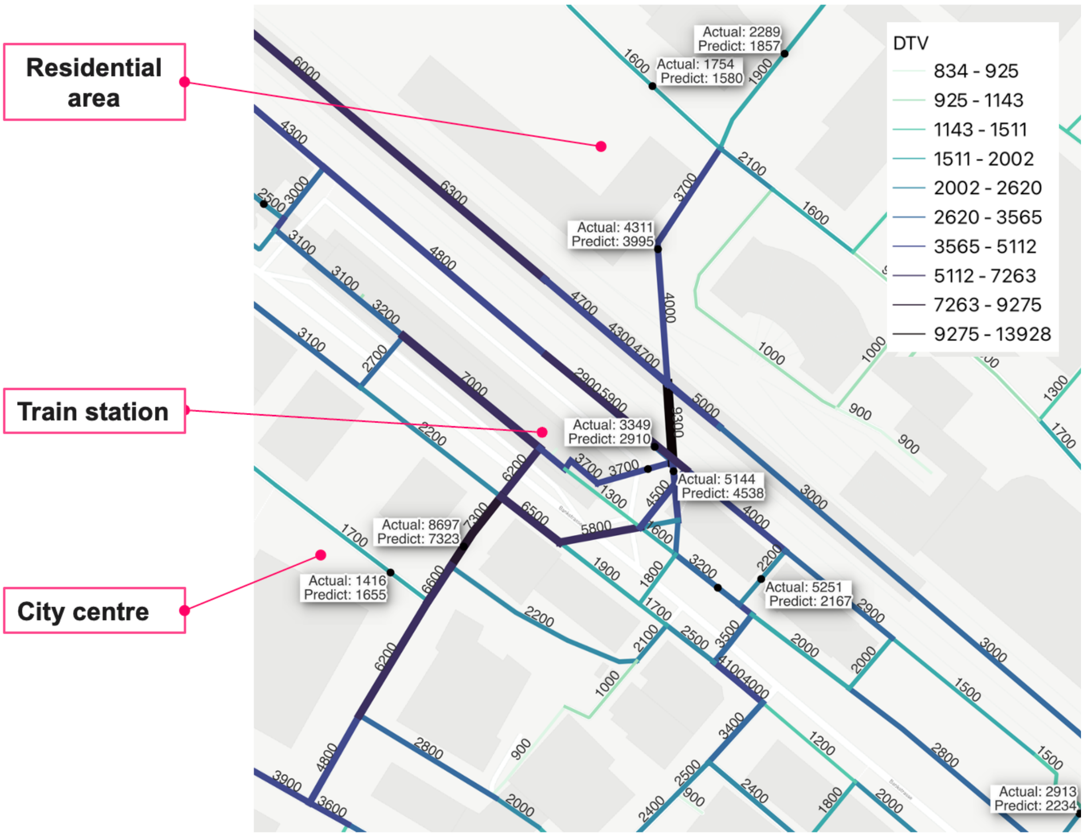
Figure 5: Comparison of pedestrian traffic counts and modelled data

Lucerne (R^2=0.6) **Uster (R^2=0.54)**



A visual impression of the model results for the case of Uster is shown in Figure 6. In the pedestrian underpass predicted flows match the counts well with a relative deviation between 7% to 11% depending on the selected cross-section. A relative deviation of 15% is observed along the street leading from the railway station towards the city center and the main shopping area to the southwest.

Figure 6: Predicted daily pedestrian flow in Uster in and around the train station



4. CONCLUSIONS

The conclusions drawn from the research work correspond to the research questions formulated in Chapter 1.

A model to describe pedestrian traffic flow in and around railways station in Switzerland, calibrated using cross-sectional count data, has been developed. This model shows that, by utilizing publicly available spatial data on population and workplaces, sales transaction data from stores within the station, and SBB data on public transport passenger volumes, pedestrian traffic in and around stations can be modeled with reasonable accuracy. The models developed for Lucerne and Uster stations explain 60% and 54% of the observed variance, respectively and allow for the differentiation of route-related pedestrian traffic volumes based on trip purpose.

Although the model results for applications in Lucerne and Uster show a similar structure, the parameter estimates can vary substantially, and the set of variables with significant parameter estimates differs. The findings underscore the importance of detailed, location-specific data to enhance pedestrian traffic models and indicates that models should not be transferred between different station types. A conclusive statement on the transferability of these models to other stations of the same type cannot be made based on this work. It is recommended that this question be explored in future research.

No data is available on sales transactions outside train stations. Therefore, the number of visitors was estimated based on the number and type of workplaces reported per hectare grid. Counting the number of pedestrian visitors at two supermarkets in Uster has shown that retail footfall rates are highly influenced by the location of the store. It can be assumed that this applies to the number of visitors to other types of workplaces as well. Similarly, we expect pedestrian mode share to vary substantially depending on the location. Currently, limited data is available to describe visitor numbers, often relying on small sample sizes. As a result, research into location-specific pedestrian traffic rates for workplaces holds significant potential for improving pedestrian traffic models.

The models are sensitive to changes in both the structure of the pedestrian network and trip generators and attractors. As a result, they provide the ability to quantify pedestrian traffic volumes to assess design scenarios, such as adding new infrastructure (e.g. underpasses) to integrate stations more effectively into the urban environment, evaluating changes in pedestrian volumes induced by new urban developments and forecasted increases in public transport ridership. Case study applications for these purposes have been prepared but are not included in this short paper. They are, however, planned to be discussed in a full paper.

ACKNOWLEDGEMENTS

This research was funded by the research fund of the Swiss Railways (SBB / CFF / FSS).

REFERENCES

Bafna, S. (2003). Space Syntax: A Brief Introduction to Its Logic and Analytical Techniques. *Environment & Behavior*, *35*(1), 17–29. https://doi.org/10.1177/0013916502238863

Cooper, C. H. V. (2018). Predictive spatial network analysis for high-resolution transport modeling, applied to cyclist flows, mode choice, and targeting investment. *International Journal of Sustainable Transportation*, 1–11. https://doi.org/10.1080/15568318.2018.1432730

Cooper, C. H. V., Harvey, I., Orford, S., & Chiaradia, A. J. F. (2019). Using multiple hybrid spatial design network analysis to predict longitudinal effect of a major city centre redevelopment on pedestrian flows. *Transportation*. https://doi.org/10.1007/s11116-019-10072-0

Lerman, Y., Rofè, Y., & Omer, I. (2014). Using Space Syntax to Model Pedestrian Movement in Urban Transportation Planning: Using Space Syntax in Transportation Planning. *Geographical Analysis*, *46*(4), 392–410. https://doi.org/10.1111/gean.12063

Olszewski, P., & Wibowo, S. S. (2005). Using Equivalent Walking Distance to Assess Pedestrian Accessibility to Transit Stations in Singapore. *Transportation Research Record*, 8.

Pestalozzi, C., Bucheli, D., & Sauter, D. (2022). *Empfehlungen zur Zählung des Fussverkehrs* (No. 1720). Bundeamt für Strassen.

Porta, S., Strano, E., Iacoviello, V., Messora, R., Latora, V., Cardillo, A., Wang, F., & Scellato, S. (2009). Street Centrality and Densities of Retail and Services in Bologna, Italy. *Environment and Planning B: Planning and Design*, *36*(3), 450–465. https://doi.org/10.1068/b34098

SBB. (2024). *Bahnhöfe*. SBB Zahlen & Fakten. https://reporting.sbb.ch/bahnhoefe

Scherr, W., Manser, P., & Bützberger, P. (2020). SIMBA MOBI: Microscopic Mobility Simulation for Corporate Planning. *Transportation Research Procedia*, *49*, 30–43. https://doi.org/10.1016/j.trpro.2020.09.004

Sevtsuk, A. (2021). Estimating Pedestrian Flows on Street Networks: Revisiting the Betweenness Index. *Journal of the American Planning Association*, 1–15. https://doi.org/10.1080/01944363.2020.1864758

Sevtsuk, A., & Basu, R. (2022). The role of turns in pedestrian route choice: A clarification. *Journal of Transport Geography*, *102*, 103392. https://doi.org/10.1016/j.jtrangeo.2022.103392

Sevtsuk, A., & Kalvo, R. (2024). Modeling pedestrian activity in cities with urban network analysis. *Environment and Planning B: Urban Analytics and City Science*, 23998083241261766. https://doi.org/10.1177/23998083241261766
Sevtsuk, A., Kollar, J., Pratama, D., Basu, R., Haddad, J., Alhassan, A., Chancey, B., Halabi, M., Makhlouf, R., & Abou-Zeid, M. (2024). Pedestrian-oriented development in Beirut: A framework for estimating urban design impacts on pedestrian flows through modeling, participatory design, and scenario analysis. *Cities*, *149*, 104927. https://doi.org/10.1016/j.cities.2024.104927
Van Eggermond, M., Mavros, P., Erath, A. L., Helle, V., Acebillo, P., & Xu, S. (2022). Forecasting district-wide pedestrian volumes in multi-level networks in high-density mixed-use areas. *2022*. Space Syntax Conference, Bergen, Norway.