Insights Into the Correspondence Between Land Use Pattern and Spatial Distribution of Rail Transit Services

Zhongqi Wang1 · Qi Han1 · Bauke de Vries1 · Li Dai1

Received: 19 June 2020 / Accepted: 20 April 2021 / Published online: 1 May 2021
© The Author(s) 2021

Abstract
The identification of the relationship between land use and transport lays the foundation for integrated land use and transport planning and management. This work aims to investigate how rail transit is linked to land use. The research on the relationship between land use and rail-based transport is dominated by the impacts of rail projects on land use, without an in-depth understanding of the reverse. However, it is important to note that issues of operation management rather than new constructions deserve greater attention for regions with established rail networks. Given that there is a correspondence between land use patterns and spatial distribution of heavy railway transit (HRT) services at such regions, the study area (i.e., the Netherlands) is partitioned by the Voronoi diagram of HRT stations and the causal relationship between land use and HRT services is examined by structural equation modeling (SEM). The case study of Helmond (a Dutch city) shows the potential of the SEM model for discussing the rail station selection problem in a multiple transit station region (MTSR). Furthermore, in this study, the node place model is adapted with the derivatives of the SEM model (i.e., the latent variable scores for rail service levels and land use characteristics), which are assigned as node and place indexes respectively, to analyze and differentiate the integration of land use and HRT services at the regional level. The answer to whether and how land use affects rail transit services from this study strengthens the scientific basis for rail transit operations management. The SEM model and the modified node place model are complementary to be used as analytical and decision-making tools for rail transit-oriented regional development.

Keywords  Land use · Rail transit · Land use and transport integration · Regional development

*Zhongqi Wang
wangzhongqi90@163.com

Extended author information available on the last page of the article
Introduction

Rail-based transport, as an efficient mass transit system, has competitive advantages over road-based transport, which is beset by environmental pollution, economic loss, safety and health risks arising from traffic congestion. As rail transit not only reduces traffic congestion but the related investment can be cost-effective (Litman, 2007), it is growing in importance to the sustainable development in highly populated and economically active regions, such as the megacities in the world (Hayashi et al., 2011). Since integrating land use and transport has gained wide acceptance in urban planning and management (Acheampong & Silva, 2015), the ongoing studies continuously review the relationship between land use and rail transport, most of which focus on the impacts of rail projects on land use. However, after the lengthy processes of land cover change (e.g., the conversion of unbuilt areas to built-up areas) and rail line construction, issues of operations management rather than new constructions deserve greater attention for regions with established rail networks. The answer to whether and how land use affects rail transit service will strengthen the scientific basis for rail transit operations management. The supply of rail transit services, given the stable level of technology and operation management for a certain period of time, is largely subject to land use driven demand. It can be seen from regions with established rail networks that discrepancy exists in the spatial distribution of rail transit services (e.g., the number and the capacity of rail stations), which generally corresponds to the spatial distribution of human activities. For example, a major city always serves as a regional transport hub, and low-ranked stations are usually located in relatively remote or sparsely populated areas. This study adopts a spatial organization perspective on how the spatial distribution of rail transit services is linked with land use patterns. The relationship between land use and rail transit service levels can be used for the discussion on facility optimization (see 4.2) and spatial development (see 4.3).

Land use characteristics are important determinants of rail transit use, together with socio-economic and demographic characteristics (Jun et al., 2015; Limtanakool et al., 2006; Loo et al., 2010). The linkages between land use and rail transit are commonly established through rail stations and their immediate vicinity, such as the surrounding area of stations within a fixed distance (i.e., catchment area) in demand/ridership estimation. However, pre-defined surroundings of rail stations could result in spatial fragmentation and discontinuity, and overlaps in areas with a high station density. Moreover, the geographical scope of land use analysis affiliated to rail stations cannot reflect the impact of land use patterns on the rail transit provision. The decision to open or close a station and the station capacity planning affect not only the station and its surrounding area but also the nearby stations and their surrounding areas. Now that a rail station cannot be isolated from other stations (especially nearby stations), space divisions, which are linked with rail stations, should be interconnected. Accordingly, the spatial relation between rail stations and space divisions is built on the Voronoi diagram in this study to analyze the relevance of the spatial configurations of land use and rail network.
The Netherlands is a highly developed region with a dense rail network. Useful lessons on the connection between land use and rail transit can be drawn from the situation in the Netherlands, which may shed light on densely populated areas undergoing large-scale and fast urbanization, e.g., growing megacities in Asia. In this study, we limit ourselves to heavy railway (HR) regarding rail-based transport, and ‘land use’ is used as a broad term covering human activities, demographics, and socio-economics. Land use characteristics are extracted from Voronoi cells of heavy railway transit (HRT) stations. Structural equation modeling (SEM) was used to test and define the relationship between land use characteristics and HRT service levels. With the knowledge of the land use and HRT nexus, the issue of whether a city deserves more HRT stations can be viewed from a new perspective. It is generally believed that the more rail stations in a specific geographical area (usually within an administrative division), the more convenient for inhabitants to gain access to rail services. However, excessive rail stations and resulting narrow interstation spacing are always bound up with higher operating and maintenance costs, and lower train speed and frequency. The SEM model of this study can assist in comparing the HRT service levels of different sets of HRT stations and optimizing the layout of HRT stations. Additionally, the node-place model (Bertolini, 1996, 1999) was modified based on the outcomes from SEM (i.e., the latent variable scores). The modified node place model can provide urban planners and managers with the overall situation of the synergy between land use and HRT services.

The rest of this paper is organized as follows. In Sect. 2, related work is reviewed. Data and methods are introduced in Sect. 3. Section 4 applies the SEM model of land use characteristics and HRT service levels to a case study and presents the node-place analysis for all Dutch HRT stations. Conclusions and avenues for future research are followed in Sect. 5.

**Literature Review**

Though land use and rail-based transport constitute a two-way interaction, empirical research typically only addresses the unidirectional relationship due to the long-term nature of some processes, such as land development and rail infrastructure construction. The studies on the relationship between land use and rail-based transport are biased towards the effect of rail transport on land use, e.g., the effect on population and employment, property values, and land use/land cover change. The strength of the relationship between rail transport and land use could vary at different periods. Given the limitations imposed by the availability of consistent long-term data, the conclusions of the relationship appear to be diverse in literature. The studies with long track records (at least half a century) in many countries confirm that rail transport drove population growth, such as the Netherlands (Koopmans, 2012), UK (Alvarez et al., 2013), Finland (Kotavaara et al., 2011), Turkey (Akgüngör et al., 2011), and Portugal (da Silveira et al., 2011). However, in a short span of time (shorter than 30 years), the studies of rail’s effect on employment for US cases reach inconsistent conclusions – the link between rail access and employment exists in some cases, e.g., McMillen and Lester (2003), but not in others, e.g., Bollinger and Ihlanfeldt (2003). Some cases report that rail infrastructure initiates
land cover change (Cervero & Landis, 1997) and new residential developments (Ratner & Goetz, 2013), and that the accessibility and environmental effect of rail stations contribute to the property value (Debrezion et al., 2007). In addition to the temporal element, the variances in the significance of rail’s effect on land use could also be caused by spatial scale and methodology, which is beyond the scope of this literature review, as well as insignificant or negative effects.

The other direction of the interaction, i.e., the effect of land use on rail transport, put a spotlight on how land use and related features affect rail ridership or demand. Built environment factors, socio-economic factors, and network-dependent factors are the main determinants of rail ridership or demand. The existing studies develop around the aspects mentioned above in whole or in part (Brons et al., 2009; Chan & Miranda-Moreno, 2013; Kuby et al., 2004; Sung & Oh, 2011; Zhao et al., 2013). Density, diversity, and design of built environment (3D), which describes the number of activities, the diversity of activities, and the properties of a station area (e.g., bikeway and sidewalk facilities), respectively, are proposed by Cervero and Kockelman (1997) as principles of transit-oriented development (TOD), as the 3D factors can affect the travel demand. Moreover, socio-economic factors, such as income, age, employment, and car ownership, can increase or decrease the probability that a person will choose rail transit (Hartholt, 2016). The ridership attraction of rail is also associated with network-dependent factors, which include network accessibility and access to stations. The network accessibility is determined by the characteristics of stations and their transit service supply, including station type (terminal or transfer), the number of lines, the distance between stations, service frequency, fare levels, centrality, etc. The access to stations is affected by bus connectivity (e.g., the number of bus routes and feeder route frequencies), park-and-ride availability, street network, etc. The rail transit supply depends on the demand/ridership; however, the provision of rail transit services is also bound by technology, operational management, and cost–benefit considerations. Identifying the role of land use in rail transit from a service provision perspective, which differs from the demand/ridership perspective, would offer more practical advice for rail operations.

SEM is a multivariate statistical framework to model structural relationships between multiple variables. Since the attempt of Kitamura et al. (1992) in travel behavior research, the use of SEM has been escalating in the studies on transport and land use, and it shows considerable potential in activity-based travel demand modeling (Golob, 2003). In addition to behavioral sciences, SEM can be used for exploring the interaction between transport investment and economic growth (see Jiang et al., 2017), and the identification of the relationship between airports’ catchment area size and flight network scale by SEM (Yang et al., 2016) provides inspiration for transport facility planning issues. The node place model (Bertolini, 1996, 1999) is a flexible analytical tool, which can illustrate the competition and complementarity between land use and transport and then unveil the development dynamics (Chen & Lin, 2015; Chorus & Bertolini, 2011). The model describes transport status as node index and land use status as place index, based on which several typologies of rail stations have been developed to support decisions on setting policy and investment priorities of stations (Reusser et al., 2008; Vale, 2015; Vale et al., 2018; Zemp et al., 2011). Besides Bertolini’s original model, the node place model now
has been extended by the appearance of variant node and place indexes (Monajem & Nosratian, 2015; Papa & Bertolini, 2015), and by the introduction of a third index, such as the ‘design index’ (Vale et al., 2018).

Data and Methods

Study Area and Data

The Dutch HR network approximately totals 2800 km, most of which are double-track electrified lines (Lentink et al., 2017). According to the EU transport statistics, the Netherlands has deployed around 180 km of HR network per million people (730 km of HR network per 10,000 square km), and the network utilization rate (thousand train-km per line-km) is 50.9% (Pocketbook, 2016). Dutch HR claimed to transport 1.2 million passengers per day in 2016 (Smeets, 2017), contributing significantly to public transport. The 9292 database provides Dutch public transport information, in which the schedules and geographic information are collected from all Dutch public transport companies and compiled into General Transit Feed Specification (GTFS) files. For analyzing the HRT services, the data about HRT was extracted from the 9292 database of 2014, working with PostgreSQL in Python. In 2014, there were 399 HRT stations in operation in the Netherlands.

Statistics Netherlands (CBS) has published spatial statistical data per square. In the database CBSvierkant100, the Netherlands is covered by equally sized grid cells of 100 m × 100 m, each of which contains statistics on demography, housing, energy, social security, and proximity to facilities. The resolution of another similar database CBSvierkant500 is 500 m × 500 m. There are considerable overlaps between these two databases; however, CBSvierkant500 contains some specific attributes. CBSvierkant100, as the one with a finer resolution, was given priority to the selection of attributes. We chose the number of inhabitants, the number of households and the number of houses from CBSvierkant100, together with the environmental address density (in Dutch: Omgevingsadressendichtheid, OAD) which measures urbanity by the number of addresses per square kilometer from CBSvierkant500. Moreover, the database Wijk-en buurtkaart 2015 provides two ratio indexes at the neighborhood level, the average income per inhabitant and the passenger cars per household. The values of the ratio indexes were assigned to the grid cells of CBSvierkant100 by location in a GIS environment (Quantum GIS, version 2.18). The database Leefbaarometer (livability monitor) 2.0 provides a comprehensive indicator to show the quality of life at 100 m grid resolution. The quality of life index is a composite of 100 indicators relating to five aspects: houses, residents, services, safety, and physical environment (for methodological and technical details, refer to Leidelmeijer et al., 2014). Income and car ownership have long been held to affect public transit patronage (Kitamura, 2009). A decent living environment expressed by the high quality of life index, including built, natural, social, cultural, and economic environments, is thought to be self-sufficient at small geographical scales without special needs for long-distance commute by train. The human development dimension of land use in this
study is reflected by the quality of life index, the average income per inhabitant, and the passenger cars per household. The quality of life index was only available for the year 2014, while the other six indexes for 2015 were used.

The study area of this research is an integral and continuous region, and the spatial proximity determines the land use feature extraction (LUFE) areas of HRT stations. The Voronoi diagram, which can describe spatial relationships between points, was created from the HRT stations for defining LUFE areas. Figure 1 shows the study area, the Dutch HR network, the HRT stations, and the Voronoi diagram of HRT stations. For each LUFE area, dividing the sums of the values of land use indexes mentioned above by its area or the number of its grid cells, the population density (by area), the household density (by area), the house density (by area), the average environmental address density (by the number of grid cells), the average quality of life (by the number of grid cells), the average income (by the number of grid cells) and the average car ownership (by the number of grid cells) were calculated for SEM.

Structural Equation Modeling

An SEM model is composed of two parts: the measurement model (MM) and the structural model (SM). The measurement model examines the relationship between
latent variables and their measures (observed variables) through confirmatory factor analysis, and the structural model tests the hypotheses on the relationship between latent variables. This study holds that there is a causal link between the land use characteristics of an HRT station’s LUFE area (hereinafter called the LUFE area characteristics) and the HRT service level of the station. The SEM technique was used to analyze the structural relationship between the observed variable (land use indexes or HRT service indexes) and the latent construct (LUFE area characteristics or HRT service level). The LUFE area characteristics and the HRT service level cannot be directly observed but can be inferred from observed variables. The LUFE area characteristics, as an exogenous variable, is independent in the SEM model and measured by the land use indexes presented in Sect. 3.1. The HRT service level is an endogenous variable, which is determined by the LUFE area characteristics. The construct of HRT service level has three dimensions: the HRT service frequency, which is the departure frequency at a station; the HRT service density, which is the number of stations with a non-change connection with a station; and the HRT service extent, which is the length of non-change routes at a station. The HRT service frequency, density and extent of 399 Dutch HRT stations were obtained by querying the 9292 database. The SEM model of this study only concerns two constructs, and the path analysis of the relationship between them is straightforward. Besides SEM, some other techniques such as multiple regression analysis and Bayesian network can be used to analyze the influence of land use indexes on HRT service indexes. However, the SEM framework is more suited for this study, as land use characteristics and HRT service level accord with latent constructs in SEM. Moreover, SEM is scalable, flexible, and extensible for incorporating more constructs and variables into further research on land use and transport interaction, which is in line with the land use transport feedback cycle (Wegener, 2004).

The initial SEM model was built and analyzed with SPSS AMOS (version 22). Through Pearson correlation analysis in SPSS and tests targeting the recommended goodness-of-fit in AMOS, the variables of LUFE area characteristics, which show high collinearity within the construct and insignificant associations with the variables of the HRT service level, were excluded from the SEM model without violating the theoretical expectations. The population density, the average environmental address density, the average car ownership, and the average

\[
\begin{bmatrix}
  x_{j1} \\
  x_{j2} \\
  x_{j3} \\
  x_{j4}
\end{bmatrix} =
\begin{bmatrix}
  \lambda_1 \\
  \lambda_2 \\
  \lambda_3 \\
  \lambda_4
\end{bmatrix}
\begin{bmatrix}
  \varepsilon_1 \\
  \varepsilon_2 \\
  \varepsilon_3 \\
  \varepsilon_4
\end{bmatrix} +
\begin{bmatrix}
  \epsilon_1 \\
  \epsilon_2 \\
  \epsilon_3 \\
  \epsilon_4
\end{bmatrix}
\]

(1)

\[
\begin{bmatrix}
  y_{j1} \\
  y_{j2} \\
  y_{j3}
\end{bmatrix} =
\begin{bmatrix}
  \mu_1 \\
  \mu_2 \\
  \mu_3
\end{bmatrix}
\begin{bmatrix}
  \delta_1 \\
  \delta_2 \\
  \delta_3
\end{bmatrix} +
\begin{bmatrix}
  \delta_1 \\
  \delta_2 \\
  \delta_3
\end{bmatrix}
\]

(2)

\[Y_j = \gamma X_j + \zeta\]

(3)
quality of life were finally retained to characterize the land use of LUFE areas. Figure 2 shows the final SEM model. The measurement models of LUFE area characteristics and HRT service levels are given by Eq. (1) and (2) respectively, and Eq. (3) presents the structural model. The variables are indicated in Fig. 2. $\lambda$ and $\mu$ are the matrices of coefficients for the effects of the latent variables on the observed variables. $\epsilon$ and $\delta$ are the vectors of measurement errors. $\gamma$ is the coefficient that represents the effects of the exogenous variable on the endogenous variable. $\zeta$ is the residual of the endogenous variable. SEM lays the groundwork for further analysis in this study, and the development and application of the SEM model are illustrated in Fig. 3.

Because the observed variables are measured on different scales, the input data of the SEM model were standardized to Z-scores. However, as the multivariate kurtosis is 152.557, the standardized data are not normally distributed, to which the Chi-square test for overall fit is sensitive (Powell & Schafer, 2001). AMOS provides the Bootstrapping, which is a method for resampling from a given sample of data with replacement, to address the multivariate non-normality (Gu et al., 2016). Bootstrapping in AMOS can adjust the $p$-value associated with the Chi-square test (Bollen-Stine bootstrap method), generate parameter estimates and standard errors of parameter estimates (Bias-corrected confidence intervals and Percentile-corrected confidence intervals), and it can test the significance of individual parameters based on the sampling distribution. The Bollen-Stine bootstrap $p$-value, parameter estimates and standard errors of the SEM model were obtained with 2000 bootstrap samples.

**Node place model**

The node place model, in terms of demand and supply, is a cross-sectional analysis tool for the degree of coordination between land use driven demand and transport driven supply. The node index corresponds to the public transport service provision at

![Fig. 2 SEM model for the relationship between LUFE area characteristics and HRT service level](image-url)
rail stations, and the place index measures the intensity and diversity of human activities in station areas (Bertolini, 1999). The node place model identifies five ideal–typical situations (as illustrated in Fig. 4) in accordance with the position in the diagram to distinguish the degree of coordination between land use and transport.

Given that the combination of maximum (minimum) node value and maximum (minimum) place value is the extremity of equally developed land use and transport, similar values denote the balance between transport supply and demand if node and place indexes are measured on the same scale, which forms the ‘balanced’ areas along the middle line in the diagram. Station areas ‘under stress’, which are at the top of the middle line, indicate both land use potential and transport potential are used to the fullest. Due to the highly utilized and limited space, the conflicts between claims of land use and transport on space would appear in the further development of ‘stressed’ station areas. ‘Dependent’ station areas, which are at the bottom of the middle line,
indicate underdeveloped situations with fewer human activities and transport services. ‘Stressed’ and ‘dependent’ station areas are isolated from ‘balanced’ station areas as two extreme situations. Compared with the transport demand arising from land use, the transport services of the ‘unbalanced nodes’ above the middle line are surplus, and the transport services of the ‘unbalanced places’ below the middle line are insufficient. Driven by the idea of approaching equilibrium, the dynamic trends of unbalanced situations moving towards balanced situations (including balanced, stressed and dependent) can be identified from the static node-place diagram.

According to Bertolini’s (1999) original node place model, the node index is measured based on nine indicators which fall into four categories: the accessibility by train (including the number of directions, the daily frequency of services and the number of stations within 45 min of travel), the accessibility by bus, tram and underground (including the number of directions and the daily frequency), the accessibility by car (including the distance from the closet motorway access and the car parking capacity) and the accessibility by bicycle (including the number of free-standing bicycle paths and the bicycle parking capacity). The measurement of place index needs the number of residents, the number of workers in retail/hotel and catering, the number of workers in education/health/culture, the number of workers in administration and services, the number of workers in industry and distribution, and the degree of functional mix. However, due to data scarcity or different viewpoints on the contributing factors of land use and transport interaction, alterations to the node and place measurements have been made in varying degrees in the later applications of the model (see Monajem & Nosratian, 2015; Papa & Bertolini, 2015; Vale et al., 2018).

Though the LUFE area characteristics and the HRT service level are not measurable in the SEM, quantification of these two variables can be achieved by mathematical methods. This study borrowed the analysis framework of the node place model to extend the use of the SEM’s results, which can give an overall picture of spatial developments from an integrated land use and transport perspective. On the basis of the SEM model established in 3.2, the node index of this study combines the observed

![Node place model (Bertolini, 1999)](image)
variables of HRT service level, i.e., the departure frequency ($y_1$), the number of stations with a non-change connection ($y_2$) and the length of non-change routes ($y_3$); the place index of this study is a composite of the observed variables of LUFE area characteristics, i.e., the population density ($x_1$), the average environmental address density ($x_2$), the average car ownership ($x_3$) and the average quality of life ($x_4$). The node/place index in previous studies is conventionally assigned by the equally-weighted average of node/place-affiliated variables. However, this study used the data imputation function provided by AMOS to calculate the latent variable scores as node and place values. The regression imputation predicts the latent variable as a linear combination of the observed variables for each case when the maximum likelihood is applied for model fitting and the model parameters adopt their maximum likelihood estimates. The resultant latent variable scores were scaled to the range 0–1.

The node place model is featured by the brief and concise illustration of the ‘balanced’ and ‘unbalanced’ situations. However, it is still being discussed whether the representation of the balance between land use related demand (place) and transport service supply (node) in the node place diagram is the middle line or a concave curve with the best fit (see Reusser et al., 2008; Vale, 2015). The urban development in highly developed regions with established rail networks has been in a stabilization phase – that is, the land use situations and the HRT management and operation remain stable when there is no external intervention such as substantial investments and new technologies, and therefore most couples of node and place can be regarded as ‘balanced’ in a broader sense (including dependent, stressed, and balanced). As SEM is a data-fitting approach, the relationship defined by the SEM model, the parameters of which are estimated by the maximum likelihood method, is consistent with the idea that the majority corresponds to the state of ‘balance’ in regions at the slow pace of change. Adjusting the node and place indexes to the range 0–1, the middle line of the node place diagram used in this study has a theoretical basis for indicating the ‘balanced’ situation, while the ‘unbalanced’ situations are likely to be conservative.

**Relationship Between Land Use and HRT**

**Correspondence Between Land Use Pattern and Spatial Distribution of HRT Services**

The SEM model illustrated in Fig. 2 was calibrated by the departure frequency, the number of stations with a non-change connection, the length of non-change routes of 399 HRT stations, and by the population density, the average environmental address density, the average quality of life, the average car ownership of their corresponding LUFE areas. Figure 5 shows the final model. For the measurement models to be identified, the loadings of average population density and departure frequency are set as 1.0. The regression weights among parameters in the model all reach significance at ***$p < 0.001$ level. The model fit was evaluated by the minimum discrepancy divided by its degrees of freedom ($\chi^2$/df), the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), the
Bentler-Bonett normed fit index (NFI), the Tucker-Lewis coefficient (TLI), the comparative fit index (CFI), the root mean square of approximation (RMSEA) and the Bollen-Stine bootstrap $p$-value, the values of which are presented in Table 1. Table 1 indicates that the model fits well with the data, which substantiates the causal relationship between LUFE area characteristics and HRT service levels. From a spatial perspective, the spatial distribution of HRT services can be explained by the land use pattern. Through the connection between HRT stations and Voronoi partitions of the study area, there is a correspondence between the land use pattern and the spatial distribution of HRT services.

**HRT Station Selection**

The number and location of rail stations are important factors affecting rail use (Givoni & Rietveld, 2011). The quantified relationship given by the SEM model can further be used for the issue of whether a city deserves more transit stations. More stations will improve access to rail services. However, the additional incremental travel time by more stops thereby makes rail transit less attractive, and the role of relatively small stations is liable to be diminished by a nearby high-ranked station. An elaborate configuration of high-ranked stations and low-ranked stations

**Table 1** Model fit evaluation

<table>
<thead>
<tr>
<th>Good-of-fit index</th>
<th>Reference</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$/df</td>
<td>$&lt; 3$</td>
<td>2.996</td>
</tr>
<tr>
<td>GFI</td>
<td>0 (no fit) to 1 (perfect fit)</td>
<td>0.972</td>
</tr>
<tr>
<td>AGFI</td>
<td>$-\infty$ (no fit) to 1 (perfect fit)</td>
<td>0.940</td>
</tr>
<tr>
<td>NFI</td>
<td>0 (no fit) to 1 (perfect fit)</td>
<td>0.989</td>
</tr>
<tr>
<td>TLI</td>
<td>0 (no fit) to 1 (perfect fit)</td>
<td>0.988</td>
</tr>
<tr>
<td>CFI</td>
<td>0 (no fit) to 1 (perfect fit)</td>
<td>0.993</td>
</tr>
<tr>
<td>RMSEA</td>
<td>$&lt; 0.10$</td>
<td>0.071</td>
</tr>
<tr>
<td>$p$-value</td>
<td>$&gt; 0.05$</td>
<td>0.143</td>
</tr>
</tbody>
</table>
is therefore essential to provide better rail transit services for a localized region, which can keep down operation costs while improving train speed. When adjustments (e.g., the closure of an HRT station) are proposed to rail stations within a region, the explanatory capacity of the model in Fig. 5 for the causal relationship enables it to be used to analyze the varying HRT service levels over changes in LUFE areas. The case of Helmond demonstrates the use of the SEM model for the HRT station selection problem in a multiple transit station region (MTSR). The HRT stations of an MTSR are commonly located in the same administrative entity. Helmond is a city in the Southern part of the Netherlands, with a population of 91,539 (January 2019, source: CBS) and an area of 54.57 km². There are four stations in operation: Helmond station (the main station), Helmond Brandevoort station, Helmond’t Hout station, and Helmond Brouwhuis station. The four stations’ LUFE areas refer herein as the Helmond railway region, distinct from the administrative scope of Helmond (see Fig. 6).

For quantification and comparison of the HRT service levels provided by stations, rail network accessibility (RNA) is proposed as a comprehensive index to synthesize the departure frequency ($y_1$), the number of stations with a non-change connection ($y_2$) and the length of non-change routes ($y_3$), which can be given as follows:

$$Ac_i = y_{i1}^{\alpha_1} \times y_{i2}^{\alpha_2} \times y_{i3}^{\alpha_3}$$  \hspace{1cm} (4)

where $Ac_i$ is the RNA of station $i$, $\alpha_1$, $\alpha_2$, $\alpha_3$ are parameters. The RNA of a station can generally be indicated by the ridership. The available ridership of 249 HRT stations was used for the intercept-free linear regression, where $\alpha_1=1.151$, $\alpha_2=0.274$, $\alpha_3=0.162$, and $R^2=0.675$. The HRT service level of an MTSR depends on all the affiliated stations. The regional rail network accessibility (RRNA), as given in Eq. (5), is proposed to evaluate the HRT service level at the MTSR level.

$$\text{Access} = \sum_{i=1}^{n} w_i \times Ac_i$$  \hspace{1cm} (5)

where $n$ is the number of HRT stations within an MTSR, $w_i$ is the weight of station $i$. The weight of station $i$, as shown in Eq. (6), takes into account both the population weight ($w_p^i$) and the area weight ($w_A^i$). The population in the proximity of an HRT station directly explains the importance of the station against other stations within the same MTSR. In addition, an HRT station with a relatively excessive LUFE area is not conducive to offering HRT service access to distant inhabitants. The negative impact of an oversize LUFE area on the RRNA is reflected in the greater area weight attaching to the corresponding HRT station.

$$w_i = \frac{w_p^i \times w_A^i}{\sum_{j=1}^{n} w_p^j \times w_A^j}$$  \hspace{1cm} (6)

As a small city with densely accumulated HRT stations (four stations in 54.57 km²), Helmond may be questioned whether maintaining such many HRT stations is necessary. In order to improve the effectiveness and efficiency of HRT operations, multiple scenarios are designed for Helmond to seek the optimal configuration of
HRT stations in the light of RRNA. The RRNA of the base scenario (four stations) is the benchmark for three-station scenarios (scenarios 1, 2 and 3 with one station being closed, see Table 2), two-station scenarios (scenarios 4, 5 and 6 with two stations being closed, see Table 2) and one-station scenario (scenario 7 with the main Helmond station, see Table 2). In the case of Helmond, the area of the MTSSRs of different scenarios fluctuates within a narrow range (the smallest MTSSR of scenario 7 shrank by 18.8%, compared to the largest one of the base scenario). A significant alteration to the structure of rail networks will be manifest in the substantial change of MTSSRs, which is beyond the scope of this study. The SEM model for HRT selection discussion here only applies to areas with a high station density, where opening or closing a station will not bring much change to the MTSSR.

The comparisons of the eight scenarios are shown in Table 2, which indicates the combination of Helmond, Helmond’t Hout and Helmond Brandevoort (scenario 3) provides the highest RRNA. In scenario 3, the closure of Helmond Brouwhuis affects the RNA of the other stations and the RRNA of the Helmond railway region, which is along with most of the LUFE area of Helmond Brouwhuis being absorbed into the LUFE area of Helmond station. From the obvious lower density of population and environmental address (OAD) of Helmond Brouwhuis’ LUFE area (see Table 2), it can be inferred that the proximity of a low-ranked station serving a comparatively sparsely populated area to a high-ranked station weakens the role of the low-ranked station. Closing such a low-ranked station would be an option to improve the RRNA with the transfer of the station’s functions to the nearby high-ranked station.

**Overall Picture of the Synergy Between Land Use and HRT**

As mentioned in 3.3, the modified node place model provides a framework to discuss the integration of land use and HRT services based on the derivatives of the SEM model. The values of the node index for 399 Dutch HRT stations and the values of the place index for the corresponding LUFE areas are mapped in the node place diagram (Fig. 7a). The averages of the node and place indexes are added to the diagram as reference lines. In the diagram, the ‘dependent’ locations mainly accumulate in the area where node and place indexes are lower than average. Further developments in ‘dependent’ station areas need additional incentives either from land use or HRT upgrades. The ‘stressed’ station areas, such as Amsterdam Centraal (Central), Utrecht Centraal, Amsterdam Sloterdijk and Rotterdam Centraal (see Fig. 7a), represent the saturation levels of development regarding both land use and HRT. The typical ‘unbalanced place’ station areas in Fig. 7a, including Den Haag HS, Amsterdam Amstel, Diemen Zuid, Utrecht Zuilen and Utrecht Maliebaan, should improve the HRT service level to fill in the gap between the strong land use driven demand indicated by the higher place indexes and the inadequate HRT service provision. All the ‘stressed’ and ‘unbalanced place’ station areas mentioned above are located in the Randstad, which is a metropolitan conurbation centering on Amsterdam, Rotterdam, The Hague, Utrecht (the four largest Dutch cities) and their surrounding areas. The Randstad is the most densely populated region with the most vibrant economy in the Netherlands, and...
<table>
<thead>
<tr>
<th>No</th>
<th>Scenario</th>
<th>Region area (km²)</th>
<th>RRNA*</th>
<th>No</th>
<th>Scenario</th>
<th>Region area (km²)</th>
<th>RRNA*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>327.75</td>
<td>1079.34</td>
<td>2</td>
<td></td>
<td>339.5</td>
<td>1080.88</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>308.5</td>
<td>1141.86</td>
<td>4</td>
<td></td>
<td>308</td>
<td>1045.69</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>296.25</td>
<td>1117.20</td>
<td>6</td>
<td></td>
<td>307.5</td>
<td>1127.36</td>
</tr>
</tbody>
</table>

Table 2: Comparisons of configurations of HRT stations in Helmond

Insights into the Correspondence Between Land Use Pattern...
<table>
<thead>
<tr>
<th>No</th>
<th>Scenario</th>
<th>Region area (km²)</th>
<th>RRNA*</th>
<th>No</th>
<th>Scenario</th>
<th>Region area (km²)</th>
<th>RRNA*</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td></td>
<td>276</td>
<td>1071.36</td>
<td>base</td>
<td></td>
<td>340</td>
<td>1092.93</td>
</tr>
</tbody>
</table>

Stations (Scenario_base) | Brandewoort (a) | Helmond’t Hout (b) | Helmond (c) | Brouwhuis (d) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average population density</td>
<td>595.7195572</td>
<td>579.6470588</td>
<td>666.6930693</td>
<td>365.3140097</td>
</tr>
<tr>
<td>Average OAD density</td>
<td>1001.549815</td>
<td>1092.329412</td>
<td>1299.247525</td>
<td>680.173913</td>
</tr>
</tbody>
</table>

*RRNA is regional rail network accessibility
the components of the Randstad region are therefore expected to have higher values of place index, which conforms to the findings of the node place analysis. HRT operations management of ‘unbalanced place’ station areas in Randstad can draw lessons from ‘stressed’ station areas to promote more balanced development. The ‘unbalanced node’ station areas, such as Schiphol and Zwolle have enormous potential for further development, which benefits from the sufficient HRT service provision.

Figure 7b continues the discussion on the HRT stations in Helmond within the framework of node place analysis. Besides the original four HRT stations in Helmond, which are extracted from Fig. 7a, the situations of scenario 3 (see 4.2) are delineated in Fig. 7b. Figure 7b indicates the adjustment of HRT stations in the Helmond railway region would not upset the ‘dependent’ balance between ‘node’ and ‘place’. The node place model provides a macro perspective on the integration of land use and HRT, which is based on the SEM model taking a deeper dive into the relationship between them. The two models used in this study are not only methodologically linked, but their functions complement each other. The node place diagram displays the way towards ‘balance’ for ‘unbalanced’ station areas. The SEM model can further provide advice on how to improve ‘balanced’ station areas (including stressed, balanced and dependent) in the light of RRNA, as exemplified by the Helmond case.
Conclusions and Discussion

This study identified and validated the causal relationship between land use and HRT service by SEM, given that there is a correspondence between land use patterns and spatial distribution of rail transit services. The spatial perspective applied...
in this study enhances the understanding of the nexus between land use and rail transport. The study area, i.e., the Netherlands, is partitioned by the Voronoi diagram of all Dutch HRT stations. The land use is characterized by the population density, the average environmental address density, the average car ownership, and the average quality of life. The HRT service level provided by a station is evaluated by rail service frequency, density, and extent. The SEM model can provide a reference to the supply of HRT services based on the LUFE area characteristics. The case of Helmond shows the potential of the SEM model for the discussion on HRT station selection in an MTSR. Moreover, in this study, the node place model is adapted with the latent constructs of the SEM model (i.e., LUFE area characteristics and HRT service levels). The modified node place model was used to analyze and differentiate the integration of land use and HRT services at the regional level. The full picture of the integration of land use and HRT given by the node place model can facilitate the design and implementation of regional development strategies to promote more balanced land use and HRT services. The SEM model and the node place model are complementary to be used as analytical and decision-making tools for rail transit oriented regional development. Regions, especially growing megacities, which expect transit-oriented development, can draw lessons from the correspondence between land use pattern and spatial distribution of HRT services in the Netherlands. More case studies with various geographical scopes and rail network densities can create a case study library for case-based reasoning in a new planning area, which would proceed the findings of this study in a more generalized manner.

The place dimension of rail stations have uncertain boundaries, and how the rail stations are linked with their surrounding environment is determined by the study purpose. Conventional catchment areas exclude the non-immediate surroundings of rail stations from the analysis of land use and rail transport. The catchment areas of rail stations are usually defined as a walkable distance ring of stations for transit ridership prediction, transit impacts evaluation, and policy formulation in the planning of transit-oriented development (Guerra et al., 2012). In this study, the extraction of land use features depends on the proximity to HRT stations, and the Voronoi diagram of HRT stations covers the entire study area, through which the spatial structure of land use is linked with the HRT networks. The relationship between land use and HRT services can be derived from the Voronoi partitions by SEM. The commonalities and differences in the LUFE areas of this study and catchment areas need to be further explored. The size of LUFE areas and the quantity indicators on land use characteristics of them, such as the number of inhabitants, are highly relevant to the station density. The relationship between land use and rail transit can be better revealed in the density of land use indexes (see 3.1) than in quantity indicators, especially in highly populated and economically active areas, where a higher station density is accompanied by smaller LUFE areas. However, without quantity indicators, the SEM model of this study is not applicable to discussions on HRT selection with structural changes to rail networks (for example, closing a station in areas with a low station density), which can cause significant changes to the MTSR. Clarifying the role of the quantity indicators on land use characteristics in the determination of rail service levels is needed to broaden and deepen the conclusions of this study.
The nature of the causal relationship between land use and HRT services in this study is to match transport driven supply with land use driven demand. Under the framework of HRT supply and demand, the HRT services in this study are only related to the attributes of HRT systems (i.e., the departure frequency, the number of stations with a non-change connection and the length of non-change routes), without the consideration of access to HRT services. However, besides network accessibility, people’s willingness to choose rail transit is influenced by access to rail transit services, including pedestrian access, bicycle access, automobile access, and bus transit access (Brons et al., 2009). When the scope of rail transport discussion is extended from rail systems themselves to a broader discussion on rail transit use, other transport modes connecting rail transit stations, which are absent in this study, should be considered.

References


**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Authors and Affiliations

**Zhongqi Wang¹· Qi Han¹· Bauke de Vries¹· Li Dai¹**

Qi Han
q.han@tue.nl

Bauke de Vries
b.d.vries@tue.nl

Li Dai
l.dai@tue.nl

¹ Information Systems in the Built Environment group, Department of Built Environment, Eindhoven, University of Technology the Netherlands, 513, 5600 MB, Eindhoven, Netherlands