Understanding public transport and built environment integration at the neighborhood scale: Towards a method for holistic quantitative assessment

Christian M. Marti, IVT, ETH Zurich
Ulrich Weidmann, IVT, ETH Zurich

Conference paper STRC 2016
Understanding public transport and built environment integration at the neighborhood scale: Towards a method for holistic quantitative assessment

Christian M. Marti
ETH Zurich
Inst. for Transport Planning and Systems IVT
HIL F 13.2
Stefano-Franscini-Platz 5
8093 Zurich
Phone: +41 44 633 68 16
Fax: +41 44 633 10 57
demail: christian.marti@ivt.baug.ethz.ch

Prof. Dr. Ulrich Weidmann
ETH Zurich
Inst. for Transport Planning and Systems IVT
HIL F 13.1
Stefano-Franscini-Platz 5
8093 Zurich
Phone: +41 44 633 33 50
Fax: +41 44 633 10 57
demail: weidmann@ivt.baug.ethz.ch

May 2016

Abstract

This paper presents an approach for the holistic quantitative assessment of public transport and built environment integration at the neighborhood scale. This integration is important to achieve a sustainable urban development and is a widely accepted policy principle, but current methods for its assessment are insufficient and lack a clear theoretical base. Therefore, a novel approach is developed based on the premise that public transport and built environment integration is achieved if their elements are attuned to each other as much as possible, and that any assessment of integration should therefore measure how well this is achieved based on an investigation of reciprocal effects. For this, a qualitative system model of interactions between elements of the built environment and of public transport systems is built and then used to identify adequate measuring points for integration – which are the base for the development of quantitative spatial indicators. The feasibility of this approach is demonstrated on the example of two indicators that are related to pedestrian access and egress to and from public transport: influences of density on the number of potential public transport users and influences of the pedestrian network on the size of the public transport catchment area.

Keywords

1. Introduction

This paper presents an approach for the holistic quantitative assessment of public transport (PT) and built environment (BE) integration at the neighborhood scale. It is part of the larger topic of land use transport integration, which is widely seen as one of the most important means to achieve a sustainable urban future (Te Brommelstroet and Bertolini, 2010; Yigitcanlar and Kamruzzaman, 2014; Soria-Lara, Valenzuela-Montes, and Pinho, 2015). Its importance stems from the fact that mobility plays a pivotal role in urban sustainability (e.g., it affects emissions, energy consumption, opportunities in daily life, economic prosperity, and quality of life (Jenks, Williams, and Burton, 2000; Goldman and Gorham, 2006; Cheng, Bertolini, and le Clercq, 2007; German EU Presidency, 2007; Hull, 2011; Bertolini, 2012; Puppim de Oliveira et al., 2012; Pearson, Newton, and Roberts, 2014)), and that mobility patterns and travel behavior are strongly intertwined with the way cities are built – the form and structure of the built environment. Land use patterns and urban structures influence travel behavior and thus transport flows, mode choice and travel times, but these in turn define accessibilities and therefore determine location choices and – again – land use patterns (Hurd, 1903; Hoyt, 1939; Mayer, 1969; Mackett, 1985; Handy, 2005; Chang, 2006; Næss, 2006; Cao, Mokhtarian, and Handy, 2009; Ewing and Cervero, 2010; Papa and Bertolini, 2015).

PT is frequently presented as a key factor for achieving integrated transport and land use plans, often together with compact and mixed-use urban development (Peterson and Schäfer, 2004; Devereux, van der Bijl, and Radbone, 2005; Kenworthy, 2006; Hickman et al., 2009; Curtis and Scheurer, 2010; Puppim de Oliveira et al., 2012; Suzuki, Cervero, and Iuchi, 2013; Oswald Beiler and Treat, 2015). One main reason is that PT bundles movements and is therefore much more space efficient and creates less emissions than car travel, which is particularly valuable where densities are high and space is scarce. On the other hand, “mass transit needs mass” (Suzuki et al., 2013, p. 15), i.e., is not viable without a certain conglomeration of users (density). Furthermore, mixed uses tend to generate a more evenly distributed demand which allows for greater efficiency of PT (Suzuki et al., 2013; Orth, Frei, and Weidmann, 2015). Therefore, PT only works efficiently if the BE is aligned to its needs, but can also contribute to a better quality of the BE if such coordination is achieved.

For these reasons, PT and BE integration is incorporated in many planning policies worldwide (Burchell, Listokin, and Galley, 2000; Curtis and Punter, 2004; Waddell, 2011; UVEK, 2012; ARE, 2015b) and corresponding concepts and planning approaches have been developed. Important examples include eco-city (Roseland, 1997; Kenworthy, 2006), new urbanism (Katz, 1994; Leccese, McCormick, and Congress for the New Urbanism, 2000), sustainable accessibility (Bertolini, le Clercq, and Kapoen, 2005; Cheng et al., 2007; Curtis, 2008), pôle d’échanges (Menerault, 2006), smart growth (Burchell et al., 2000; Downs, 2005; Handy, 2005;
Edwards and Haines, 2007), and transit-oriented development (TOD) (Calthorpe, 1993; Cervero et al., 2004; Dunphy et al., 2004; Evans and Pratt, 2007; Curtis, Renne, and Bertolini, 2009; Cervero and Sullivan, 2011). However, their implementation is not always successful and planning reality deviates strongly from what would be adequate given the theoretical knowledge developed. While institutional barriers and unsuitable planning practices play an important role (Curtis, 2008; Te Brommelstroet and Bertolini, 2010; Switzer, Bertolini, and Grin, 2013), another reason is the lack of objective assessment approaches for PT and BE integration (Renne and Wells, 2005; Evans and Pratt, 2007; Dur, Yigitcanlar, and Bunker, 2014; Federer, 2014; Hale, 2014).

In recent years, this problem has been successfully addressed on a network level with the development of quantitative accessibility-based approaches for PT and BE integration (Cheng et al., 2007; Durousset et al., 2009; Curtis and Scheurer, 2010; Keller et al., 2011; Silva, Reis, and Pinho, 2014; Singh et al., 2014; Papa and Bertolini, 2015; Singh et al., 2015; Vale, 2015). However, such methods are not available for the smaller neighborhood scale. This is an important shortcoming because there are crucial interactions between PT and BE also at this scale. For example, detailed density distribution relative to PT stop location influences PT patronage, location and mix of uses influence PT demand distribution, and road space organization such as pedestrian crossings, street layout, and segregation type affect PT performance (Currie, Ahern, and Delbosc, 2011; Currie and Delbosc, 2011; Carrasco, Fink, and Weidmann, 2012). Because PT supply is spatially discrete, access and egress legs are prerequisites for any ridership at all; walking and (to a lesser extent) cycling are the main modes for access and egress, and their attractiveness and competitiveness depends on the structure, quality and safety of the urban environment, on local activity range, as well as on the provision of designated infrastructure (Filion, McSpurren, and Appleby, 2006; Thorne, Filmer-Sankey, and Alexander, 2009; Carmona et al., 2010; Grob and Michel, 2011; Adkins et al., 2012). PT operation, layout, and design in turn affect local quality aspects of the built environment such as accessibility, legibility, permeability, noise, and safety (Burns, 2005; Devereux et al., 2005; COST TU1103, 2015; Marti et al., forthcoming).

Existing attempts to define what PT and BE integration at the neighborhood scale is and how it can be assessed quantitatively are either focused on individual aspects such as walking access (Schlossberg and Brown, 2004) or sustainable mode share (Hale, 2014), are only partially quantitative and operationalized (Evans and Pratt, 2007; Renne, 2009; ITDP, 2014), or cover a broad range of sustainability criteria beyond integration (Dur et al., 2014). All of them consider characteristics and requirements of PT with far less detail than those of the BE, and indicators are selected based on expert opinion and without a theoretical framework. The latter also means that it remains unclear what exactly integration is and whether it is really measured by these approaches. This could only be achieved by a systematic analysis of the interactions between
PT and BE at the neighborhood scale as the base for assessment criteria development or selection.

In summary, PT and BE integration is highly relevant for achieving a sustainable urban development and is therefore an accepted policy principle. However, there is currently no clear understanding of what exactly such an integration means at the neighborhood scale. Accordingly, there is also no objective and holistic quantitative assessment methodology. This paper addresses this research gap. It presents a new quantitative assessment approach for PT and BE integration that is based on an analysis of interactions. Section 2 presents the research approach, with a focus on the method for developing quantitative indicators for integration. Section 3 demonstrates the feasibility this approach with the aid of two concrete indicators. And section 4 ends the paper with conclusions and an outlook on future research steps.

2. Research approach

2.1 Overview

The main premise is that PT and BE integration is achieved if their elements are attuned to each other as much as possible under the given conditions, i.e., if mutual requirements are met and reciprocal effects are balanced. Therefore, any assessment of integration should consider how well this is achieved, based on an investigation of reciprocal effects. To achieve this, a qualitative system model of interactions between elements of the built environment and of public transport systems has been developed. It is used to identify adequate measuring points for integration, which are the base for developing quantitative spatial indicators for integration.

2.2 Qualitative system model

The system model is based on prior projects at the chair for transport systems at IVT, complemented with a literature review. It consists of 65 elements and 178 influences and is structured in eight thematic sectors and four variable type layers as depicted in Figure 1. Influences are only considered within the same layer and across layers in the direction from input to results (shown as black arrows in Figure 1). While secondary effects (i.e., in the other direction) exist in the long term, they are not present when analyzing one specific situation and therefore are beyond the scope of this paper.
Results are further structured into intermediate and final results. There are six intermediate and two final results for PT and BE each, as depicted in Figure 2 with the influences between them. Note that also other elements (variables, points of influence, and inputs) directly influence intermediate and final results, but are not depicted in Figure 2 to improve readability.
2.3 Identification of measuring points

Measuring points are identified in the system model by starting from intermediate results and analyzing effect chains. For each intermediate result, a model excerpt with complete effect chains is created. Additionally, direct effects on final results from other elements than intermediate results are also analyzed.

All influences that cross the “hemisphere” boundary (black and white semicircles in Figure 1) and that are connected to elements that can be influenced (i.e., are not only linked to input
elements) are defined as measuring points. Altogether, 18 such measuring points were identified in the model:

1. Influence of density on the number of potential PT users
2. Influence of the pedestrian network on the size of PT catchment area
3. Influence of the bicycle network on the size of PT catchment area
4. Influence of road speed limit on PT operating speed (and thus attractiveness and productivity)
5. Influence of mixed traffic, street and intersection layout and operation, and priority measures on PT operating speed
6. Influence of type and mix of uses on variation of PT demand and thus dwell time and operating speed
7. Influence of mixed traffic, street and intersection layout and operation, and priority measures on PT operational quality (reliability)
8. Influence of type and mix of uses on variation of PT demand and thus dwell time and operational quality (reliability)
9. Influence of type and mix of uses on variation of PT demand and thus productivity
10. Influence of road speed limit and mixed traffic type on PT safety
11. Influence of BE on quality of pedestrian PT access and egress
12. Influence of BE on quality of cycling PT access and egress
13. Influence of PT system type and vehicles, alignment, and stop design on BE legibility
14. Influence of PT system type and vehicles and alignment on BE conformity of scales
15. Influence of PT system type and vehicles, alignment, and stop design on BE conformity of design
16. Influence of PT system type and vehicles and alignment on distribution of public and road space
17. Influence of PT alignment, frequency and operating speed on pedestrian permeability and thus walkability
18. Cumulative influence of PT attractiveness on modal shift

As an example, Figure 3 depicts three specific measuring points related to pedestrian access and egress to and from PT stops together with the elements and influences used to derive them. The figure is a combination of the effect chain analysis of intermediate results PT patronage and PT access quality (walking); it excludes aspects related to bicycle access that are also part of the effect chain of PT patronage.

Patronage is affected by where potential PT users are or want to go relative to PT stops, and how they can access and egress these stops using the pedestrian network. The former essentially deals with density and its distribution, while the latter concerns how well the network allows walking access and egress. The quality of walking access to PT is a more general element of
PT attractiveness and affected by the neighborhood’s walkability, an element that itself is influenced by a variety of BE elements.

Figure 3  Example: Identification of measuring points
2.4 Indicator development

For each of the 18 measuring points, a quantitative indicator is developed. Indicators therefore reflect the magnitude of influences between the system “hemispheres”. This also means that they do not predict the outcome of any variables in the model – this is not their purpose.

The respective qualitative model excerpt is analyzed for elements that could be used as a variable in the indicator. This also includes elements that are not directly considered for measuring point identification. In parallel, a literature review on existing indicators and potentially useful measuring approaches is conducted. Based on this, the rationale of the indicator is summarized and performance measure and normalization approach are defined.

Indicators are of increasing form, i.e., higher values represent a better integration, and are normalized to values between zero and one. Indicators are formulated generically and parameters such as thresholds or functions necessary for specific applications are listed. While some of them are universal, e.g. can be defined based on an analysis of physical, operational, organizational and logical values or limits, others need to be estimated from literature or data analysis. Indicator development for measuring points 1 and 2 depicted in Figure 3 is explained in detail in section 3.

3. Indicators

3.1 Indicator 1: Influences of density on the number of potential PT users

3.1.1 Theoretical rationale and background

Influences of density and its distribution

Activity density strongly influences the number of potential users of a PT stop: the number and location of activities around a stop basically determines the number of potential PT users that may access this stop or egress from it (Gutiérrez, Cardozo, and García-Palomares, 2011). Therefore, higher density is beneficial for PT because it increases patronage and thus economies of scale and productivity. However, there are two distinct influencing aspects that should be considered:

- absolute level of activity density around a PT stop and
- density distribution relative to PT stop location.
The reason for the second aspect to be important is that the distance PT users are willing to travel to access a stop is limited. Most PT users access PT systems on, and the share of people willing to walk to a stop decreases with the access distance (Zhao et al., 2003). In fact, “spatial accessibility to the transit feature is the primary determinant of transit use and only in the presence of such accessibility will a user consider other factors such as cost, comfort, security, or other factors” (Biba, Curtin, and Manca, 2010, p. 350). Similar relationships can be assumed for egress.

Therefore, an assessment of the impact of density on the number of potential PT users should consider both the absolute value of density and its distribution relative to PT stops simultaneously.

**Measuring density**

The most commonly available data related to activity density are residential and job density; the latter often includes attributes on the types of businesses that are present, which can be used to correct for uses that are generating many trips (e.g., retail). This information could also be extracted form data on points of interests.

Data sources exist both in aggregated (e.g., number of residents per hectare grid cell) and disaggregated (e.g., residents per building or parcel) form. For the analysis of the neighborhood scale, aggregated data is in general not appropriate, since a distance difference of e.g. 100 meters is already highly relevant. Therefore, disaggregated data is necessary in order to consider fine-grained structures within the built environment. If such data is not available, it can also be estimated from aggregated sources (see, e.g., Greger (2015)).

Both aggregated and disaggregated data provides point measures for the number of activities. In order to analyze density distribution, a continuous representation of density is needed. A common approach to generate a smooth density surface from point data is kernel density estimation, which adequately represents variation in population or services (Gatrell et al., 1996; Porta et al., 2009; Lewis, 2015; Neutens, 2015). The basic idea is to estimate density at point $s$ based on data points $s_i$ (each point representing, e.g., an activity or a person) by weighting them with distance $d_i$ between $s$ and $s_i$ so that data points further away from $s$ contribute less to the estimated density at $s$. This analysis is conducted using all data points $s_i$ with $d_i$ smaller than a so-called bandwidth $h$. The distance weight is applied with a kernel function $k(\cdot)$, which must integrate to one so that the overall count of density values estimated with the kernel within an analysis area remains the same as when simply summing up the number of points within that area.

The exact functional form of $k(\cdot)$ does not have a large effect on the estimation accuracy (Gatrell et al., 1996; O'Sullivan and Wong, 2007; Danese, Lazzari, and Murgante, 2008; Lewis,
A typical kernel for geographic density estimation (Gatrell et al., 1996; Neutens, 2015; O'Sullivan and Wong, 2007) is the quartic function based on Silverman (1986); it is the only available kernel in ArcGIS Spatial Analyst (Quantitative Decisions, unknown; ESRI, 2016) and one of the available options in QGIS (QGIS Development Team, 2016), GRASS GIS (Menegon and Blazek, 2015), and SAGA (Conrad, 2010). The kernel density estimation using a quartic kernel can be formulated as follows (Gatrell et al., 1996):

\[
\hat{p}(s) = \sum_{d_i \leq r} \frac{3}{\pi h^2} \left(1 - \left(\frac{d_i}{h}\right)^2\right)^2
\]

where \(\hat{p}(s)\) is the estimated density at point \(s\) (in absolute terms, i.e., the sum of all \(\hat{p}\) equals the total number of points) and \(n\) is the number of data points \(s_i\) that fulfil \(d_i \leq h\). Thus, the summation is over all pairs \(s\) and \(s_i\) where \(d_i\) does not exceed \(h\). “The region of influence within which observed events contribute to \(\hat{p}(s)\) is therefore a circle of radius \(h\) centred on \(s\). At the point \(s\) \((d_i = 0)\), the weight is simply \(\frac{3}{\pi h^2}\) and drops smoothly to a value of zero at distance \(h\)” (Gatrell et al., 1996, p. 260). In this approach, data with counts (e.g. of activities or people per point) can simply be treated as multiple data points at a single location (O'Sullivan and Wong, 2007).

The selection of bandwidth has a much larger effect on the estimation results than the kernel (Gatrell et al., 1996; O'Sullivan and Wong, 2007). The bandwidth defines how much the data are smoothed by the kernel density estimation. “Large […] bandwidths will result in more smoothed patterns, whereas small bandwidths will emphasize local peaks and troughs” (Neutens, 2015, p. 17). Therefore, the danger in choosing a bandwidth too large is that local patterns are lost, while in the opposite case, local variations may be overemphasized (O'Sullivan and Wong, 2007). Approaches for both fixed and variable bandwidth selection have been presented (e.g., Brunsdon (1995)), but in spatial analysis, fixed values are by far the most common – one of the main reasons being that available GIS platforms until today only include that option (Conrad, 2010; Menegon and Blazek, 2015; ESRI, 2016; QGIS Development Team, 2016). Furthermore, most methods for fixed bandwidth selection (e.g., Jones, Marron, and Sheather (1996)) have only been applied in non-spatial statistics.

Most applications of kernel density estimation for spatial analysis rely on experiments with different bandwidths to choose an appropriate value (O'Sullivan and Wong, 2007; Spencer and Angeles, 2007; Kloog, Haim, and Portnov, 2009; Porta et al., 2009), either by visual inspection of the results or by some form of sensitivity analysis. The choice of bandwidth range is often based on a theoretical discussion of the extent of influence of one data point, or it is related to some physical observation such as block size (Kloog et al., 2009). One method repeatedly mentioned is to use distance between data points as a reference. O'Sullivan and Wong (2007)
suggest “that bandwidth should be of the same order as typical distances between areal unit centroids” (p. 165) and Danese et al. (2008) use mean nearest-neighbor distance between centroids (in both cases, aggregated data was used, therefore “centroid” refers to census blocks or similar areas).

Kernel density estimation theoretically allows to compute density at every point within an analysis area; however, it is generally operationalized with raster representations of continuous surfaces (Neutens, 2015). Therefore, a raster grid resolution, or cell size, must be defined. Large grid cells would jeopardize the goal of the kernel density estimation by not capturing local detail, whereas the main concern with small cells is computational cost (Spencer and Angeles, 2007). Furthermore, the location of the grid can affect estimates per cell, but this becomes negligible if grid resolution is substantially smaller than the bandwidth (O'Sullivan and Wong, 2007). As a rule of thumb, O'Sullivan and Wong (2007) suggest to use a maximum cell size of 0.1 x bandwidth.

Another important consideration for kernel density estimation are edge effects. These tend to distort estimates close to the boundary of an analysis area because there are no data points beyond the border influencing these locations. This problem can either be addressed by adjusting the estimation algorithm, or by constructing a “guard area” with a width at least equal to the bandwidth around the analysis area, which contains data points which influence the density estimation; for this guard area itself, no density estimation is conducted (Gatrell et al., 1996).

**Evaluation of absolute density values**

Considering absolute value of density, previous indicators for the orientation of the built environment towards public transport (e.g., the spatial index for neighborhood land use transport integration by Dur et al. (2014) or the TOD index by Singh et al. (2014) and Singh et al. (2015)) have simply used a “the more the better” approach. While this reflects the fact that “denser is better” for PT, it invariably renders the best results in core areas of cities with generally high densities and much lower results for areas with moderate density, e.g. most residential zones. Therefore, when assessing neighborhood scale PT and BE integration, it seems more promising to consider how high density is relative to what is actually achievable in the given situation, i.e., what density could be achieved in the analysis area.

This requires to define achievable benchmark density values for certain situations, $p_{min}$. If the density within analysis area $A$, $p_A(A)$, is below the benchmark, it should be evaluated as insufficient, whereas exceeding the benchmark should not be penalized since it is beneficial for PT. This can be formulated in the form of a density benchmark achievement value in the range $[0, 1]$: 

\[ \text{achievement} = \min\left(1, \frac{p_A(A)}{p_{min}}\right) \]
\[ b(p_A(A), p_{\text{min}}(A)) = \begin{cases} 
1, & p_A(A) \geq p_{\text{min}}(A) \\
\frac{p_A(A)}{p_{\text{min}}(A)}, & \text{otherwise} 
\end{cases} \]

Note that \( p_A(A) \) is simply the summation of all activities within analysis area \( A \) divided by that area. \( p_{\text{min}}(A) \) depends on many characteristics of \( A \) such as the uses or zoning within \( A \) or its location (e.g., distance to the CBD or closest subcenter if it is part of an agglomeration). A viable approach is to define categories for analysis areas and use a certain percentile of densities in similar situations as the benchmark value of what is achievable. In such comparisons, it is important to only consider built up areas for \( A \), because including e.g. lakes, forests, parks, or agricultural land would yield lower density values.

**Evaluation of density distribution**

The walk access or egress distance to or from a PT stop is one of the main factors that determine the probability that potential PT users actually use PT. Because most PT users access and egress on foot, walk distance thus affects patronage substantially and in some cases more than factors such as cost, comfort, or security (Murray et al., 1998; Beimborn, Greenwald, and Jin, 2003; Zhao et al., 2003; Gutierrez and Garcia-Palomares, 2008; Biba et al., 2010; Garcia-Palomares, Gutiérrez, and Cardozo, 2013). In planning practice, often a fixed threshold of “willingness to walk” is used, such as the common 400m or \( \frac{1}{4} \) mile radius (Biba et al., 2010). However, this does not adequately represent the continuous decay of willingness to walk to or from PT with increasing access or egress distance. A common approach to overcome this limitation is the use of a distance-decay function \( z(d_{s,t}) \), where \( d_{s,t} \) is the pedestrian network distance between point \( s \) and the closest PT stop \( t \) (Walther, 1973; Zhao et al., 2003; Gutiérrez et al., 2011). While this function is usually used for the development of PT ridership prediction models, it essentially represents the relationship between likelihood of PT use and access or egress distance. Normally, \( z(d_{s,t}) \) is estimated based on an analysis of observed PT access and egress leg lengths for distance intervals, normalized by the population or number of activities located at each interval distance and multiplied with a constant so that \( z(0) = 1 \). Different functional forms can be used for distance-decay functions (Osth, Lyhagen, and Reggiani, 2016), but in direct PT demand models, the exponential form often provides a good fit for observed data (Zhao et al., 2003; Gutiérrez et al., 2011). Furthermore, often a threshold distance is defined where only very few users are willing to walk and the function is cut off there.

Since the decay function represents the share of potential users actually accessing (or degressing from) PT depending on access (or egress) distance, it can be interpreted as an ideal distribution of potential users in space around a PT stop. Therefore, it can be used as a reference to assess how well density distribution is oriented towards PT. For this, a reference density function used, which defines the “ideal” density at every spatial point \( i \) based on its distance to the closest PT.
stop. For use together with the raster grid representing density distribution, it is computed for raster cell $c$ as follows:

$$p_{\text{ref}}(c) = m \times z(d_{c,t})$$

where $m$ is a scaling factor so that $\sum_{i=1}^{n} p_{\text{ref}}(c_i) = \sum_{i=1}^{n} \hat{p}(c_i)$, i.e., the overall number of activities in the analysis area is the same for the density distribution computed using kernel density estimation and the reference density. Therefore,

$$m = \frac{\sum_{i=1}^{n} \hat{p}(c_i)}{\sum_{i=1}^{n} z(d_{c_i,t})}$$

where $n$ is the number of raster cells in the analysis area.

### 3.1.2 Performance measure

There are two scale levels considered by indicator 1: the disaggregated fine raster used for density distribution analysis, and the aggregated density over the entire analysis area (e.g., a neighborhood development). The indicator is a combination of two performance measures, one for each of these scales.

First, density distribution is evaluated for each raster cell $c$ using the deviation of the density estimated with the kernel from the reference density value:

$$q(c) = \hat{p}(c) - p_{\text{ref}}(c)$$

where $q(c)$ is the density distribution performance measure for cell $c$. Positive values of $q(c)$ indicate that the density in cell $c$ is higher than the reference density, while negative values indicate the opposite. This can be presented in a map which can be used to identify locations with deviations from reference density.

Second, the absolute values of this cell-based performance measure are aggregated over the analysis area $A$ and multiplied with the density benchmark achievement value $b(p_A(A), p_{\text{min}}(A))$:

$$I_{1,A} = b(p_A(A), p_{\text{min}}(A)) \left( 1 - \frac{\sum_{i=1}^{n} |q(c_i)|}{2 \sum_{i=1}^{n} \hat{p}(c_i)} \right)$$

where $I_{1,A}$ is the result of indicator 1 for analysis area $A$ and $n$ is the number of cells in $A$. The denominator is used to normalize the indicator result to the range $[0, 1]$ with 1 denoting $\sum_{i=1}^{n} |q(c_i)| = 0$, i.e. no deviation from the reference density ($\sum_{i=1}^{n} |q(c_i)|$ can be at most equal to $2 \sum_{i=1}^{n} \hat{p}(c_i)$). Note that $\sum_{i=1}^{n} \hat{p}(c_i) = p_A(A)$. 

13
3.1.3 Parameters

Two settings and two parameters are needed for the application of indicator 1:

- raster grid resolution and bandwidth for kernel density estimation
- benchmark values for achievable density $p_{min}(A)$
- distance-decay function $z(d_{s,t})$

The following subsections present how these parameters are defined for analysis in Switzerland.

**Raster grid resolution and bandwidth for kernel density estimation**

Bandwidth will be chosen based on the mean nearest-neighbor distance (Euclidean) between activity points, as suggested in section 3.1.1. Raster grid resolution will be chosen accordingly: cell size should not exceed 0.1 x bandwidth, and small cells are generally more desirable as long as computation remains feasible. A quick analysis based on the city of Zurich yielded a value of 40m for bandwidth; accordingly, raster resolution would be 4m. However, these values will be separately evaluated for each concrete analysis area, and a range of bandwidth values will be used and respective results compared to conclude whether the “rule of thumb” bandwidth = mean nearest-neighbor distance is adequate for application with this indicator.

**Benchmark values for achievable density $p_{min}(A)$**

For the definition of $p_{min}(A)$, characteristics of the analysis area $A$ need to be evaluated, and density values for comparable situations collected. Achievable could be defined, for example, as the 75th percentile of density values in such comparable situations.

**distance-decay function $z(d_{s,t})$**

The decay function depends on local circumstances (how far are people willing to walk to and from PT in a specific country, city, etc.) and needs to be adapted to context. For a preliminary analysis, the following, simplified form is suggested (see Figure 4).
The constant value of 1 for the range [0m, 100m] of $d_{s,t}$ is chosen because an immediate decay starting at $d_{s,t} = 0$ seems unrealistic. It should be noted that this value is chosen arbitrarily and further investigation as to its validity is necessary.

The value of $d_{\text{max}}(t)$ is a function of mode and service characteristics of PT stop $t$. For Switzerland, ARE (2015a) provides a classification. It contains 15 stop classes based on the combination of three “modes” (railway hub, railway stop, and urban mode) and five service headway groups (< 5min, 6-9min, 10-19min, 20-39min, and 40-60min). It also suggests maximum catchment area Euclidean distances for each class. The classification is summarized in Table 1.

Table 1  Suggested values for $d_{\text{max}}(t)$ in the Swiss context [m]

<table>
<thead>
<tr>
<th>stop type</th>
<th>headway [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 5</td>
</tr>
<tr>
<td>railway hub</td>
<td>1500</td>
</tr>
<tr>
<td>rail stop (railway, rapid transit)</td>
<td>1500</td>
</tr>
<tr>
<td>urban mode (tram, bus)</td>
<td>1000</td>
</tr>
</tbody>
</table>

Source: Adapted from methodology for assigning PT rating classes (“Güteklassen”) (ARE, 2015a); values of 1500m were chosen for stop categories where no class D was reached.
The values in Table 1 reflect upper thresholds of any catchment by PT. Comparing them with international values (Biba et al., 2010; El-Geneidy et al., 2014) shows that they are rather high and suggests that if at all, the values should rather be used as network distance thresholds than Euclidean distance.

For a rough validation of the values in Table 1 and to gain some insights into walking distances to and from PT in Switzerland, data from the microcensus mobility and traffic 2010 (BFS/ARE, 2012) were analyzed. The dataset contains 211’359 trips with 310’193 legs in total. Among many other variables, data for each leg contains mode, origin and destination coordinates, and Euclidean distance. Data for trips includes a main mode. Access and egress legs to and from PT were identified by selecting all trips with PT (only train, tram and bus was used) as the main mode of transport, and for these trips selecting walking legs which are the first leg of the trip and directly precede a PT leg (access) or which are the last leg of the trip and directly follow a PT leg (egress). For each access and egress leg, also the respective type of PT that was accessed or egressed from was recorded. Only legs of trips that start and end within urban areas (spatial structure core municipality of agglomeration (“Agglomerationskerngemeinde”) or other municipality of agglomeration core (“Übrige Gemeinde der Agglomerationskernzone”) within Switzerland were considered. Furthermore, legs with Euclidean distance smaller than 10 meters or larger than 2 kilometers were excluded because they are likely mistakes in the dataset or leisure trips (e.g., hikes). Finally, legs which did not contain the most accurate level of origin and destination data as defined by the microcensus were also excluded. With these restrictions, 12’524 access and egress legs were used for analysis.

For all analyses, the weight given by the microcensus for the respective respondent undertaking a leg was used. Only Euclidean distances were analyzed because routing distances were stated not to be accurate for walking legs (Scherer Ohnmacht, 2012). Results are presented in Table 2, Figure 5, and Figure 6.
Table 2  Euclidean walking access and egress distances to and from PT in Switzerland [km]

<table>
<thead>
<tr>
<th>mode</th>
<th>N</th>
<th>quantile</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.10</td>
<td>0.25</td>
<td>0.50</td>
<td>0.75</td>
<td>0.90</td>
<td>mean</td>
<td></td>
</tr>
<tr>
<td>train access</td>
<td>113</td>
<td>0.119</td>
<td>0.198</td>
<td>0.342</td>
<td>0.510</td>
<td>0.945</td>
<td>0.418</td>
<td></td>
</tr>
<tr>
<td>train egress</td>
<td>111</td>
<td>0.115</td>
<td>0.172</td>
<td>0.314</td>
<td>0.538</td>
<td>0.982</td>
<td>0.438</td>
<td></td>
</tr>
<tr>
<td>tram access</td>
<td>2246</td>
<td>0.054</td>
<td>0.106</td>
<td>0.189</td>
<td>0.327</td>
<td>0.545</td>
<td>0.271</td>
<td></td>
</tr>
<tr>
<td>tram egress</td>
<td>2203</td>
<td>0.058</td>
<td>0.108</td>
<td>0.190</td>
<td>0.326</td>
<td>0.550</td>
<td>0.269</td>
<td></td>
</tr>
<tr>
<td>bus access</td>
<td>3982</td>
<td>0.052</td>
<td>0.099</td>
<td>0.177</td>
<td>0.309</td>
<td>0.520</td>
<td>0.256</td>
<td></td>
</tr>
<tr>
<td>bus egress</td>
<td>3869</td>
<td>0.056</td>
<td>0.103</td>
<td>0.184</td>
<td>0.320</td>
<td>0.544</td>
<td>0.270</td>
<td></td>
</tr>
</tbody>
</table>

Source: BFS/ARE (2012)

Figure 5  Euclidean walking access and egress distances to and from PT in Switzerland

Source: BFS/ARE (2012)
Figure 6 Empirical cumulative distribution function of Euclidean walking access and egress (combined) distances for different PT modes in Switzerland

Source: BFS/ARE (2012)

Note that the figure does not depict distance-decay functions because the values are not normalized with the number of activities at a certain distance from PT.

This analysis of observed Euclidean access and egress distances in Switzerland supports the notion stated above that values in Table 1 are on the high side. Therefore, for application of indicator 1, they will be used as network (rather than Euclidean) distance thresholds $d_{\text{max}}(t)$ in the simplified distance-decay function shown in Figure 4:

$$z(d_{s.t}) = \begin{cases} 
1, & d_{s.t} < 100 \\
\frac{100 - d_{s.t}}{d_{\text{max}}(t) - 100} + 1, & 100 \leq d_{s.t} \leq d_{\text{max}}(t) \\
0, & d_{s.t} > d_{\text{max}}(t)
\end{cases}$$
3.2 Indicator 2: Influences of the pedestrian network on the size of PT catchment area

3.2.1 Theoretical rationale and background

As section 3.1.1 shows, the distance to and from PT strongly affects the probability that potential PT users actually use PT. Therefore, often the area within a certain network or airline distance (e.g., the ) from PT stops is defined as their catchment or service area. The size of the catchment area depends on the pedestrian network – or more precisely, the so called detour factor, which is defined as the ratio between network and Euclidean distance:

\[
u = \frac{d}{g}\]

where \(u\) is the detour factor, \(d\) is the network distance, and \(g\) the Euclidean distance.

For PT, the catchment area should be as big as possible, since this maximizes the area for potential locations of activities with access to PT. Therefore, the ideal pedestrian network for PT provides direct, i.e. detour-free, access to PT from every point around it – in such a theoretical case, the catchment area would be a circle around a PT stop. In reality, this is almost never fulfilled entirely; hence, the aspiration should be to keep the detour as small as possible.

Different approaches exist that could be used to assess the performance of a pedestrian network related to providing access to PT. The most prominent are the ratio between actual and theoretically possible catchment areas of a PT stop, the so-called pedshed ratio or stop coverage ratio (Schlossberg and Brown, 2004; Foda and Osman, 2010), the network ratio (ratio of length of the street network within walking distance to PT and total length of street network in circular catchment area) (O’Neill, Ramsey, and Chou, 1992; Hsiao et al., 1997), detour and obstacle factors (Olszewski and Wibowo, 2005; Jiang, Zegras, and Mehndiratta, 2012), and parcel-network method (using population data per parcel and network distances) (Biba et al., 2010).

Most of these approaches have been developed to provide an estimate of PT ridership, whereas here the goal is to evaluate how well the pedestrian network provides access to PT. The explicit use of catchment areas (e.g., pedshed ratio) requires the definition of a maximum viable walking network access distance – which is not adequate because there is no absolute threshold but rather a decay of probability to walk with increasing distance, as shown in section 3.1.1. Therefore, each distance counts, i.e. for each distance the size of the respective catchment area should be maximized.

Basically, this means that the pedestrian network should provide access to the closest PT stop with as little detour as possible for every point within the built-up area. To assess this, a
combination of methods assessing walking distances to PT stops from every activity, such as Biba et al. (2010) (using parcels as activity proxies), with the detour factor for each of these distances is chosen, following a similar approach as Meeder (2015). As an approximation of “every point”, all locations of activities are used. For each activity location, the detour factor for access to the closest PT stop is measured. This approach also conveniently accounts for areas such as parks or farmland, where there are no activities and therefore direct access to PT is not necessary, but that may act as barriers if they are located between activities and a PT stop.

3.2.2 Performance measure

As with indicator 1, there are two scale levels to consider here: disaggregated for each activity point and aggregated for each PT stop (and the area for which that stop is the closest).

First, for every point with an activity \( s_i \), the detour factor is defined as

\[
u(s_i) = \frac{d_{s_i,t}}{g_{s_i,t}}
\]

where \( d_{s_i,t} \) and \( g_{s_i,t} \) are network and Euclidean distance between \( s_i \) and the closest (with network distance) PT stop \( t \), respectively. This disaggregated performance measure can be visualized on a map, e.g. by attributing different colors to the range of \( u \).

Second, for a PT stop \( t \), the mean of all \( u(s_i) \) for all \( s_i \) for which \( t \) is the closest PT stop (with network distance) is calculated:

\[
U_t = \frac{1}{n} \sum_{i=1}^{n} u(s_i)
\]

where \( U_t \) is the mean detour factor for stop \( t \) and \( n \) is the number of \( s_i \) for which \( t \) is the closest PT stop (with network distance). Indicator 2 then linearly transforms \( U_t \) to a value in the range \([0,1]\), with 1 representing the “ideal” result of direct, detour-free access from every point, and 0 a mean detour factor equal or larger than a reference detour factor \( U_{ref} \):

\[
I_{2,t} = \begin{cases} 
\frac{1 - U_t}{U_{ref} - 1} + 1, & U_t \leq U_{ref} \\
0, & \text{otherwise}
\end{cases}
\]

where \( I_{2,s} \) is the result of indicator 2 for stop \( t \). \( U_{ref} \) needs to be defined based on what is acceptable in a specific context.

Note that for indicator 2, the activity points are not weighted with the number of activities at each point. This is because, as stated above, using activity points for the calculation is simply

20
an approximation for “every point in built up area” and conveniently excludes undeveloped areas such as parks. Indicator 1 already assesses where activities (considering the number of activities as density) are located relative to PT stops.

### 3.2.3 Parameters

Only one parameter needs to be defined for application of indicator 2: the reference detour factor $U_{ref}$. Several sources state acceptable detour factors or factor ranges for pedestrian friendly neighborhoods. For example, ASTRA (2015) states that a detour factor of 1.1 for train and 1.2 for tram and bus is acceptable for PT access and egress trips. Meeder (2015) conducted an international literature review and found values between 1.2 and 1.5 described as thresholds for pedestrian friendly neighborhoods. He also observed that for regular grid patterns, the detour factor is never larger than 1.5 for routes spanning at least two blocks. Furthermore, he analyzed detour factors from PT stops to all activity points (including residences) for several neighborhoods in Switzerland and found average values ranging from 1.24 to 1.39. Given these insights, a threshold value of $U_{ref} = 1.4$ seems adequate for the Swiss context.

### 4. Conclusions and outlook

In this paper, a novel approach for neighborhood-scale PT and BE integration assessment that is based on a qualitative model of interactions has been presented. The feasibility of quantitative indicator development derived from this theoretical and qualitative base has been demonstrated with two concrete indicators related to pedestrian access and egress.

The logical next research step is application of these two indicators to case studies to complete the demonstration of concept feasibility and allow for comparison with existing approaches. Furthermore, 16 more measuring points have been identified in the qualitative system model and quantitative indicators will be developed for each of them. Given that each indicator reflects the magnitude of a mutual influence, and all indicators together cover all the mutual influences discovered in the qualitative model, the combination of all indicators will deliver a holistic assessment of integration. To assess integration as a whole, a composite metric is needed. Therefore, once all indicators are developed, they will be combined to a spatial index of PT and BE integration at neighborhood scales.

There are also several future research needs related to the two indicators presented in this paper. For indicator 1, a classification of analysis areas $A$ with criteria such as distance to CBD, closest subcenter, or PT hub, uses, zoning types, and the type of PT stops within $A$ and subsequent definition of benchmark values for achievable density $p_{min}(A)$ based on a broad analysis of Swiss density values seems promising. Such a classification with preset values of $p_{min}(A)$
would make application of indicator 1 possible without first analyzing density at comparable situations. Another important improvement of indicator 1 would be the precise estimation of distance-decay functions \( z(d_{s,t}) \) for Switzerland – based on an extended analysis of network distances of walk access and egress trips in the microcensus dataset and normalization with activity numbers relative to \( d_{s,t} \) for the same analysis area (essentially the entire urbanized land in Switzerland). Ideally, such decay functions would be estimated for different PT stop types, for example following the classification presented in section 3.1.3. For indicator 2, obstacles such as steps, pedestrian crossings, over- and underpasses, traffic signals, as well as elevation could be considered in the detour factor, for example following the approach of “equivalent walking distance” by Olszewski and Wibowo (2005) or using the values provided by ASTRA (2015). Furthermore, a literature review on the disutility of detour factors should be conducted to ensure that the linear relationship that was assumed in the performance measure of indicator 2 (section 3.2.2) is realistic.

The two indicators presented here are a first step towards a method for holistic quantitative assessment of PT and BE integration at the neighborhood scale and for an increased understanding of what this integration constitutes exactly. The result – a spatial index of integration – will be useful for different applications. It will facilitate the identification of decisive factors for integration by comparing index results for different cases and highlighting the role of specific indicators (individual indicator values). For projects with explicit integration goals, for example approved measures of the agglomeration programs in Switzerland, ex-post analysis with the index could be used to evaluate how successfully these goals have been achieved. For project development where integration is one of the goals, alternatives could be compared for their integration performance. Even without alternatives to compare, suitability of a specific project could be evaluated using benchmark values from other, similar cases. And individual indicator values could be used to identify strong and weak points (topics and locations) influencing the index result, which could inform manual intervention design. Alternatively, a large number of alternatives could be generated with parameter variation, which would open up powerful opportunities to employ evolutionary algorithms for finding optimized solutions, using the index as the fitness function.

5. References


ARE (2015a) ÖV-Güteklassen, Berechnungsmethodik, Aktualisierung 02.2015, Bundesamt für Raumentwicklung, Bern.


Zhao, F., L.F. Chow, M.T. Li, I. Ubaka, and A. Gan (2003) Forecasting transit walk accessibility - Regression model alternative to buffer method, Transportation Research Record, 1835 34-41.