Spatio-Temporal Investigation of Public Transport Demand Using Smart Card Data

Robert Klar[1,2] · Isak Rubensson[3,4]

Abstract

Policymakers must find efficient public transport solutions to promote sustainability and provide efficient urban mobility in the course of urban growth. A growing number of research papers are applying Geographically weighted regression (GWR) to model the relationship between public transport demand and its influential factors. However, few studies have considered the rapid development of journey inference from ticket transaction data. Similarly, the potential of GWR to analyze spatio-temporal changes that reflect changes in transportation supply and thus provide a measure for evaluating the local success of transport supply changes has yet to be exploited. In this paper, we use inferred journeys from smart card inferences as the dependent variable and analyze how public transport demand responds to a set of explanatory variables, emphasizing transport supply. Consequently, GWR and its successor Multiscale Geographically Weighted Regression (MGWR) are applied to analyze the spatially varying impact of transport supply changes for seven consecutive time frames between autumn 2017 and spring 2020, allowing conclusions about local changes in transport demand, as well as the benchmarking of transport supply changes. The (M)GWR framework's predictive power is evaluated by training the model with past transport supply data and testing the model with data from the following consecutive years. The conducted analyses reveal that the (M)GWR model, using inferred journeys and transport supply data, can retrospectively predict the impact of transport supply changes on travel behavior and thus provides conclusions about the success of transport policies.

Isak Rubensson contributed equally to this work.

✉ Robert Klar
robert.klar@liu.se

Extended author information available on the last page of the article

Introduction

Public transport plays a vital role in society as it provides affordable basic mobility for everyone and efficient and sustainable urban travel. At its best, public transport contributes to more efficient land use development, increased air quality, and improved space allocation within the city (Litman, 2022; Miller et al., 2016). Moreover, public transport's role in decarbonization and urbanization is crucial in meeting the European Union's (EU) commitments to global climate action under the Paris Agreement, as transportation is one of the major and fast-growing emitters of CO2 (Giannakis et al., 2020) Public transport is a crucial driver of transportation sustainability and a tool to address the sustainability impacts of car dependency by providing energy-efficient transportation in urban environments that can compete with the speed of personal transportation (Miller et al., 2016). The benefits of public transportation are not limited to the environmental aspects of sustainability. Public transportation also provides social sustainability benefits due to equal access, affordability, and positive impacts on human health due to reduced noise pollution (Schiller & Kenworthy, 2017). In addition, public transport also contributes to economic sustainability by reducing traffic congestion and mobility barriers while providing enhanced accessibility (Miller et al., 2016; Litman & Burwell, 2006). Consequently, public transportation affects a city's economy, residents' quality of life, and degree of sustainability (Vuchic, 2008). With increasing urbanization and growing environmental concerns, the public transport policymaker must find even more effective solutions weighing trade-offs between investment costs, environmental impacts, finance, and sustainability (May et al., 2000). Towards meeting these challenges and understanding the effect of different policies, robust forecasting of public transport demand is of utmost importance.

In recent years direct models of public transport demand using geographically weighted regression (GWR) methods gained traction in the literature (Tang et al., 2021; Ma et al., 2018; Chiou et al., 2015; Cardozo et al., 2012; Blainey, 2010; Yang et al., 2020; He et al., 2019; Blainey & Preston, 2013). GWR models commonly yield better model fit and reflect the non-stationary quality of independent variables (Cardozo et al., 2012; Blainey, 2010; Blainey & Preston, 2013). Improving further on the GWR, newer studies let the data decide if coefficients should vary spatially or be globally constant, introducing the framework of mixed GWR models. Mixed GWR was introduced (Tang et al., 2021; Yang et al., 2020; He et al., 2019) with further improvements in model fit. In addition, Ma et al. (2018) extended the model to encompass daily temporal variation (GTWR). An additional important development was the advent of the multi-scale geographically weighted regression (MGWR), which allows the parameters to have different numbers of relevant neighboring areas (Lyu et al., 2020; An et al., 2022; Cao et al., 2021).

However, most earlier studies use boardings' or entrances in the public transport system as the dependent variable, sometimes limiting the data to a subset of all public transport, e.g., only metro stations. Since the policy planner aims to facilitate users' complete journeys encompassing the whole public transport system, direct demand models should have the same scope. Further, forecasting and understanding

long-term temporal variation (year on year) of geographical differences in public transport demand, e.g., the impacts of significant changes in supply, could be improved.

In this study, our case is the Stockholm County Public Transport system, where journey inference from automated fare collection is in effect, and we can infer journeys as long back as 2017. Among the advantages of this approach is that we can consider actual complete journeys rather than boardings. From the usage pattern of the electronic tickets, we can also attribute these journeys to their correct home zone.

This study aims to add to the existing body of research in two ways: applying GWR and MGWR on whole journeys and training and evaluating a predictive GWR model over consecutive time frames. We start our study by showing the development of public transport usage from 2017 to 2020, indicating, among other factors, the discernible impact of a substantial infrastructure investment in the commuter rail system (a new railway tunnel, two new central stations, and increased supply) introduced during the summer of 2017. Then we use this new rich data set to construct direct demand models with inferred journeys as the dependent variable. We develop three types of direct demand models using Ordinary Least Square (OLS), GWR, and MGWR techniques. The models are estimated for seven consecutive timeframes, from spring 2017 up to spring 2020, just before the outbreak of Covid19. Last, we assess the predictive power of GWR versus OLS by training models on one timeframe and then forecasting subsequent time frames.

The remainder of this paper is structured as follows. The literature review is presented in "Literature Review", providing a concise overview of the various factors affecting public transport usage ("Factors Affecting Public Transport Usage") and a brief summary of GWR applications within transport science ("GWR Based Modelling within Transport Science"). The study area and the data, as well as the public transport system within Stockholm County, are presented in "Data". The applied methodology is introduced in "Methodology", covering journey inference ("Journeys Per Capita as Dependent Variable"), the feature selection process ("Feature Selection"), and the (M)GWR estimation framework ("(Multiscale) Geographically Weighted Regression"). The results are presented in "Results". Spatio-temporal trends are presented first ("Spatio-Temporal Trends"), followed by results from the (M)GWR framework ("(M)GWR"), including model fit comparison ("Model Fit Comparison") and model prediction ("Prediction Results"). Finally, the results of this study are critically discussed in "Discussion", followed by concluding remarks in "Conclusion".

Literature Review

The aim of the literature review in this section is to identify which factors influence the use of public transport ("Factors Affecting Public Transport Usage") and to what extent these have already been used to predict transport use using GWR ("GWR Based Modelling Within Transport Science").

Factors Affecting Public Transport Usage

A clear understanding of how and to what extent various factors affect public transport demand is needed if public transport is to contribute to solving transport-related environmental problems (Holmgren, 2013).

The various factors determining the public transport usage rate are extensively examined in the existing literature. In general, most research results conclude that the public transport use rate is positively correlated with denser built environments, spatial heterogeneity of land use, higher accessibility, and more walking-supportive urban design (Tsai et al., 2012). The impact of the built environment is further investigated by Gascon et al. (2020), stating that higher density (there defined as areas with high street length density and connectivity, population density, and density of public transport stations) of the residential or work area clearly increased the odds of using public transport. The results by De Vos et al. (2020) further imply that raising public transport service frequency, living in an urban area, and limited access to a car positively affect PT use. Consequently, the use of public transport tends to increase in more densely populated areas, while there is an inverse relationship between car use and population density. Travel mode choices are also influenced by the degree of centralization of employment and facilities, as a greater centralization encourages public transport use Balcombe et al. (2004).

The quality of service may be defined by a wide range of attributes which can be influenced by planning authorities and transport operators by certain transport supply policies. Some of these attributes (frequency, access time - walking time to the next stop or in-vehicle time) are measurable and are incorporated in many demand forecasting models as these attributes have a direct effect on the transport demand (Paulley et al., 2006). The importance of public transport service frequency and walking distance, which have significant effects on the probability of choosing PT over car on commute trips, is also confirmed by Lunke et al. (2021). Further studies by Yeboah et al. (2019) reveal that travel frequency, sociodemographic characteristics, travel context (travel planning stages, preferred mode of transportation), and the provision of information are significant predictors of public transport use.

Besides transport supply, the density of the built environment, and land use, socioeconomic factors also play a major role, as socioeconomic factors are relevant for explaining differences regarding mode choice and travel behavior between areas (Balcombe et al., 2004). It appears that changes in public transport use among residents are related to changes in car ownership, the number of adults and children per household, household income, educational background, and quality of and satisfaction with public transport (De Gruyter et al., 2022). Recent research about the explanation of travel behavior based on socioeconomic factors during the COVID-19 pandemic in Stockholm reveals that the choice of public transport is heavily influenced by socioeconomic variables (Almlöf, Rubensson, Cebecauer, & Jenelius, 2020). The study concludes that people with fewer resources are more dependent on public transport and, consequently, are more likely to use public transport.

In summary, drawing on the work by Gascon et al. (2020), key elements of strategies to enhance public transport usage should include the improvement of the nearby built

environment, promotion of and planning for more densely populated urban areas, incentives to use public transport through barrier-free and pedestrian-friendly access to stations, investment in and promotion of high quality, reliable and affordable public transport services, and the curtailment of private motorized vehicles in high-density areas of the cities. Vice versa, the most important factors favoring car usage over public transport are car access (both physical via the road network and socioeconomic), travel time, travel cost, trip importance, trips outside the city center, weather, flexibility, and accessibility to PT stations (Nguyen-Phuoc et al., 2018). Consequently, well-planned transport policies that consider all the various factors are crucial to promote public transport to enhance its demand and support more sustainability. Such policies could include more and better service, attractive fares, and convenient ticketing, full multimodal and regional integration, high taxes and restrictions on car use, and land use policies that promote a compact development to support public transport (Buehler & Pucher, 2012).

GWR Based Modelling Within Transport Science

Already in the years of their early development starting in 1996, GWR-based models have been tested and applied for various applications such as the prediction of car ownership rates based on socioeconomic variables (Brunsdon et al., 1996) as well as the spread of illnesses based on socioeconomic variables (Brunsdon et al., 1998). Furthermore, since the advent of GWR within spatial science, GWR models have been widely used within the field of transport to examine the spatial variation of several target variables of interest, such as traffic accidents (Hadayeghi et al., 2010; Pirdavani et al., 2014), accessibility (Du & Mulley, 2006, 2012; Dziauddin et al., 2015), annual average daily traffic (Zhao & Park, 2004), and public transport ridership (Chow et al., 2006; Chiou et al., 2015), as well as car ownership estimation (Clark, 2007). In recent years, GWR has been extensively applied within the field of station-revel ridership modeling, as each station can be viewed as a geographic unit that can be covered by a local model. These local models estimate the number of passengers boarding at each station as a function of the station characteristics and the surrounding area that they serve (Cardozo et al., 2012) or the built environment (Gao et al., 2022), but also as a function of the land use near the stations and the characteristics of the transit system surrounding the stations (Marques & Pitombo, 2022). These station-level models have been applied both on a national scale (Blainey, 2010) and on an urban scale (Cardozo et al., 2012; Blainey & Mulley, 2013; He et al., 2019; Marques & Pitombo, 2022). In addition to further developments in the fields of application, the methodology itself has also evolved. For example, recent research applied multiscale geographically weighted regression (MGWR) to further enhance the model fit compared to GWR. Compared to standard GWR, MGWR allows the relationship between the response and a covariate to vary locally, regionally, and or not vary at all Oshan et al. (2019). Consequently, MGWR assigns different bandwidths for each explanatory variable. MGWR models have been recently applied for modeling ridership on metro station level (Yang et al., 2020; Liu et al., 2023), to evaluate rail transit station accessibility (Li et al., 2023), to assess the correlation between transport and property values

(Liu et al., 2022) or the built environment (Zhu et al., 2023), and to examine the impact of shared mobility trips on taxi and public transit ridership (Tang et al., 2021).

Recent research also underlines the potential of GWR as a tool to examine spatio-temporal impacts. Geographically and temporally weighted regression has been successfully applied to explore the spatio-temporal influence of built environment on transit ridership (Ma et al., 2018). Similar to the study we conducted, smart card data was used to determine public transportation ridership per zone. A local model was created to determine the spatiotemporal impact of the built environment on public transit ridership. Our study aims to tie into this by using whole journeys instead of boardings and by retrospectively examining the effects of past infrastructure investments and estimating the possible effects of future infrastructure investments.

Data

In this study, Stockholm County's public transport is investigated, and this section outlines what data is included.

This study is performed for Stockholm County, Sweden, population of 2.39 m. Stockholm comprises 26 municipalities, with Stockholm City as the largest (pop. 0.98m). The land-based public transport has four main modes; a central metro network (100 stations, 1.3 m daily boardings), a long-range commuter train network (54 stations, 0,4 m boardings), five light rail/tram systems (117 stations, 0.2m boardings) and, spanning the whole county, an extensive bus network (6700 stops, 1,2m boardings). Figure 1 represents Stockholm municipality and its public transport
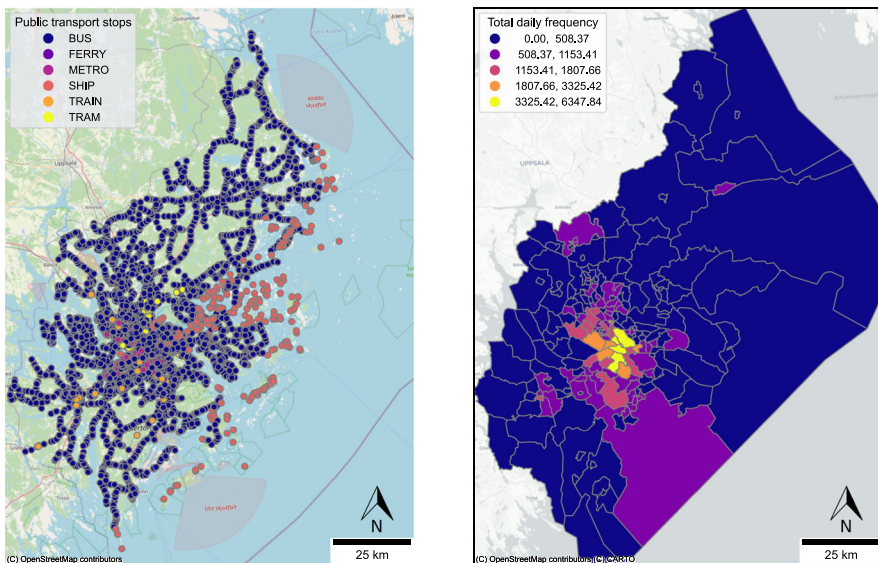
(a) Public transport network in Stockholm   (b) Derivated zonal transport supply

Fig. 1 Overview of Stockholm County and its transport supply

network. Subfigure a presents the whole public transport network within Stockholm municipality, and Subfigure b illustrates the derivated zonal transport supply for each of the 218 administrative zones termed as "planområden" in Sweden. These zones are further referred to as small areas. The Fisher Jenks classification method, also known as Jenks natural breaks algorithm, is used to classify the derived zonal transport supply into a predefined number of homogeneous classes (North, 2009).

The applied data of this study is divided into four distinct categories: Transport supply, Centrality measures, Land-use, and Socioeconomic. The transport supply data consists of various transport supply information on the stop level, the transport network, and car ownership data. The centrality of each small area within the road network and the distance of each small area to the central transport hub of Stockholm as well as their inverses to give small areas in proximity a higher weight, are covered in the centrality measures category. The land-use characteristics contributing to enhanced urban density are categorized as land-use, while all social, demographic, or income-related variables are categorized as socioeconomic.

The data applied in this study are described in this section. Table 1 presents the utilized data and their descriptive statistics. In addition to the presented distance and centrality measures depicted in Table 1, the inverse Euclidean and the inverse network distance are computed as well. The aim of these two additional explanatory variables is to give small areas in proximity to the main transportation hub (T-Centralen) a higher weight.

The data for the timeframe spring 2017 to spring 2020 are obtained from Statistics Sweden, Stockholm Public Transport Administration, and OpenStreetMap. Table 2 presents the sources and time scale for each classification. All data is organized geographically into the 218 administrative zones.

Methodology

This section briefly review the performed methodology, including feature selection, the theoretical framework behind (M)GWR, model evaluation and visualization process as well as the retroactive prediction framework. All analyses are performed within the Python ecosystem using the mgwr package (Oshan et al., 2019). A brief overview of the applied methodology is illustrated in Fig. 2. At the time of the study, prediction with MGWR is not yet available; therefore prediction is limited to GWR and OLS.

Journeys Per Capita as Dependent Variable

In this study, we use the maturing methodologies of mining the rich data from ticket transactions (Pelletier et al., 2011). This type of data makes it possible to follow, with very high resolution, public transport users' behavior over time by inferring their trips and journeys. Competing sources for this type of data, such as travel survey data and automatic passenger count data, are disadvantaged by the cost of collection and difficulty in tracking full travel patterns. The methodologies for inferring travel behavior from ticket transactions have had several applications in recent years.

Table 1 Overview about the collected data

Classification	Variable	Mean	Standard Deviation	Min	Max
Dependent variable	Journeys per capita ratio	0.35	0.42	0.00	3.13
Transport supply	PT stop count	51.71	62.86	18.25	480.00
	PT stop density	6.06	6.91	0.00	41.72
	Metro stops	0.46	1.67	0.00	14.00
	Railway stops	0.22	0.50	0.00	4.00
	Tram stops	0.29	1.50	0.00	15.00
	Ferry stops	0.52	5.31	0.00	77.00
	Total frequency	665.48	896.35	0.00	6648.32
	Maximum Frequency	145.03	84.65	0.00	357.73
	Number of lines	12.13	11.94	0.00	92.00
	Road network length	71.84	96.79	0.00	713.42
	Car ownership rate	30.73	7.99	0.00	49.75
Centrality measures	Closeness centrality	0.01	0.002	0.00	0.02
	Mean Pagerank	1e-05	2e-06	0.00	2e-05
	Total Pagerank	0.005	0.005	0.00	0.03
	Network distance till centre	26.77	18.62	1.11	100.23
	Euclidean distance till centre	21.68	15.69	0.87	89.76
Land-use	Industrial area	0.25	0.47	0.00	4.75
	Built-up area	2.54	3.57	0.01	30.92
	Industrial density	0.04	0.08	0.00	0.59
	Built-up density	0.25	0.20	0.00	0.93
Socioeconomic	Population	10731	17324	6	129766
	Population density	1860.32	2642.96	0.64	19178.37
	Single households	23.99	10.51	9.77	100.00
	Shared households	60.03	12.55	0.00	90.23
	Proportion of minors	25.15	5.83	0.00	37.67
	Proportion of elders	17.03	6.48	0.00	41.67
	Proportion of workers	57.82	6.57	41.67	100.00
	Higher education	31.89	11.07	11.20	100.00
	Mean yearly net income	276011.22	62288.88	145489.00	511113.00

Table 2 Sources and scales for the collected data

Classification	Source	Time Scale
Journeys per capita	Stockholm Public Transport Administration	For each of the 7 time frames
Transport supply	Stockholm Public Transport Administration	For each of the 7 time frames
Centrality mesures	OpenStreetMap	Constant
Land-use	Statistics Sweden	For each of the 7 time frames
Socioeconomics	Statistics Sweden	For each of the 7 time frames

Kholodov et al. (2021) assessed ticket fare elasticities for different social groups. Cats & Ferranti (2022) showed how to cluster users based on their travel patterns. Chen & Zhou (2022) demonstrated the effects of a fare increase on travel demand, crowding, equity, and chosen travel start time. Yap et al. (2017) showed in a noteworthy study that the planner must apply inferring algorithms with caution when studying public transport disruptions due to the seemingly erratic traveler behavior with continuously changing route choices (to find optimal decisions in a changing environment).

Kholodov et al. (2021) developed the specific methodology used in this study for trips and journeys in the public transport system of Stockholm. Stockholm has a tap-in smart-card system, where users tap their smart-card ticket carrier on the busses and trams or at the turnstiles in metro- and commuter rail stations. There is no tap-out when exiting a vehicle or station. The algorithm works with three fundamental assumptions; first, each tap-in defines the chosen destination for the previous tap-in. Second, the interval between two tap-ins determines if the second tap-in is a transfer point or a destination where the traveler performed some activity. Third, each smart-card has a decided home stop inferred from the most frequent place for first validations on a day. Output from the model is a database where each card has a defined home stop, and all inferred journeys and their corresponding trip legs are listed.

In our analysis, we study two periods during the years, spring (weeks 3-6) and autumn (weeks 39-42). We chose these weeks because they correspond to periods of stable high public transport demand. It is also a benefit that the spring period of
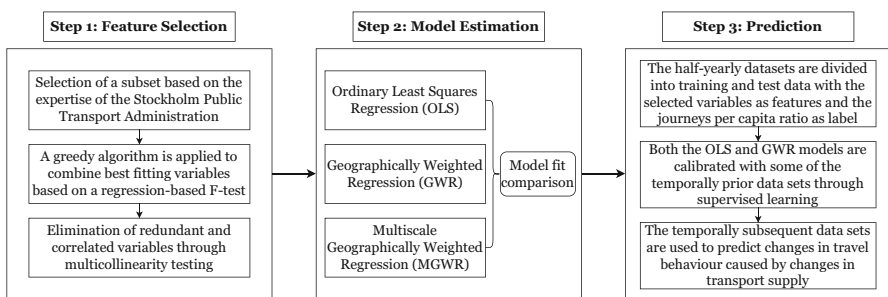

Fig. 2 Brief overview of the methodology applied in this paper. Step 1: Feature selection refers to "Feature Selection", Step 2: Model Estimation refers to "(Multiscale) Geographically Weighted Regression", and Step 3: Prediction refers to "Retroactive Prediction of Changes in Transport Supply"

2020 is before the outbreak of Covid19. All journeys are extracted and clustered per administrative zone for the chosen periods. The total number of journeys is then divided by zonal population and the number of days to arrive at journeys per capita for a representative day. The journey per capita variable is the dependent variable in our analysis.

Feature Selection

Feature selection is often characterized by its three-fold objective - enhanced prediction performance of the predictors, reduction of the training time, and a better understanding of the underlying process and the results (Guyon & Elisseeff, 2003). Consequently, prediction accuracy is improved while the risk of overfitting is reduced. Feature selection is a challenging task within local modeling (Lu et al., 2014) as the decision to include and exclude a potential variable is enhanced by the question on which scale the variable should be included (local or global). In order to overcome these challenges, a subset of variables is chosen based on expertise from the Stockholm public transport administration and a greedy algorithm evaluating the feature score of each variable based on a regression-based F-test. The remaining variables in the subset are further reduced by removing redundant variables as well as by performing multicollinearity tests. The spatial scale of the variables and the best-fitting model are determined by performing an AIC-based procedure.

(Multiscale) Geographically Weighted Regression

This section presents the theoretical background of the OLS, GWR, and MGWR models applied in this study and outlines how results are evaluated and mapped.

Geographically Weighted Regression

Regression is the most frequently used statistical modeling approach for the analysis of spatial data (Fotheringham & Rogerson, 2008). Traditional Ordinary least squares (OLS) regression models are based on the assumptions that observations are independent of one another and that the relationships being modeled are the same everywhere within the study area of which the data is collected (Charlton et al., 2009). However, these assumptions may not be validated for spatial data as Tobler's first law of geography states that "Everything is related to everything else, but near things are more related to distant things" (Tobler, 1970). Spatial data is characterized by the existence of spatial autocorrelation (spatial dependence) as well as spatial heterogeneity (spatial non-stationary) (Longley et al., 2005). As spatial data exhibits these spatial characteristics, it is required to allow the parameter estimates in the model to vary over space. "The specification of a model that allows the parameter estimates to vary over space is the essence of geographically weighted regression" (Fotheringham & Rogerson, 2008). Consequentially, GWR calibrates a separate regression model at each location through a "data-borrowing scheme that distance-weights observations from each location serving as a regression point" (Oshan

et al., 2019). Consequently, the results can be projected to a map to visualize the spatial distribution of the estimated parameters (Matthews & Yang, 2012).

$$\gamma_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + e_{ik}, \quad i = 1, ..., n \tag{1}$$

The GWR model takes the spatial component of the data into account by incorporating the location i in its equation. In contrast to OLS regression, the coefficients β_{i0} for each predictor x_{ik} vary over space. For each location i, the value of the dependent variable γ_i is calculated by Eq. (1).

$$\hat{\beta}_i = [X'W(i)X]^{-1} X'W(i)y, \tag{2}$$

Equation (2) expresses the GWR estimator for local parameter estimates at location i in matrix form where X is a n by k matrix of explanatory variables, $W(i) = diag[w_1(i), ..., w_n(i)]$ is the n by n diagonal weights matrix that weights each observation based on its distance from location i, $\hat{\beta}_i$ is a k by 1 vector of coefficients, and y is a k by 1 vector of observations.

The GWR model is mainly configured by the selection of the kernel function, the kernel type (fixed vs adaptive), and the bandwidth. In order to construct $W(i)$ and to calculate β_{i0} the selection of a distance weighting concept is required that is linked to the selection of the kernel function and the kernel type. The kernel function specifies how and to which extent the weights are decreased.

The selection of the bandwidth has the highest impact on the results of GWR as it defines the spatial range of the kernel. The optimal bandwidth is computed using the corrected Akaike information criterion (AICc) that is based on the use of a golden section search optimization routine (Oshan et al., 2019). The corrected AIC outperforms other bandwidths selection methods as it offers a compromise between the goodness-of-fit of the model by minimizing the estimation error of the dependent variable and model complexity as there is a penalty in the criterion for the effective number of parameters in the model (Wheeler & Páez, 2010).

$$AIC_c = 2n \log_e \left(\frac{RSS}{n} \right) + n \log_e (2\pi) + n \left(\frac{n + tr(S)}{n - 2 - tr(S)} \right) \tag{3}$$

Equation (3) defines the corrected AIC where n is the sample size, S is the hat matrix, and RSS is the residual sum of squares.

Multiscale Geographically Weighted Regression

Recent efforts to improve the GWR method have resulted in MGWR allowing the conditional relationships between the dependent and explanatory variables to vary at different spatial scales, while GWR treats these relationships at the same scale.

Fotheringham et al. (2017). Consequently, the varying bandwidths of MGWR express the degree of spatial heterogeneity associated with the relationship of each explanatory variable to the dependent variable (Comber et al., 2020).

$$\gamma_i = \sum_{k=1}^{p} \beta_{bwk} x_{ik} + e_{ik}, \quad i = 1, ..., n \tag{4}$$

In the case of MGWR, Eq. (4) enhances Eq. (3) with the specific bandwidth β_{bwk} used for the calibration of the kth conditional relationship. Compared to GWR, bandwidth selection is performed using Back-fitting algorithms that sequentially calibrate a series of univariate GWR models based on the partial residuals from the previous iteration until the MGWR model converges to a solution, and thus recasting MGWR as a generalized additive model (GAM) as bandwidth selection in MGWR has a higher complexity (Oshan et al., 2019; Yu et al., 2020).

Model Evaluation and Visualisation of Results

In this study, the R2 value, the adjusted R2 value, and the Akaike information criterion (AICc) as defined in equation four are applied to evaluate and compare the model fit. The adjusted R2 value is presented in Eq. (5)

$$R_{adj}^2 = \left[\frac{(1 - R^2)(n - 1)}{n - k - 1} \right] \tag{5}$$

where R2 is determined by $R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$ whereas SS_{res} is defined as the residual sum of squares $SS_{res} = \sum_{j=1}(y_j - f_j)^2$ and SS_{tot} as the total sum of squares $SS_{tot} = \sum_{j=1}(y_j - y)^2$. Compared to R2 and R_{adj}^2, lower values for AICc indicate both better model fit and lower model complexity, as a reasonable trade-off between goodness of fit and degrees of freedom or model complexity is achieved (Fotheringham & Oshan, 2016).

Mapping local GWR results is often considered a challenge as parameter estimates and T-values can take both positive and negative numbers and as it is not necessarily required to map all parameter estimates and attached significance values to generate an effective visualization of the overall quality and the most relevant characteristics of a GWR model (Matthews & Yang, 2012). In order to solve this issue, a bivariate choropleth mapping approach suggested by Mennis (2006); Matthews & Yang (2012) is applied that limits the presentation of results only to those areas of the study area that are statistically significant by masking out areas that have local T-values between -1.96 and +1.96.

Retroactive Prediction of Changes in Transport Supply

The predictive capacities of the journeys per capita ratio using both OLS and GWR to investigate the impact of certain transport policy changes are evaluated by training the

model with the data from a past time frame and then testing the model with data from a subsequent time frame following the concept of supervised learning. Supervised learning entails learning a mapping between a set of input variables X and an output variable y, and applying that mapping to predict outputs for unseen data (Cunningham et al., 2008). The transport policy changes include mainly adjustments in transport supply expressed as the total frequency. This variable takes all operating lines of each small area into account and computes their total frequency. It thus expresses how many daily public transport options the average user has. The journeys per capita ratio is evaluated for seven consecutive time frames (spring 2017 to spring 2020), and the whole transport supply data is provided by Stockholm public transport administration on a yearly base resulting in four different transport supply data sets. Consequently, three prediction cases are assessed that apply the 2017 autumn data to predict the journeys per capita ratio based on the updated transport supply test data in spring 2018. This process is repeated for autumn 2018 to spring 2019 and autumn 2019 to spring 2020. This approach evaluates the capabilities of GWR beyond its explanatory capabilities. This approach thus enables public transport providers (such as the Stockholm Public Transport Administration) to draw conclusions on whether GWR can be used in the future to predict the local impact of changes in transport supply on travel behaviour, expressed as the journeys per capita ratio.

Results

The results based on the methodological procedure described in "Methodology" are presented in this section. The first subsection focuses on the spatio-temporal development of public transport use in the Stockholm region from spring 2017 to spring 2020. The second subsection is dedicated to the results of the (M)GWR framework. This subsection compromises the presentation of the computed parameter estimates for all three methods (OLS, GWR, and MGWR), the presentation and the comparison of the goodness of fit statistics, as well as the retrospective prediction results using historical transport supply data.

Spatio-Temporal Trends

The aggregated journeys per capita ratios for each of the seven time frames on the municipality level within Stockholm County are presented in Fig. 3. The comparison of the change over time, as well as the total change during the whole study period, reveals two trends: Regions that already had strong public transport competitiveness (high journeys per capita ratio) could even enhance their public transport usage, while regions with already low levels of public transport competitiveness even further declined in terms of public transport usage. Consequently, the discrepancy between competitive and non-competitive regions in public transport usage further increased within Stockholm County. The other trend is distinguished by a spatial pattern showing that areas with increased public transport competitiveness are clustered as a north-south corridor towards Uppsala.
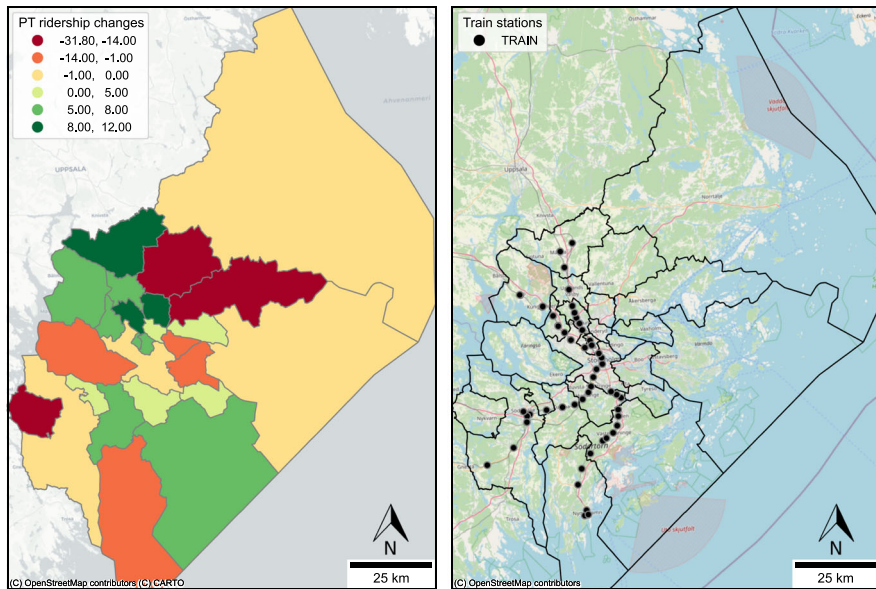
(a) Transit ridership development in % per municipality

(b) Spatial distribution of the train stations

	Spring 17	Autumn 17	Spring 18	Autumn 18	Spring 19	Autumn 19	Spring 20	Change over time	Total change (%)
Österåker	0,22	0,17	0,20	0,17	0,15	0,14	0,15		-31,8%
Vallentuna	0,17	0,16	0,15	0,14	0,13	0,13	0,13		-23,5%
Nykvarn	0,07	0,07	0,06	0,07	0,06	0,07	0,06		-14,3%
Lidingö	0,23	0,22	0,24	0,22	0,23	0,20	0,21		-8,7%
Nynäshamn	0,17	0,17	0,17	0,15	0,17	0,18	0,16		-5,9%
Ekerö	0,23	0,25	0,23	0,25	0,22	0,23	0,22		-4,3%
Nacka	0,36	0,36	0,35	0,35	0,34	0,34	0,35		-2,8%
Stockholm	1,07	1,07	1,05	1,04	1,04	1,05	1,06		-0,9%
Värmdö	0,30	0,31	0,28	0,30	0,29	0,30	0,30		0,0%
Södertälje	0,30	0,30	0,29	0,30	0,29	0,30	0,30		0,0%
Norrtälje	0,13	0,14	0,13	0,14	0,13	0,13	0,13		0,0%
Danderyd	0,45	0,48	0,53	0,47	0,48	0,46	0,46		2,2%
Tyresö	0,35	0,35	0,34	0,35	0,35	0,36	0,36		2,9%
Salem	0,27	0,28	0,26	0,27	0,26	0,28	0,28		3,7%
Vaxholm	0,23	0,24	0,23	0,24	0,22	0,24	0,24		4,3%
Huddinge	0,45	0,41	0,45	0,47	0,45	0,47	0,47		4,4%
Haninge	0,32	0,32	0,31	0,32	0,31	0,33	0,34		6,3%
Upplands Väsby	0,29	0,31	0,29	0,30	0,29	0,28	0,31		6,9%
Järfälla	0,29	0,31	0,29	0,30	0,30	0,31	0,31		6,9%
Solna	0,66	0,71	0,68	0,69	0,69	0,72	0,71		7,6%
Upplands-Bro	0,26	0,28	0,25	0,27	0,25	0,28	0,28		7,7%
Sundbyberg	0,51	0,53	0,51	0,53	0,50	0,53	0,55		7,8%
Botkyrka	0,38	0,42	0,40	0,42	0,39	0,42	0,41		7,9%
Sigtuna	0,33	0,35	0,32	0,35	0,32	0,34	0,36		9,1%
Täby	0,20	0,20	0,19	0,25	0,23	0,22	0,22		10,0%
Sollentuna	0,42	0,46	0,45	0,46	0,44	0,47	0,47		11,9%

(c) Temporal development of the journeys per capita ratio on municipality level

Fig. 3 Spatio-temporal trends of public transport ridership development within Stockholm County. Subfigure a presents a spatial rendering of total changes in public transport ridership between spring 2017 and spring 2020. Subfigure b shows train stations. In Subfigure c changes in public transport ridership at half yearly intervals on the municipality level are presented. Generally, it is noticeable that passenger numbers are rising in municipalities where the train network and the service frequency have been expanded

Figure 3a maps the total change in percent during the study period and thus helps to understand the spatial allocation of the areas marked by either increased or declined public transport competitiveness. During the studied period, the commuter rail was heavily expanded with a new tunnel through the city center, making the two most central stations more accessible to destinations in the city. Furthermore, the frequency of its services was increased. The extended train network is depicted in Fig. 3b. These improvements seem to have had an effect on the journeys per capita in the commuter rail adjacent areas.

(M)GWR

Following the procedure presented in "Feature Selection", a model consisting of four explanatory variables was selected to reflect the journeys per capita ratio as the dependent variable. The comparison of the four selected variables in Table 3 with Table 1 in "Data" reveals that two of the selected explanatory variables (total frequency and car road network) are related to transport supply, while the inverse network distance, which reflects the proximity of a small area to Stockholm's central public transport hub, is a centrality measure, and the variable proportion of elders is a socioeconomic variable. The combination of these four explanatory variables is characterized by the best model fit, expressed by the lowest AIC_c value, without running into multicollinearity issues.

The selected explanatory variables and their estimated coefficients, as well as measures presenting their statistical significance, are depicted in Table 3. All selected variables are characterized by their low p-values verifying their statistical significance. The achieved t-values prove their significance as well. All estimators have their presumed sign yielding that public transport is positively impacted by a distinct transport supply expressed as the total frequency, the proximity to T-Centralen, the largest and most central transport hub within Stockholm County as well as elderly people tend more towards using public transport. In contrast, an expanded car infrastructure tends to decrease public transport usage. Out of all explanatory variables, the total frequency representing the transport supply has the biggest impact on public transport ridership.

The distribution of the local coefficients of the resulting GWR and MGWR models, as well as their bandwidths, are presented in Table 4. In the case of GWR, an adaptive bisquare kernel is applied, and the AICc criterion is applied as a search criterion for the optimal bandwidth. In the case of MGWR, the back-fitting algorithm based on GAM

Table 3 Results of the OLS model

Variable	Full name	Coefficients	Standard error	t-value	p-value
β_0	Intercept	-0.000	0.043	-0.000	1.000
β_1	Total Frequency	0.655	0.067	9.771	0.000
β_2	Car road network	-0.138	0.044	-3.119	0.002
β_3	Inverse network distance	0.161	0.067	2.411	0.016
β_4	Proportion of elders	0.103	0.043	2.381	0.017

Table 4 Comparison of the results between GWR and MGWR, the former with fixed bandwidth (Bw) of 59

Variable	GWR							MGWR						
	Bw	Mean	Std	Min	Median	Max		Bw	Mean	Std	Min	Median	Max
β_0	59	-0.178	0.297	-1.146	-0.166	0.421		63	-0.204	0.231	-0.603	-0.212	0.235
β_1	59	1.022	0.462	0.248	0.976	1.939		16	0.991	0.532	-0.305	0.942	2.414
β_2	59	-0.625	0.727	-2.514	-0.217	0.171		92	-0.552	0.286	-1.030	-0.586	-0.031
β_3	59	-0.104	0.438	-1.454	-0.054	0.736		111	0.180	0.093	-0.138	0.223	0.236
β_4	59	0.082	0.180	-0.458	0.071	0.585		62	0.066	0.167	-0.363	0.055	0.352

fitting methods is applied. The optimum computed bandwidth for the GWR model is 59, meaning that only the closest 59 neighboring small areas are considered regarding the construction of the weight matrix. While GWR is characterized by a constant bandwidth of 59 among the explanatory variables, MGWR has varying bandwidths taking 16 to 111 nearest neighbors into account. The smallest bandwidth of 16 in case of total frequency results (β_1) in more localized parameter estimates as a small bandwidth supports the analysis at a finer spatial scale. The largest bandwidth of 111 in the case of the inverse network distance (β_3) leads to more stationary parameter estimates. These assumptions are also reflected in the parameter estimates that vary from 0.991 (β_1) to 0.180 (β_3).

The statistically significant parameter estimates with t-values below or above 1.96 are illustrated in Fig. 4. The spatial distribution of the estimated local coefficients in Fig. 4a reveals that the total frequency (β_1) has a concentric distribution with high parameter estimates within the center and declining estimates in the outskirts. While all significant parameter estimates of (β_1) are positive, the car road network (β_2) parameter estimates in Fig. 4b are all negative with high values in the center and lower values in the outskirts.

Districts with a dense road network are particularly attractive for car usage leading to lower public transport ridership. The inverse car road network distance (β_3) in Fig. 4c has negative parameter estimates in most of the small areas around the city center. The proximity of a small area to the hub of public transport is often characterized by a high transport supply, as neighboring small areas are often characterized by a higher transport supply. Consequently, the proximity expressed as the negative network distance positively correlates with increased public transport ridership. Some of the small areas have negative parameter estimates violating this assumption. A comparison between Figs. 4c and 5a reveals these small areas are characterized by low local R^2_{local} values indicating the model fit of the local models at these particular small areas is below average. This indicates that the local models in these small areas miss some relevant information. The spatial distribution of the estimated local coefficients in 4d symbolizes the core concept behind local modeling - varying parameter estimates.

In some small areas, the proportion of elders (β_4) contributes to increased public transport ridership while it leads to decreased transport ridership in other small areas. The comparison of Fig. 4a and c reveals that small areas with a distinct transport supply encourage elderly people to use public transport. In contrast, small areas with insufficient transport supply have a discouraging impact on the use of public transport. This effect might be directly correlated with the required walking distance to public transport stops.

Figure 5 illustrates the distribution of the actual journeys per capita ratio and local R2 values. Comparing these two Figures indicates that these R^2_{local} values are caused by two effects - spatial outliers and edge effects. Spatial outliers encompass those small areas with attribute values that significantly differ from their counterparts in the neighboring small areas. The edge effects are caused by more heterogeneous local neighborhoods in terms of distance as the nearest neighbor search along the edge of the study area needs to cover more distance to find the same number of nearest neighbors as a small area within the city center (Farber & Páez, 2007).
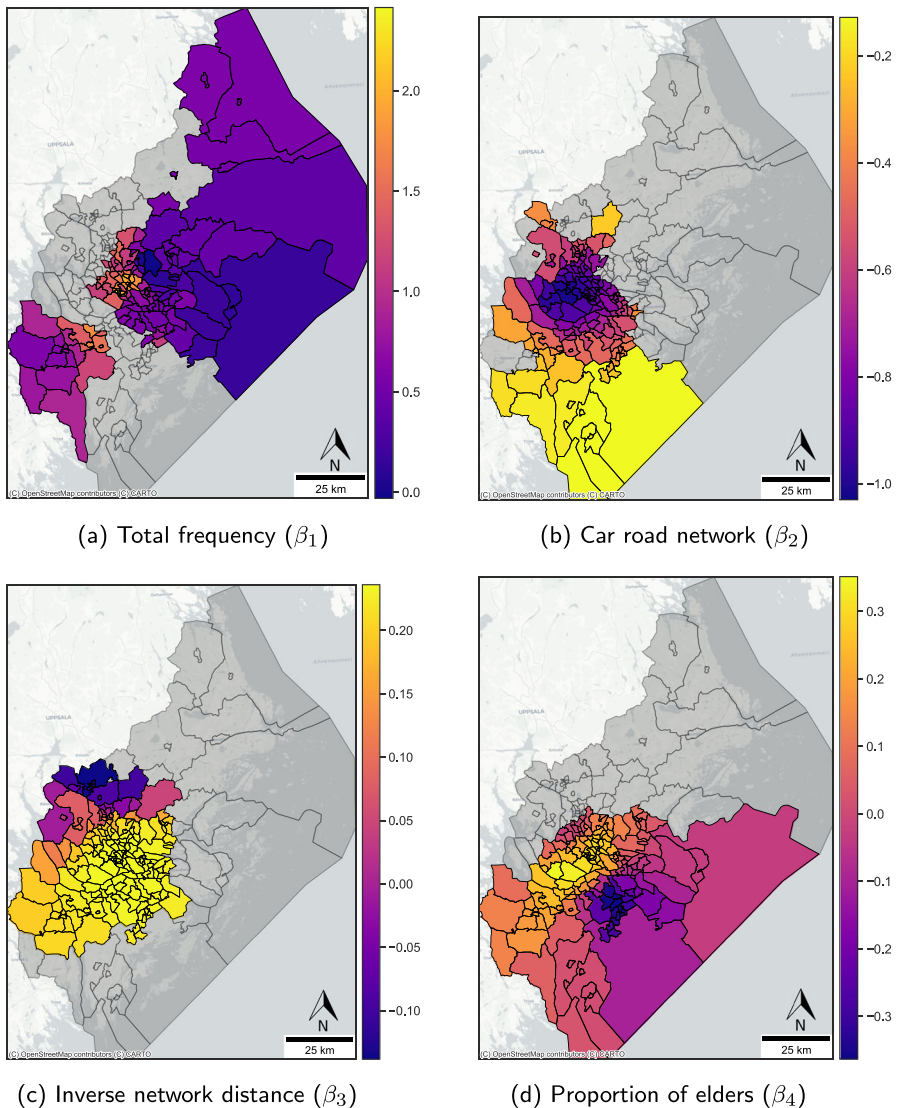
(a) Total frequency (β_1)

(b) Car road network (β_2)

(c) Inverse network distance (β_3)

(d) Proportion of elders (β_4)

Fig. 4 Distribution of the significant parameter estimates with t-values ≤ -1.96 or $\geq +1.96$

Model Fit Comparison

The comparison of the goodness of fit for all three models for each time frame is demonstrated in Table 5. It indicates that the local GWR and MGWR models are superior to the global OLS model in all test cases. For all time frames, the OLS model has the highest AIC_c and SS_{res} value while having the lowest R^2 and R^2_{adj} scores. The comparison between the GWR and the MGWR model reveals that both models have a more comparable goodness of fit. The discrepancy between the two local models is
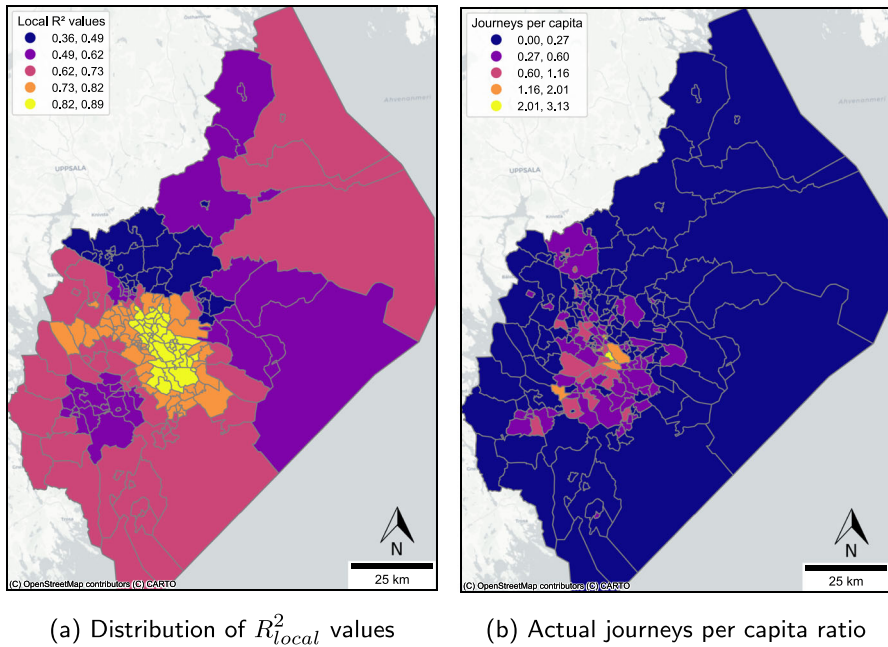
(a) Distribution of R^2_{local} values   (b) Actual journeys per capita ratio

Fig. 5 Distribution of the R^2_{local} values and the actual journeys per capita ratio

lower than between the global and the local models. It is noteworthy that the ranking of the models regarding the goodness of fit varies with the different time frames. In the time frames of autumn 2017 and spring 2018, the GWR model has the highest model fit. In the remaining five time frames, the MGWR model is the superior model in terms of goodness of fit. Either way, these results reveal that a local model is more beneficial to express the journeys per capita ratio as a function of several explanatory variables.

Prediction Results

Following the procedure outlined as step 3 in Fig. 2 and its respective "Retroactive Prediction of Changes in Transport Supply", both the OLS regression and the GWR models are calibrated following the concept of supervised learning with known historical transport supply data as features and the calculated journeys per capita ratio as the label to obtain the optimum parameter values that best reflect the journeys per capita ratio. The models are then tested with *unseen* data from the temporally subsequent time frame with updated transport supply to assess whether they are capable of assessing the local impact of transport supply changes on the journeys per capita ratio. The aim of this procedure is to test whether GWR can be used beyond its exploratory capability by testing its predictive capability and benchmarking it against OLS regression. The performance of the predictive capabilities of the models is tested using the R^2 & R^2_{adj}

Table 5 Goodness of fit results

Time frame	OLS				GWR				MGWR			
	SS_{res}	R^2	R^2_{adj}	AIC_c	SS_{res}	R^2	R^2_{adj}	AIC_c	SS_{res}	R^2	R^2_{adj}	AIC_c
Spring 2017	77.144	0.646	0.639	404.591	36.046	0.835	0.793	340.210	33.491	0.846	0.804	333.980
Autumn 2017	89.816	0.588	0.580	437.748	37.337	0.829	0.781	359.559	39.730	0.818	0.770	365.473
Spring 2018	85.920	0.606	0.598	428.080	38.574	0.773	0.823	367.032	41.609	0.809	0.761	369.685
Autumn 2018	84.209	0.614	0.606	423.696	38.183	0.825	0.780	353.058	29.330	0.865	0.815	349.641
Spring 2019	79.817	0.634	0.627	412.018	41.088	0.812	0.769	355.105	37.036	0.830	0.784	354.433
Autumn 2019	84.656	0.612	0.604	424.849	50.333	0.769	0.730	375.014	41.006	0.812	0.761	376.783
Spring 2020	83.404	0.617	0.610	421.600	50.075	0.770	0.732	373.876	40.188	0.816	0.771	357.809

error metrics as these are less prone to outliers and as R^2_{adj} takes the number of input features into account.

Table 6 illustrates the predicted journeys per capita ratio after training both models with the autumn data. The models are then updated with the consecutive transport supply data of the next year, and the journeys per capita ratio for each small area is computed based on the estimated coefficients derived from the training process using the autumn data and the new input from the temporal subsequent spring test data.

Figure 6a presents the estimated journeys per capita ratio obtained from the GWR model, while Fig. 6b presents the estimated journeys per capita ratio obtained from the OLS model. The comparison of the two Subfigures reveals that the predicted journeys per capita ratios obtained from the OLS model are slightly higher around the city center as well as the eastern outskirts. This observation is reflected in Fig. 6d by the higher derivation of the actual and the predicted journeys per capita ratios compared to Fig. 6c. The prediction error expressed as the discrepancy between the actual and the predicted value is more distinct in the case of OLS. This phenomenon is portrayed in Fig. 6d as the deviation there is more frequent and distinct through higher values. While the deviation between the actual and the predicted journeys per capita ratio in the case of GWR is mainly located within the city center, the OLS deviation is distributed over the whole study area.

The superiority of the local model in terms of prediction power is also confirmed in Table 6 comparing the prediction accuracy between the OLS and the GWR model. In all three prediction cases, the GWR model is characterized by a higher prediction accuracy. Unfortunately, the prediction capacity of MGWR could not be computed due to restrictions in the applied Python library, but the results in "Model Fit Comparison" imply that the prediction results of MGWR might be slightly superior compared to GWR.

Discussion

The implemented (M)GWR model, in conjunction with the journey inference from smart card ticketing and the selected explanatory variables, enables the evaluation of the spatio-temporal impact of a wide array of policy changes in public transport. Whether it be changes in transport supply, including the creation of new public

Table 6 Comparison of the prediction accuracy between OLS and GWR

Time frame		OLS		GWR	
Training data	Testing data	R^2	R^2_{adj}	R^2	R^2_{adj}
Autumn 2017	Spring 2018	0.601	0.596	0.829	0.781
Autumn 2018	Spring 2019	0.632	0.626	0.825	0.780
Autumn 2019	Spring 2020	0.617	0.612	0.769	0.730

(a) GWR predicted journeys per capita

(b) OLS predicted journeys per capita

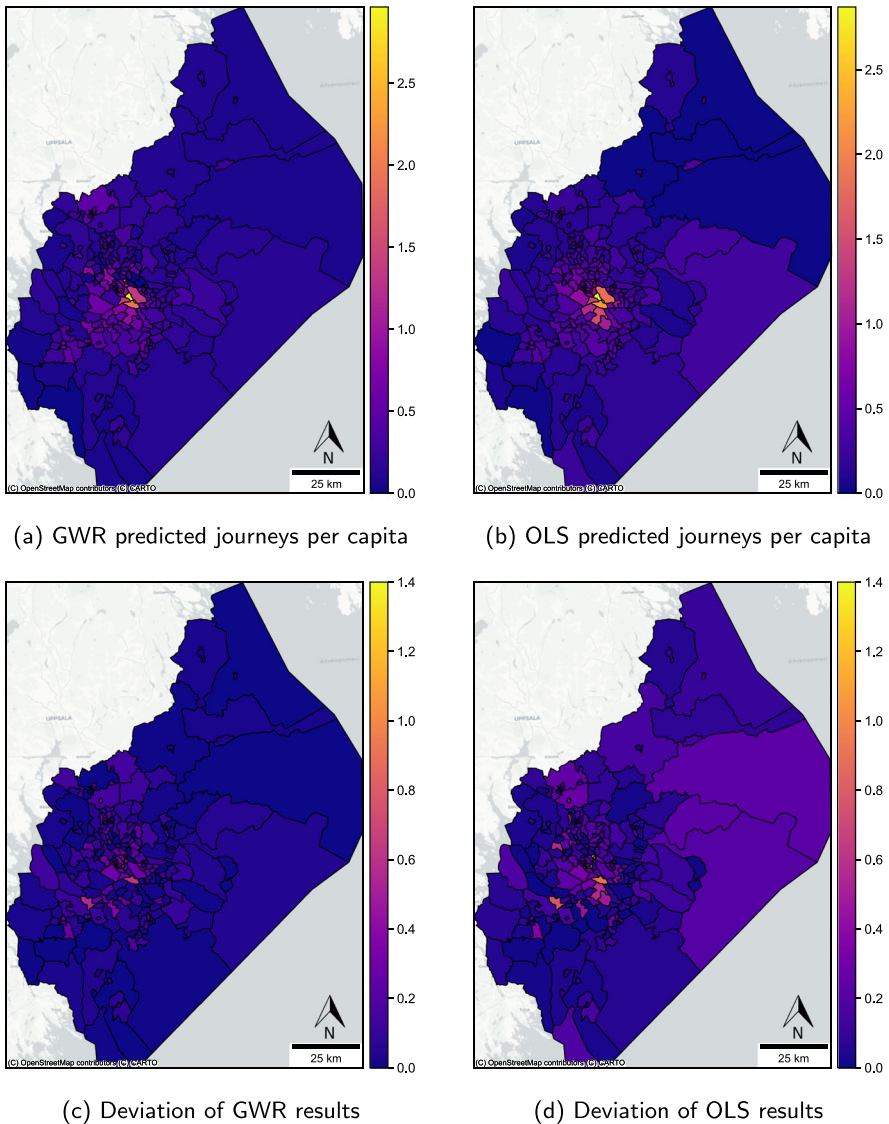(c) Deviation of GWR results

(d) Deviation of OLS results

Fig. 6 Comparison of the GWR and OLS prediction results

transport stations, fleet expansion through the purchase of new buses or trains, or changes in operational strategies. Key applications include:

- Benchmarking different areas on how journeys per capita develop given input levels of public transport supply, population concentration, and road network quality.
- Intentional goal-oriented work with set targets, planned activities for the meeting of said targets, and rigorous follow-up assessment of the activities' efficiency.

- Historical understanding of demand trends, investigating how the competitiveness of public transport has developed over time both for the whole Region and for smaller areas.

Further, it allows setting realistic and achievable goals for increasing competitiveness to expand the efficiency of the public transport network and to promote sustainability. Better goals for competitiveness and an efficient forecast on journeys per capita might lead to an increase in public transport share and therefore decreased car dependence. Lower levels of car use are related to lower levels of environmentally harmful carbon emissions. The computation of the journeys per capita ratio and its comparison with transport supply may also help to identify areas with an oversaturation of public transport supply. These areas can then be optimized, and the total amount of traffic could be even further reduced. Besides promoting environmental sustainability, assessing public transport supply and ridership by area and comparing it with the socioeconomic data by area may also help to further promote social sustainability by ensuring that all societal groups obtain a sufficient level of service.

The selected features within the feature selection process described in "Feature Selection" correspond to a large extent to the expectations of the literature research outcomes as outlined in "Factors Affecting Public Transport Usage". The total frequency per area reflects the transport supply as well as the accessibility and density of the stops, the car road network corresponds to car accessibility, while the inverse network distance reflects the centrality of the respective small area to the main transport hub. Nevertheless, we would have expected a slightly stronger inclusion of the socioeconomic variables, which did not exist except for age (with low parameter values).

Generally, the predictive power of the model and the validity of the findings are highly dependent on the correctness of the inferred journeys. If the users mix public transport journeys with other motorized journeys, the estimated start and end locations of the journeys might be incorrect. However, these are known theoretical problems with known state-of-the-art remedies. The journeys per capita ratio is calculated as the ratio between starting journeys on the stop level and the population of the area based on a desired spatial scale. The performed methodology is based on a spatial join of the starting journeys at the public transport stops within the study area. It has to be kept in mind that aggregation results of point-based measures to spatial units are always a source of statistical bias referred to as the modifiable areal unit problem (MAUP) (Longley et al., 2005).

Another challenge of this study is the uneven distribution of the area and the population within the planområden. In particular, small areas with low populations but good public transport accessibility are prone to errors in the calculation of the journeys per capita ratio as their public transport infrastructure also supplies neighboring areas. In order to solve that problem, a buffer-based approach is applied to take the neighborhood into account. The algorithm considers the population of the neighborhood if the buffer intersects with the neighboring areas. In case the neighborhood is considered as the buffer touches neighboring area, the population is computed based on the proportional overlap. One future contribution to improving this method would

be a network distance-based buffer. That buffer would catch reality better than the Euclidean distance buffer applied in this study.

Another beneficial application of the network distance would be a network distance-based computation of the nearest neighbor based on the calculated bandwidth. A network distance based nearest neighbor approach could take geography better into account and would most likely simulate spatial patterns better.

Compared to most previous studies, the prediction power of GWR is tested as well by evaluating the impact of past transport supply changes. The results of this study reveal that GWR is a superior method to OLS both in terms of goodness of fit and prediction power. The improved fit of MGWR compared to GWR raises expectations that the predictive capabilities using MGWR could be even higher, leading to future research needs.

Conclusion

The performed analyses reveal that (M)GWR, in conjunction with the derived journeys per capita ratio, is a suitable method to track back the impact of transport supply policies on a local scale. More specifically, from a transport science perspective, the performed analyses demonstrate the potential of inferred journeys to measure travel behavior at the city district level and to track local changes over time, allowing conclusions to be drawn about the success of transport policies. Furthermore, this study highlights the potential of training GWR models according to the idea of supervised learning with historical transport supply data with known changes in travel behavior to obtain a trained model to predict the impact of future transport supply changes at a local scale.

From a spatial analysis point of view, the application spectrum of GWR is enhanced beyond model fit analyses as the prediction power of GWR is retrospectively analyzed as well based on seven consecutive data frames. Compared to most recent GWR-related studies focusing on station-level ridership modeling, this study models public transport ridership on a city district scale.

References

Almlöf, E., Rubensson, I., Cebecauer, M., Jenelius, E. (2020). Who is still travelling by public transport during covid-19? socioeconomic factors explaining travel behaviour in stockholm based on smart card data. *Socioeconomic Factors Explaining Travel Behaviour in Stockholm Based on Smart Card Data* (September 8, 2020)

An, R., Wu, Z., Tong, Z., Qin, S., Zhu, Y., & Liu, Y. (2022). How the built environment promotes public transportation in wuhan: A multiscale geographically weighted regression analysis. *Travel Behaviour and Society, 29*, 186–199.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., & White, P. (2004). The demand for public transport: a practical guide. *Transportation Research Laboratory Report (TRL593)*

Blainey, S. (2010). Trip end models of local rail demand in england and wales. *Journal of Transport Geography, 18*(1), 153–165.

Blainey, S., & Mulley, C. (2013). Using geographically weighted regression to forecast rail demand in the sydney region. *Australasian transport research forum, brisbane*

Blainey, S. P., & Preston, J. M. (2013). Extending geographically weighted regression from points to flows: a rail-based case study. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 227*(6), 724–734.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis, 28*(4), 281–298.

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician), 47*(3), 431–443.

Buehler, R., & Pucher, J. (2012). Demand for public transport in germany and the USA: an analysis of rider characteristics. *Transport Reviews, 32*(5), 541–567.

Cao, X., Shi, Y., Zhou, L., Tao, T., & Yang, Q. (2021). Analysis of factors influencing the urban carrying capacity of the shanghai metropolis based on a multiscale geographically weighted regression (mgwr) model. *Land, 10*(6), 578.

Cardozo, O. D., García-Palomares, J. C., & Gutiérrez, J. (2012). Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Applied Geography, 34*, 548–558.

Cats, O., & Ferranti, F. (2022). Unravelling individual mobility temporal patterns using longitudinal smart card data. *Research in Transportation Business & Management*, 100816

Charlton, M., Fotheringham, A. S., & Brunsdon, C. (2009). *Geographically weighted regression white paper* (pp. 1–14). Kildare: National University of Ireland Maynooth.

Chen, R., & Zhou, J. (2022). Fare adjustment's impacts on travel patterns and farebox revenue: An empirical study based on longitudinal smartcard data. *Transportation Research Part A: Policy and Practice, 164*, 111–133.

Chiou, Y.-C., Jou, R.-C., & Yang, C.-H. (2015). Factors affecting public transportation usage rate: Geographically weighted regression. *Transportation Research Part A: Policy and Practice, 78*, 161–177.

Chow, L.-F., Zhao, F., Liu, X., Li, M.-T., & Ubaka, I. (2006). Transit ridership model based on geographically weighted regression. *Transportation Research Record, 1972*(1), 105–114.

Clark, S. D. (2007). Estimating local car ownership models. *Journal of Transport Geography, 15*(3), 184–197.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., . . . others (2020). The gwr route map: a guide to the informed application of geographically weighted regression. arXiv preprint arXiv:2004.06070

Cunningham, P., Cord, M., Delany, S.J. (2008). Supervised learning. *Machine learning techniques for multimedia: case studies on organization and retrieval* , 21–49

De Gruyter, C., Gunn, L., Kroen, A., Saghapour, T., Davern, M., & Higgs, C. (2022). Exploring changes in the frequency of public transport use among residents who move to outer suburban greenfield estates. *Case Studies on Transport Policy, 10*(1), 341–353.

De Vos, J., Waygood, E. O. D., & Letarte, L. (2020). Modeling the desire for using public transport. *Travel Behaviour and Society, 19*, 90–98.

Du, H., & Mulley, C. (2006). Relationship between transport accessibility and land value: Local model approach with geographically weighted regression. *Transportation Research Record, 1977*(1), 197–205.

Du, H., & Mulley, C. (2012). Understanding spatial variations in the impact of accessibility on land value using geographically weighted regression. *Journal of Transport and Land Use, 5*(2), 46–59.

Dziauddin, M. F., Powe, N., & Alvanides, S. (2015). Estimating the effects of light rail transit (lrt) system on residential property values using geographically weighted regression (gwr). *Applied Spatial Analysis and Policy, 8*(1), 1–25.

Farber, S., & Páez, A. (2007). A systematic investigation of cross-validation in gwr model estimation: empirical analysis and monte carlo simulations. *Journal of Geographical Systems, 9*(4), 371–396.

Fotheringham, A. S., & Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems, 18*(4), 303–329.

Fotheringham, A. S., & Rogerson, P. A. (2008). *The sage handbook of spatial analysis*. Sage.

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (mgwr). *Annals of the American Association of Geographers, 107*(6), 1247–1265.

Gao, F., Yang, L., Han, C., Tang, J., & Li, Z. (2022). A network-distance-based geographically weighted regression model to examine spatiotemporal effects of station-level built environments on metro ridership. *Journal of Transport Geography, 105*, 103472.

Gascon, M., Marquet, O., Grácia-Lavedan, E., Ambrós, A., Götschi, T., de Nazelle, A., et al. (2020). What explains public transport use? evidence from seven european cities. *Transport Policy, 99*, 362–374.

Giannakis, E., Serghides, D., Dimitriou, S., & Zittis, G. (2020). Land transport CO2 emissions and climate change: evidence from cyprus. *International Journal of Sustainable Energy, 39*(7), 634–647.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*(Mar), 1157–1182

Hadayeghi, A., Shalaby, A. S., & Persaud, B. N. (2010). Development of planning level transportation safety tools using geographically weighted poisson regression. *Accident Analysis & Prevention, 42*(2), 676–688.

He, Y., Zhao, Y., & Tsui, K.-L. (2019). Geographically modeling and understanding factors influencing transit ridership: an empirical study of shenzhen metro. *Applied Sciences, 9*(20), 4217.

Holmgren, J. (2013). An analysis of the determinants of local public transport demand focusing the effects of income changes. *European Transport Research Review, 5*(2), 101–107.

Kholodov, Y., Jenelius, E., Cats, O., van Oort, N., Mouter, N., Cebecauer, M., & Vermeulen, A. (2021). Public transport fare elasticities from smartcard data: Evidence from a natural experiment. *Transport Policy, 105*, 35–43.

Li, D., Zang, H., Yu, D., He, Q., & Huang, X. (2023). Study on the influence mechanism and space distribution characteristics of rail transit station area accessibility based on mgwr. *International Journal of Environmental Research and Public Health, 20*(2), 1535.

Litman, T. (2022). Evaluating public transit benefits and costs: Best practices guidebook

Litman, T., & Burwell, D. (2006). Issues in sustainable transportation. *International Journal of Global Environmental Issues, 6*(4), 331–347.

Liu, J., Chau, K., & Bao, Z. (2023). Multiscale spatial analysis of metro usage and its determinants for sustainable urban development in shenzhen, china. *Tunnelling and Underground Space Technology, 133*, 104912.

Liu, L., Yu, H., Zhao, J., Wu, H., Peng, Z., & Wang, R. (2022). Multiscale effects of multimodal public facilities accessibility on housing prices based on mgwr: a case study of wuhan, china. *ISPRS International Journal of Geo-Information, 11*(1), 57.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographic information systems and science*. John Wiley & Sons.

Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2014). Geographically weighted regression with a non-euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science, 28*(4), 660–681.

Lunke, E.B., Fearnley, N., Aarhaug, J. (2021). Public transport competitiveness vs. the car: impact of relative journey time and service attributes. *Research in Transportation Economics, 90*, 101098

Lyu, C., Liu, Y., Liu, Z., Wu, X., & Yang, X. (2020). Exploring multi-scale spatial relationship between built environment and public bicycle ridership. *Journal of Transport and Land Use, 13*(1), 447–467.

Ma, X., Zhang, J., Ding, C., & Wang, Y. (2018). A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership. *Computers, Environment and Urban Systems, 70*, 113–124.

Marques, S.d.F., & Pitombo, C.S. (2022). Transit ridership modeling at the bus stop level: Comparison of approaches focusing on count and spatially dependent data. *Applied Spatial Analysis and Policy* , 1–37

Matthews, S. A., & Yang, T.-C. (2012). Mapping the results of local statistics: Using geographically weighted regression. *Demographic Research, 26*, 151.

May, A. D., Shepherd, S. P., & Timms, P. M. (2000). Optimal transport strategies for european cities. *Transportation, 27*(3), 285–315.

Mennis, J. (2006). Mapping the results of geographically weighted regression. *The Cartographic Journal, 43*(2), 171–179.

Miller, P., de Barros, A. G., Kattan, L., & Wirasinghe, S. (2016). Public transportation and sustainability: A review. *KSCE Journal of Civil Engineering, 20*(3), 1076–1083.

Nguyen-Phuoc, D. Q., Currie, G., De Gruyter, C., & Young, W. (2018). How do public transport users adjust their travel behaviour if public transport ceases? a qualitative study. *Transportation Research Part F: Traffic Psychology and Behaviour, 54*, 1–14.

North, M.A. (2009). A method for implementing a statistically significant number of data classes in the Jenks algorithm. *2009 sixth international conference on fuzzy systems and knowledge discovery* (vol. 1, pp. 35–38)

Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information, 8*(6), 269.

Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., & White, P. (2006). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy, 13*(4), 295–306.

Pelletier, M.-P., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies, 19*(4), 557–568.

Pirdavani, A., Bellemans, T., Brijs, T., & Wets, G. (2014). Application of geographically weighted regression technique in spatial analysis of fatal and injury crashes. *Journal of Transportation Engineering, 140*(8), 04014032.

Schiller, P. L., & Kenworthy, J. R. (2017). *An introduction to sustainable transportation: Policy, planning and implementation*. Routledge.

Tang, J., Gao, F., Han, C., Cen, X., & Li, Z. (2021). Uncovering the spatially heterogeneous effects of shared mobility on public transit and taxi. *Journal of Transport Geography, 95*, 103134.

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography, 46*(sup1), 234–240.

Tsai, C.-H., Mulley, C., & Clifton, G. (2012). The spatial interactions between public transport demand and land use characteristics in the sydney greater metropolitan area. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice, 21*(4), 62–73.

Vuchic, V. R. (2008). Transport systems and policies for sustainable cities. *Thermal Science, 12*(4), 7.

Wheeler, D.C., & Páez, A. (2010). Geographically weighted regression. *Handbook of applied spatial analysis* (pp. 461–486). Springer

Yang, H., Xu, T., Chen, D., Yang, H., & Pu, L. (2020). Direct modeling of subway ridership at the station level: a study based on mixed geographically weighted regression. *Canadian Journal of Civil Engineering, 47*(5), 534–545.

Yap, M., Cats, O., van Oort, N., & Hoogendoorn, S. (2017). A robust transfer inference algorithm for public transport journeys during disruptions. *Transportation Research Procedia, 27*, 1042–1049.

Yeboah, G., Cottrill, C. D., Nelson, J. D., Corsar, D., Markovic, M., & Edwards, P. (2019). Understanding factors influencing public transport passengers' pretravel information-seeking behaviour. *Public Transport, 11*(1), 135–158.

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., & Wolf, L. J. (2020). Inference in multiscale geographically weighted regression. *Geographical Analysis, 52*(1), 87–106.

Zhao, F., & Park, N. (2004). Using geographically weighted regression models to estimate annual average daily traffic. *Transportation Research Record, 1879*(1), 99–107.

Zhu, P., Li, J., Wang, K., Huang, J. (2023). Exploring spatial heterogeneity in the impact of built environment on taxi ridership using multiscale geographically weighted regression. *Transportation*, 1–35

Authors and Affiliations

Robert Klar[1,2] 🆔 · **Isak Rubensson**[3,4] 🆔

Isak Rubensson
isak.jarlebring-rubensson@regionstockholm.se

[1] Department of Science and Technology, Linköping University, Luntgatan 2, Norrköping 602 47, Östergotland, Sweden

[2] Traffic Analysis and Logistics, Swedish National Road and Transport Research Institute (VTI), Olaus Magnus väg 35, Linköping 583 30, Östergotland, Sweden

[3] Transport and Systems Analysis, Department of Urban Planning and Environment, KTH Royal Institute of Technology, Teknikringen 10 A, Stockholm 100 44, Sweden

[4] Stockholm Public Transport Administration, Region Stockholm, Lindhagensgatan 100, Stockholm 105 73, Sweden