What cities have is how people travel
Conceptualizing a data-mining-driven modal split framework

Lee, Sujin; Lee, Jinwoo; Mastrigt, Suzanne Hiemstra van; Kim, Euiyoung

DOI
10.1016/j.cities.2022.103902

Publication date
2022

Document Version
Final published version

Published in
Cities

Citation (APA)
Lee, S., Lee, J., Mastrigt, S. H. V., & Kim, E. (2022). What cities have is how people travel: Conceptualizing a data-mining-driven modal split framework. *Cities*, *131*, Article 103902. https://doi.org/10.1016/j.cities.2022.103902

Cities 131 (2022) 103902

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

What cities have is how people travel: Conceptualizing a data-mining-driven modal split framework

Sujin Lee [a], Jinwoo Lee [a], Suzanne Hiemstra-van Mastrigt [b], Euiyoung Kim [b,*]

[a] Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Munji-ro 193, Yuseong-gu, Daejeon 34051, Republic of Korea
[b] Faculty of Industrial Design Engineering, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, the Netherlands

A R T I C L E I N F O

Keywords:
Modal split
Mode choice
Sustainable mobility
Data mining
Data-driven decision making

A B S T R A C T

As city-level modal splits are outcomes of city functions, it is essential to understand whether and how city attributes affect modal splits to derive a modal shift toward low-emission travel modes and sustainable mobility in cities. This study elucidates this relationship between modal splits and city attributes in 46 cities worldwide, proposing a two-step data mining framework. First, using the K-Means method, we classify cities into private-vehicle-, public-transit-, and bicycle-dominant groups based on their modal splits. Second, we categorize city attributes into environmental, socio-demographic, and transportation planning factors and quantify their interlocked impacts on cities' modal splits via the decision tree method. We observe that the socio-demographic factor has the highest impact on determining the cities' modal splits. In addition, high population density and employment rate are positively associated with low-emission travel modes. High gasoline tax and low public transit and taxi fares often make people reconsider possessing private vehicles. On the other hand, extreme weather conditions (e.g., hot temperatures) can prevent bicycle usage. Our contribution expands the impact of introduced city planning and policies for modal shifts toward a real-world paradigm and we present implications of the proposed framework in developing practical modal shift strategies.

1. Introduction

For decades worldwide, cities have pursued sustainable mobility, an idea that is intertwined with concerns of climate change, rapid urbanization, and road traffic safety (European Commission, 2013; United Nations, 1997; United Nations, 2015). With its awakening to the adverse effects of cities' relying primarily on motorized modes of transport, the transportation sector has proposed mobility plans and urban planning policies that focus on sustainability (May, 2015). To deemphasize automobile dependency, government authorities and related organizations have introduced policies such as congestion charges and high fuel taxes (C40 Cities, 2019; European Union, 2019). At the same time, those authorities have improved service levels of public transit, cycling, and walking, which has led to a modal shift from conventional private vehicles to low-emission modes (BitiBi.eu, 2017; Buehler et al., 2017).

The modal split – the share of daily trips made by each travel mode – is a valuable indicator to represent city functions; it involves multiple variables and their interaction in the city (Pucher, 1988; Vanoutrive, 2015). As a preliminary step toward the modal shift, it is vital to understand the characteristics and determinants of modal split and mode choice. Therefore, existing studies have investigated modal split and mode preference determinants based on household travel surveys covering multiple modes and the usage of specific travel modes. Researchers have commonly found that higher population density and mixed land use are associated with low dependency on automobiles (Buehler, 2011; Vanoutrive, 2015).

Public transit use tends to rise in dense areas where frequent services (i.e., short time headways) and high station and line densities are justified by concentrated demand in the context of economies of scale, i.e., the Mohring effect (Mohring, 1972; Yang & Zhou, 2020). In other words, demand is highly related to waiting time and proximity to facilities (Ha et al., 2020; Paulley et al., 2006). Like public transit, demand for bicycles (both shared and owned) positively correlates with the associated infrastructure, such as bike paths and docks close to the public transit stations (Eren & Uz, 2020). On the other hand, public transit and bicycles involve more weather dependency than private vehicles (Böcker et al., 2013; Eren & Uz, 2020; Liu et al., 2015).

In addition, in terms of travel behavior, high trip frequencies and

* Corresponding author.
E-mail addresses: su-jin.lee@kaist.ac.kr (S. Lee), lee.jinwoo@kaist.ac.kr (J. Lee), S.Hiemstra-vanMastrigt@tudelft.nl (S.H.-v. Mastrigt), E.Y.Kim@tudelft.nl (E. Kim).

https://doi.org/10.1016/j.cities.2022.103902
Received 9 September 2021; Received in revised form 24 February 2022; Accepted 2 August 2022
Available online 21 August 2022

long travel distances can lead to more private vehicle use (Scheiner, 2010). From a socio-economic perspective, compared to low- and mid-income groups, high-income groups use private vehicles more than other transportation modes (Santos et al., 2013; Sun et al., 2017).

In evaluating the impact of variables on travel mode demand, existing studies have dominantly used statistical models, such as regression and logit methods (Buehler, 2011; Ding et al., 2017; Ha et al., 2020; Liu et al., 2015; Ma et al., 2020; Paulley et al., 2006; Santos et al., 2013; Scheiner, 2010; Sun et al., 2017). Although these methods attempt to explicitly frame a simple relationship between variables and mode preference, they have difficulty reflecting the complexity in mode preferences, which are determined by multiple interlocking variables. In addition, only a few modal split studies (Buehler et al., 2017; Klinger et al., 2013; Liu et al., 2015; McIntosh et al., 2014; Vanoutrive, 2015) have been conducted in a small number of regions or cities within the same continent, so it is difficult to expand and apply results to general problems globally. Therefore, it is necessary to define city archetypes of modal split and to simplify relationships between these archetypes and variables for cities worldwide.

This paper aims to elucidate the impact of city attributes on modal splits of primary and comprehensive intra-city trip modes – private vehicles, public transit, and bicycles – by exploring global city cases. To this end, we propose a two-step data mining framework as a categorical approach to understanding city modal splits. We diagnose city archetypes for trip mode, first by grouping 46 cities based on their similar modal splits. To derive practical suggestions, we seek to define the concept of controllability and employ 17 explanatory variables categorized based on that concept into three factors—environmental, socio-demographic, and planning—that can help decision-makers focus on actionable schemes within relatively unchangeable conditions. In the next step, we investigate the impact of the variables on the defined city modal types. Ultimately, by providing practical suggestions from the empirical evidence, we support decision-makers and city planners in achieving their target modal splits (e.g., low share of trips by private vehicles and high share of public transit or bicycles). We address these objectives by answering the following research questions:

(i) which variables dominate in determining the current modal split of a city;
(ii) how do these variables interact with other variables and modal splits; and.
(iii) which efforts can be potentially made to achieve a city's desired modal split, different from the current.

The remainder of this paper is organized as follows. We describe a study sample of 46 cities worldwide and enumerate possible explanatory variables categorized based on their controllability by decision-makers. Next, we present an analytic framework consisting of K-Means and decision tree methods to investigate the relationship between the explanatory variables and the city-level modal split. Subsequently, we report analysis results for the selected cities and discuss findings related to key questions raised, and address the need for further research. Finally, some concluding remarks are presented.

2. Study site

Based on data availability and credibility, and to reduce regional bias across countries and continents, we selected 46 cities in 35 countries on four continents – America, Asia, Europe, and Oceania. Therefore, the number of cities selected may be uneven for some countries or continents. The study sites are listed by continent and country (cities), as follows:

- America: Brazil (Rio de Janeiro, São Paulo), Canada (Montreal, Toronto), Colombia (Bogota), United States (Chicago, Los Angeles, Miami, New York City)
- Asia: China (Beijing, Shanghai), India (Delhi, Mumbai), Indonesia (Jakarta), Israel (Tel Aviv), Japan (Tokyo), Philippines (Manila), South Korea (Seoul), Taiwan (Taipei)
- Europe: Austria (Vienna), Belgium (Brussels), Czech Republic (Prague), Denmark (Copenhagen), Estonia (Tallinn), Finland (Helsinki), France (Paris), Germany (Berlin, Frankfurt, Munich), Greece (Athens), Hungary (Budapest), Italy (Milan, Rome), Lithuania (Vilnius), Netherlands (Amsterdam), Norway (Oslo), Poland (Warsaw), Republic of Ireland (Dublin), Slovakia (Bratislava), Spain (Barcelona, Madrid), Sweden (Stockholm), Switzerland (Zurich), United Kingdom (London)
- Oceania: Australia (Sydney), New Zealand (Auckland).

3. Variables and data collection

We first define and use the modal splits as the share of daily trips by private vehicle, public transit, and bicycles, which are primary intra-city travel modes. We use 17 variables that describe the city characteristics and affect the use of modal splits. Then, to derive practical suggestions for decision-makers involving target modal split, we define the controllability of variables as to whether changes can be made within feasible periods (5 to 10 years) by interventions (e.g., government plans and policies). As shown in Fig. 1, based on controllability, we categorize the variables into three factors: environmental, socio-demographic, and planning.

Uncontrollable variables are those that are hard to change by human forces because they involve existing environmental conditions, such as climate and natural features (Lee et al., 2017); this is the environmental factor. The planning factor is close to the controllable side; it covers variables that change with city planning and policies, contributing to a city's transformation. The socio-demographic factor is located in the middle of the spectrum because it takes a long time to change via higher-level government plans.

Table 1 briefly presents the dependent variables that express the city-level modal split and the explanatory variables categorized by factor. In the following subsections, the factors and variables are described in detail.

3.1. Environmental factor

The environmental factor is the most uncontrollable one. Even if this factor does not have much room for change, it can represent constraints or opportunities through interaction with other factors. In addition, the impact of the environmental factor may appear as an adaptation to nature, which can influence travel mode choice. This study uses core land area, population density, average temperature, and annual rainfall as environmental variables to capture the impact of this most uncontrollable factor in the modal split.

The city size enables us to infer the scale of travel distances and trip frequencies (Schwanen, 2002; Vanoutrive, 2015). Long travel distances and high trip frequencies are associated with private vehicle use (Haustein & Nielsen, 2016; Jiang et al., 2021; Kenworthy & Laube, 1999). It can be inferred that city size implicitly affects mode choice. Therefore, we use the core land area to reflect the impact of the city size and spatial mobility scale on the city-level modal split. The core land area represents the functional urban area composed of livable areas in the city and contiguous regions with similar population density, defined and calculated by the OECD (2012, 2021); this core land area may be larger or smaller than the nominal administrative boundaries of the city.

Higher population density leads to a compact built environment and dense transportation facilities (Buehler, 2011; Giuliano & Narayan, 2003; McIntosh et al., 2014; Vanoutrive, 2015). Like small city size, high population density can lower dependency on the car by enhancing proximity to workplaces and communities. In previous studies, population density has been found to be an intermediate factor between environmental and socio-demographic factors because it is affected by

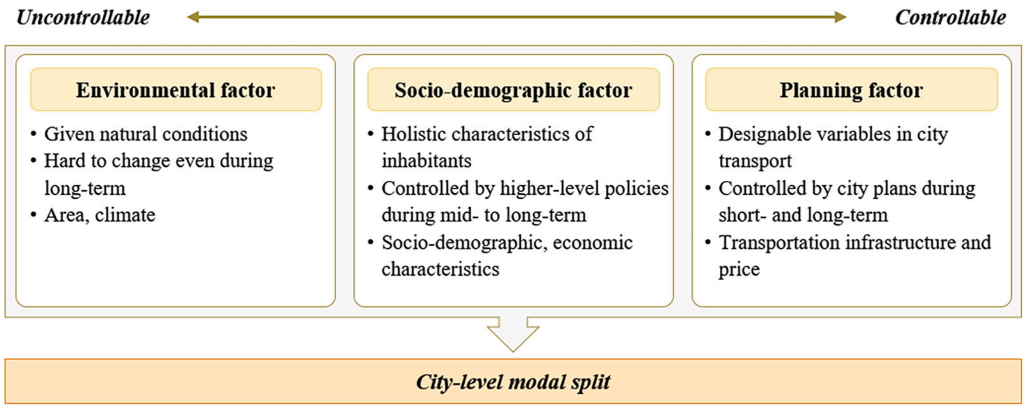


Fig. 1. Categorizing variables influencing city-level modal split based on their controllability: environmental, socio-demographic, and planning factors.

Table 1
List of employed variables.

Category	Variables	Unit	Description
Modal split	Share of private vehicles	%	Share of private vehicles
	Share of public transit	%	Share of public transit
	Share of bicycles	%	Share of bicycles
Environmental factor	Core land area	km^2	Functional area of city
	Population density	Population/ km^2	Population density in metropolitan area
	Average temperature	°C	Annual average temperature
	Average rainfall	mm	Annual average rainfall
Socio-demographic factor	Elderly rate	%	Nation-wide value, share of population aged 65 years or over
	Young rate	%	Nation-wide value, share of population aged under 15 years
	Employment rate	%	Nation-wide value, ratio of the employed to working-age population
	Hourly earning	USD	Average hourly earnings
Planning factor	Density of metro stations	Station/km^2	Number of metro stations per city size
	Density of metro network	km/ population	Length of metro lines per population
	Density of bike-sharing stations	Station/km^2	Number of bike-sharing stations per city size
	Density of bike lanes	km/ population	Length of bike lanes per population
	Number of bike-sharing programs	–	Number of bike-sharing programs operated in city
	Density of road network	km/ population	Length of road network per population
	Tax rate on gasoline	%	Nation-wide value, rate of the total gasoline tax on total price
	Public transit fare	USD	Public transit fare compared to price level[a]
	Taxi fare	USD	Taxi fare compared to price level[a]

[a] Note: Price level is relative to New York City (New York = 100).

the city's environment and social interaction (Hawley, 1972). However, we consider it an environmental variable because it is challenging to intentionally intervene in factors of both area and population in the short- to mid-terms. Similar to using the core land area instead of the nominal city size, we consider urban population density, defined as the ratio of total population to surface area in the city (OECD, 2021), to eliminate bias due to significantly low density or the presence of uninhabitable areas, included in the denominator.

Adverse weather, such as rain, snow, and extreme temperature, is one of the most important reasons for people not to travel or use specific travel modes. Regarding exposure time to external environmental conditions by mode, we can readily accept that weather can disturb cyclists and pedestrians more than public transit passengers and private vehicle users (Böcker et al., 2013; Liu et al., 2015). As such, we use average temperature and annual rainfall to focus on long-term effects of city-specific climate on modal split.

3.2. Socio-demographic factor

The socio-demographic factor usually cannot be controlled by transportation planning but only by higher-level policies in the mid- to long-terms. This factor is crucial because it reveals holistic characteristics of people living in the city from various aspects such as age, occupation, and income (Antipova et al., 2011); these characteristics shape general travel behavior. Therefore, we include elderly rate, young rate, employment rate, and hourly earning as socio-demographic variables to explain population types.

The travel behavior and mode choice characteristics are markedly different by age group. Children have shorter travel boundaries and limitations to driving ability, so their travel modes are usually affected by those of their families, who prefer safe and robust travel modes for their kids (McCarthy et al., 2017). The middle age group has more ability to afford car ownership and frequently travels more for social activities than do other age groups, leading them to prefer car ownership (Ding et al., 2017). Seniors have relatively short travel boundaries and lower trip frequency due to physical restrictions or fewer activities than people of more active ages (Alsnih & Hensher, 2003; Moniruzzaman et al., 2013; Olawole & Aloba, 2014; Rosenbloom, 2001). Seniors' mode choices are also affected by family situation, income level, physical capability, and neighborhood environment (Alsnih & Hensher, 2003; Cheng et al., 2019; Kim & Ulfarsson, 2004; Moniruzzaman et al., 2013). However, there are conflicting results for mode preference. With preferences for more comfortable modes, the elderly rely more on automobiles (Alsnih & Hensher, 2003; Rosenbloom, 2001). In contrast, fewer trips and shorter distances reduce the need for car ownership as most locations can be reached on foot or by public transit (Cheng et al., 2019; Olawole & Aloba, 2014; Wong et al., 2018). Overall, among the various age groups, the elderly and young populations have many restrictions on and other factors involved in mode choice. If the portions of these age groups are considerable, that can significantly influence the city-level modal split. We thus use the rates of the population whose ages are more than 65 or under 15 – the elderly and the young – to study impacts on modal split.

The employment rate and income level can represent the economic status of people residing in a city (Foard et al., 2013). The employment rate reflects job opportunities and employment density and, similar to population density, can be leveraged to develop further transit infrastructure and promote mixed land use; this consequently shortens travel distances (Sun et al., 2017); people in areas with high employment rates also tend to use low-emission modes. On the other hand, income level can indicate purchasing power, which is positively related to car ownership (Ding et al., 2017; Li et al., 2010; Sun et al., 2017; Zhang et al., 2017). In addition, people with high earnings usually have more social activities, leading to more trips (Giuliano & Narayan, 2003), which pushes them to own cars. Therefore, we use the employment rate and hourly earnings to elucidate the relationship between the city's economic indicators and the modal split.

3.3. Planning factor

In contrast to the other two factors, the planning factor offers practical opportunities for control during short- and long-term periods. We select nine variables in terms of transportation infrastructure, services, and price.

Improving the service quality and quantity of infrastructure of a specific travel mode can attract potential users. For example, providing built environments for public transport services enhances the accessibility to these services and increases demand for public transit (Ding et al., 2017; Kenworthy & Laube, 1999; Soltani & Allan, 2006; Zhang et al., 2017). In a similar vein, the supply of road infrastructure drives car ownership (Li et al., 2010). Therefore, transportation infrastructure and services are essential to determine the city-level modal split.

Among several transportation service and infrastructure performance measures, we focus on accessibility and capability. First, accessibility means the ease of reaching a certain transportation infrastructure, i.e., opportunities to use transportation in a given area (Litman, 2017). To quantify the accessibility, we use the densities of metro stations and bike-sharing stations (if dock-based), defined as the total number of stations divided by the territorial area. In addition, we use the number of bike-sharing programs, which provide ease of access to bicycles, even if people do not own bicycles (Ma et al., 2020).

Second, we use the metro, bike lanes, and road network densities to quantify the capability. The network density is obtained by transport infrastructure length per population, which means the maximum infrastructure capacity for current and potential users (Dingil et al., 2018). It also implies the level of investment in the transport network according to the number of inhabitants.

Taxation on car usage and subsidies for public transit represent existing push and pull measures against car dependency. As a frequently mentioned push measure for reducing automobile dependency, we first use tax rates on gasoline averaged by country (OECD, 2020; Santos et al., 2013). Second, we use public transit fares, for which subsidies can reduce trip burden and, as a pull measure, induce a modal shift away from private vehicles (Batty et al., 2015; Buehler et al., 2017). Taxi fares are also used because taxis compete with public transit and their fares indirectly affect the demand for public transit (Wong et al., 2018). To understand the value of a city's transportation among the city's overall goods and services, we normalize public transit and taxi fares by price level.

3.4. Data collection

We obtain modal split values for each city from official governments and research institutes. Data for explanatory variables mainly originate from official reports for each city and from international organizations; in cases in which it was challenging to obtain city-level data, these data are replaced with national-level values. If it is not possible to obtain national-level values, we refer to news articles. The data we collected is shared in an external repository.[1]

4. Methodology

This paper proposes a categorical approach to understand the impact of city conditions on modal splits. Existing studies analyzed the impact of influential variables on the modal split and mode choice by using conventional statistical methods such as the linear regression and logit models (Buehler, 2011; Ding et al., 2017; Ha et al., 2020; Liu et al., 2015; Ma et al., 2020; Paulley et al., 2006; Santos et al., 2013; Scheiner, 2010; Sun et al., 2017).

Those studies explicitly showed the relationship between quantifiable mode choice indices and explanatory variables based on their linearity. However, first, it is difficult to explain the exact values of multiple cities' modal splits because a modal split reflects city-specific features such as history, culture, land use, industries, and relationships with other cities or adjacent countries. Second, cities' explanatory variables interact with each other, which leads to interlocking relationships that determine the modal split. In this process, a hierarchical relationship between explanatory variables may exist, which is fundamentally accompanied by nonlinearity in determining the modal split (Yang & Zhou, 2020).

To overcome the above-mentioned limitations of using linear models, we propose two-step data mining techniques, as shown in Fig. 2. To overcome the first limitation, we classify cities with similar modal splits by the K-Means method and define the city type with their corresponding trip mode in the first step. To overcome the second limitation, we adopt a decision tree method in the next step, which evaluates the impact of determinants on the modal group classification and captures the hierarchical and conditional relationships between explanatory variables and modal groups.

4.1. City clusters based on modal split by K-Means method

We use the K-Means method to categorize similar cities based on their modal splits. The K-Means method is a clustering algorithm that groups objects into K clusters based on the similarity between the objects in the input data (Lloyd, 1982). Each object refers to a city and contains its citywide modal split information, i.e., the share of using private vehicles, public transit, and bicycles. The K-Means method forms city clusters grouped based on the similarity of modal splits – modal groups. To measure the similarity of modal splits between cities, we use the Euclidean distance, which is the most commonly used distance function.

We set the number of clusters to three ($K = 3$) because we intend to classify city types into three modal groups of private vehicle, public transit, and bicycle, with groups named according to the dominant travel mode between clusters. To utilize this method, we use the K-Means function in the fpc package in R software.

4.2. Hierarchical structure of explanatory variables by decision tree method

In the second step, we use a decision tree method to measure the impact of the explanatory variables on the modal groups determined in the first step. The decision tree method recursively finds the optimal explanatory variable and its value, classifying data into subgroups with identical values of the dependent variable. As a result, this model

[1] Data for the study sites are shared via this link (https://docs.google.com/spreadsheets/d/1CfKN3Q5Z0RWOOCJD-j1jW42PRloiYnSB_Bk-tXU3zTs/edit?usp=sharing). It includes the values of modal splits and explanatory variables described in Table 1. It also contains outcomes for cities clustered by their modal splits, which will be presented in the Results section.

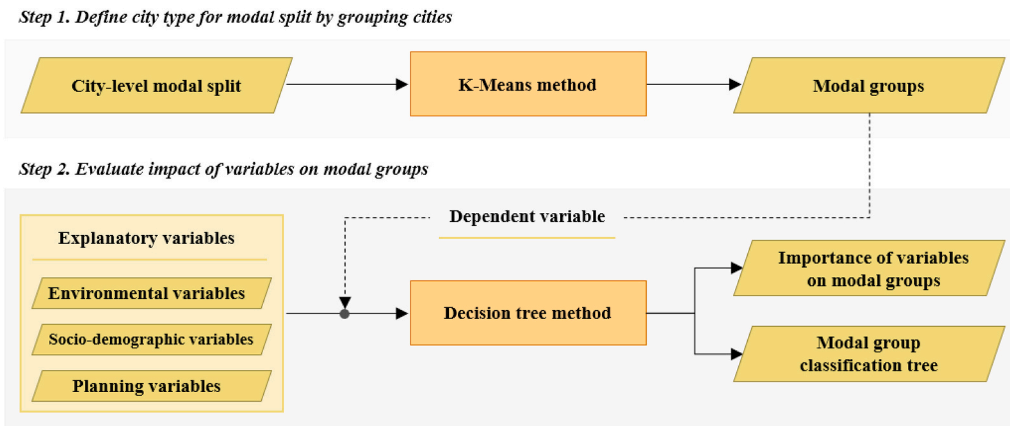


Fig. 2. Two-step methodology framework for evaluating impact of environmental, socio-demographic, and planning variables on city-level modal split.

generates combined conditions using multiple explanatory variables to explain the dependent variable (Yang & Zhou, 2020).

Among several decision tree algorithms, we use the classification and regression tree (CART) method of Breiman et al. (1984). Because it is applicable to both categorical and numeric variables, CART is suitable for our study with its categorical dependent variables of private vehicle, public transit, or bicycle groups. The 17 explanatory variables, categorized into environmental, socio-demographic, and planning factors, are employed to explain the city conditions that determine the modal groups. As a result, we obtain a hierarchical structure representing the relationship between explanatory variables and modal groups and the importance of each variable. The impact of each variable, called the variable importance, is quantified by summing all potential contributions to splitting data with the same class (e.g., cities belonging to the same modal group) into the same partition. Even if a variable is not directly used to split the data, its potential contributions might be reflected when it is used as a surrogate variable for missing values or is highly correlated with a split variable.

Using the rpart package for CART in R software (Atkinson & Therneau, 2000), we iteratively run the function by changing significant parameters within certain ranges. The minimum numbers of cases (cities) for the parent node and the child node range from 1 to 4 and 1 to 10, respectively. The maximum tree depth is set from 3 to 10, and the complexity parameter has a value between 0 and 0.1. Controlling parameters, we take a decision tree with ease of explanation and higher accuracy, measured by the number of correctly classified objects divided by the total number of objects.

5. Results

5.1. Cities by modal split cluster

For 46 cities, we implement the K-Means method with the city-level modal splits and group the cities into three clusters. Fig. 3(a) shows that each cluster has a higher share than other clusters in specific modes. **Cluster 1** had more than twice as much share of private vehicles as did other clusters; **Cluster 1** is called the "Private vehicle group." **Cluster 2** has the highest share of public transit among all clusters. **Cluster 2** is considered the "Public transit group." **Cluster 3** has ranges similar to those of **Cluster 2** in the share of private vehicles. However, a distinct difference can be found in the share of bicycles: cities in **Cluster 3** have a higher rate of mode share for bicycles; we call **Cluster 3** the "Bicycle group."

Fig. 3(b) shows that the public transit group has the largest share among the modal groups, followed by the private vehicle group and the bicycle group.

Fig. 4 shows the regional dependency, especially in certain cities in the private vehicle and bicycle groups. All cities in the USA are categorized into the private vehicle group, reflecting the high automobile dependency in the USA (Buehler, 2011; Kenworthy & Laube, 1999). Most cities in the bicycle group are in Europe. Amsterdam and Copenhagen are especially well-known for beginning bike-sharing programs (DeMaio, 2009). On the other hand, Beijing is an Asian city within the bicycle group, and China, where Beijing is located, is known for its vast bike-sharing market (Eren & Uz, 2020).

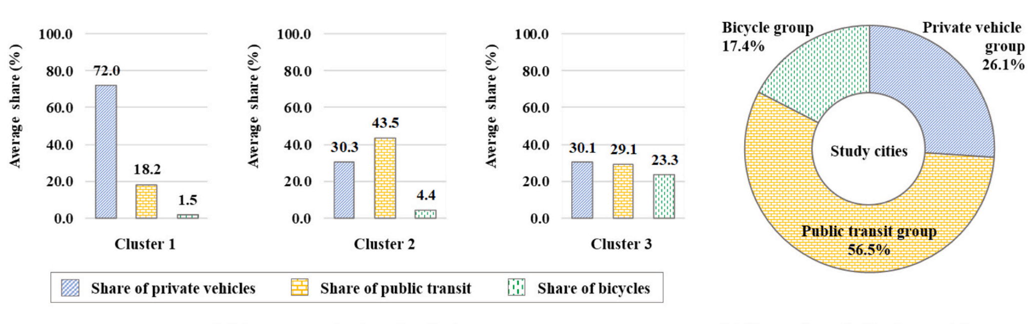


Fig. 3. (a) Average mode share and (b) proportion of cities by modal groups.

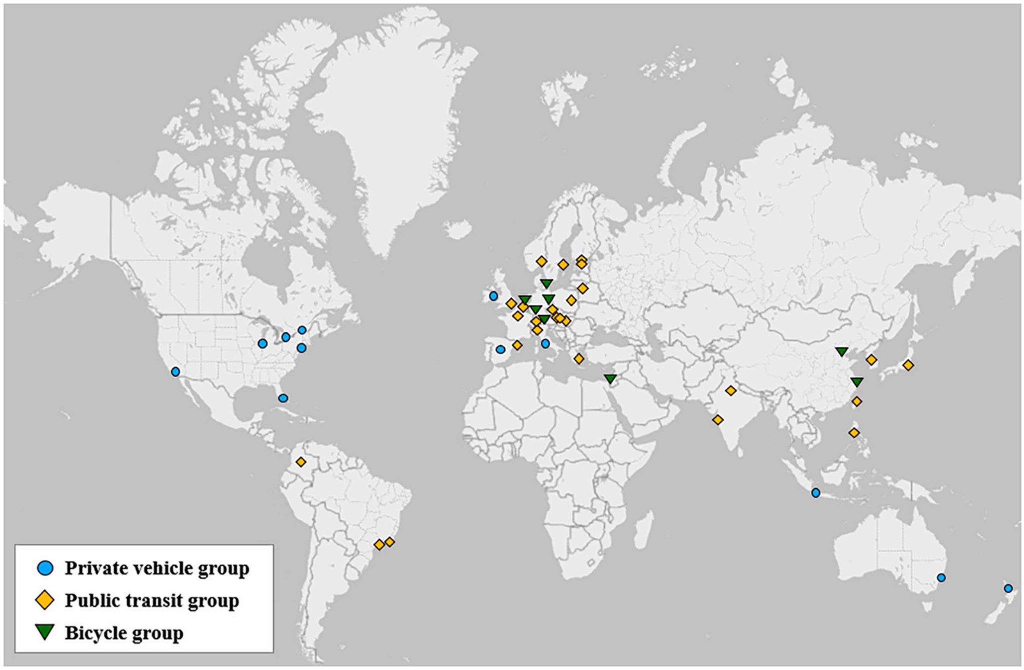


Fig. 4. Geographical distribution of cities by modal groups – private vehicle, public transit, and bicycle groups.

5.2. *Hierarchical structure of factors affecting city-level modal splits*

To understand the integrated influences determining the modal group, we use the decision tree method to draw a hierarchical structure with all variables from the three factors. Fig. 5 shows the importance of variables after quantifying their contributions to the modal group classification.

This figure shows that the socio-demographic variables, especially the elderly and young rates, have the largest impacts on modal split. Among the three factors, the socio-demographic factor is most closely related to human behaviors. Because public parties consider the general attributes of residents as they attempt to make appropriate city plans, it makes sense that the planning factor, such as bicycle- and metro-related variables, follows the socio-demographic factor.

The environmental factor is shown to be less important than other factors. Although its quantified importance is relatively insignificant, we cannot regard it as meaningless overall. People behave and plan by adapting to their given environmental conditions, and it can be revealed in terms of lifestyle, which also connects to two other factors. Therefore, we can analogize the indirect effects of the environmental factor on the modal split.

The decision tree in Fig. 6 shows with 95.65 % accuracy the nested conditions of the modal group. It classifies the cities using six variables: tax rate on gasoline, elderly rate, number of bike-sharing programs, average temperature, population density, and bike lane network density by population. Two variables shown to be important in Fig. 5, core land area and young rate, are not included because they are strongly correlated with population density and elderly rate, respectively, which are

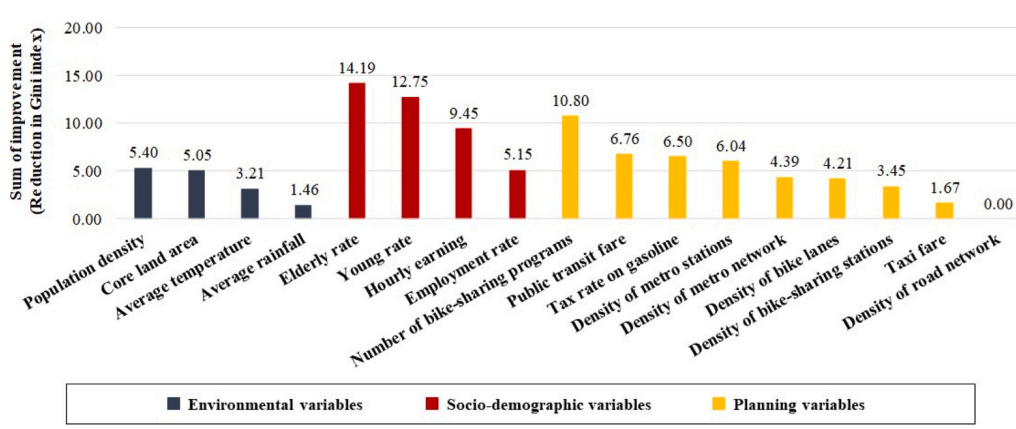
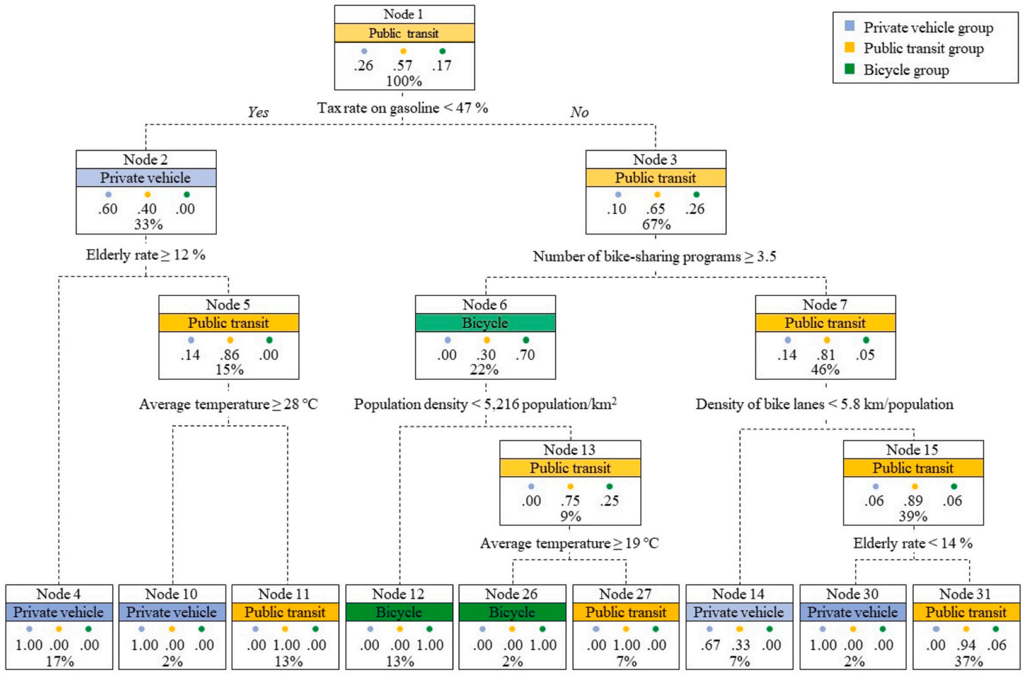


Fig. 5. Importance of variables in determining modal groups.



** Notes (1) Each node contains as many cities as the indicated percentage.*
(2)The decimals in each node expresses the proportion of cities categorized in the three modal groups.
(3) Saturation means purity of classification.

Fig. 6. Modal group classification tree with variables for three factors.

already included in the tree.

In the root node (**Node 1**), the tax rate on gasoline is first used to classify the cities. Of the cities where the tax rate on gasoline is lower than 47 %, 60 % are in the private vehicle group (**Node 2**); only 10 % of cities with higher gasoline taxes are in the private vehicle group (**Node 3**). This implies that taxes on vehicle fuel can determine whether a city is associated with high private vehicle share or not. In addition, notable differences have been observed among regions. The USA applies lower gasoline tax rates, while most European countries have high gasoline tax rates.

As the next factor splitting low gasoline tax cities, the elderly rate is used. The results show that cities with elderly population rates of more than 12 %—mainly cities in the USA and Canada—are entirely classified as private vehicle group (**Node 4**).

At **Node 5**, the cities in the public transit group are dominant; in these, both gasoline tax and the elderly rate are low. However, if they have average temperatures higher than 28 °C, they are classified one step further into the private vehicle group, not the public transit group (**Node 10**). This shows that high temperatures can be an obstacle to the use of public transit.

Among cities with higher gasoline taxes, the number of bike-sharing programs determines whether the city belongs to the bicycle group or the public transit group (**Node 3**). If there are more than three bike-sharing programs, the cities are in the bicycle group (**Node 6**). In this case, bike-sharing programs can either induce additional bike trips or be facilitated by high bicycle demand. Either or both are possible. Furthermore, as shown in **Node 12**, lower population density might cause a lower density of transportation infrastructure, but can strengthen the conditions for forming a bicycle group if the gasoline tax is too high. Conversely, higher population density and sufficient bike-sharing programs characterize the public transit group (**Node 13**), in which bike-sharing programs act as feeders for public transit.

In general, cycling is the travel mode that is most vulnerable to hot weather conditions (Böcker et al., 2013). In alignment with this, the average temperature of cities in the bicycle group is around 12 °C; this is the lowest among the modal groups. However, outcomes partitioned from **Node 13** show an exceptional case in which Tel Aviv, where the average temperature is higher than 19 °C, is in the bicycle group. Considering that annual rainfall of about 530 mm places Tel Aviv in the lowest level, this exceptional case may be due to a minor effect between temperature and precipitation.

Transportation infrastructure for a certain travel mode can affect demand for other modes. Of cities with high gasoline tax but few bike-sharing programs, 81 % belong to the public transit group, as in **Node 7**. However, in this case, if bike lane network is insufficient, cities will belong to the private vehicle group, not the public transit group, as in **Node 14**; Rome and Madrid are examples. This is because, in a city that is not bike-friendly, accessibility of public transit by bicycles is low, so the use of public transit is also hindered.

Node 15 shows that cities with elderly rates higher than 14 % are in the public transit group. The cities in **Node 31** are mainly located in Eastern and Western Europe and Eastern Asia; especially in Europe, public transit use for people over 75 years of age has grown (Alsnih & Hensher, 2003). Across **Node 2** and **Node 15**, it is observed that the elderly rate has different impacts in car-friendly cities and public transit-friendly cities; this means a non-linear relationship between the elderly rate and the modal split.

Figs. 5 and 6 show which variables or factors are the most critical among all factors in determining city-level modal splits. However, some variables identified as important in Fig. 5, such as the public transit fare, are not shown in Fig. 6 because the tree depth is insufficient. If a good quantity of city data is available, we can increase the tree depth to unveil the effects of such hidden variables. To investigate the impact of the explanatory variables that are statistically significant but not included in Fig. 6 due to the limited tree depth and overlapping impact of other variables on modal splits, we apply the decision tree method to the variables included in each factor category to reduce the total number of explanatory variables in each tree; this allows us to show more the most important variables in each factor.

5.3. Effects of environmental variables on modal groups

In Fig. 7, the decision tree uses all environmental variables: core land area, population density, average temperature, and average rainfall. Excluding the impact of socio-demographic and planning factors, it also explains how core land area and rainfall, not shown in Fig. 6, affect modal groups.

The first split variable is the core land area, representing the size of the functional region of the city. Comparing the proportions of cities in the public transit group, **Node 2** for cities with 1237 km^2 or more of core land area characterizes the public transit group less well than does **Node 3** with relatively small cities. In addition, the proportion of the private vehicle group is much higher in **Node 2**. The travel time and distance, especially for commuting, are generally proportional to the size of the city, even if these can vary with city forms (e.g., monocentric or polycentric) (Schwanen, 2002; Vanoutrive, 2015). Longer travel distances can make people rely on automobiles (Jiang et al., 2021; Scheiner, 2010; Xiao et al., 2021a; Zhang et al., 2017). This effect can be intensified by low population density (**Node 4**).

Similar to the role of population density, shown in Fig. 6, lower-density areas might mean low development density, which can be linked to insufficient public transit infrastructure. Suppose a city's condition is not amenable to public transit development or usage. In such a case, it may make sense to choose a private vehicle, especially in cities with large functional areas where people make more long-distance trips. In terms of economies of scale, dense cities with sufficient demand for trips can guarantee revenue from investment in public transit. In addition, considering the Mohring effect, high population density can contribute to positive interaction of supply of and demand for public transit. On the other hand, in small living areas (**Node 3**), high population density might play a role in shortening travel distances rather than guaranteeing demand on public transit compared to investment. The difference in the proportion of bicycle groups in **Node 3** and **Node 7** implies that high population density in a small area can make cycling more prevalent.

Even if the city size or population density is appropriate for specific modes, weather can be a barrier. Regardless of other preliminary conditions, general temperature distribution in modal groups proves that private vehicles are more practical and robust against extreme weather conditions such as high temperature and frequent rain. However, **Node 5** classifies high-temperature cities as a public transit group, which may be due to the regional dependence of southern Europe. **Node 10** explains the impact of rain on modal groups. Furthermore, adequate weather conditions for cycling are observed at **Node 6** and **Node 12**. In small and compact cities, people tend to enjoy cycling at moderate temperatures of around 10 °C. If we can consider additional cities in cold regions with lower temperatures, not grouped in the bicycle groups, it will be possible to show that extremely low temperature negatively affects bicycle use, which is not addressed in the above decision tree.

5.4. Effects of socio-demographic variables on modal groups

The decision tree in Fig. 8 uses three socio-demographic variables: hourly earnings, employment rate, and elderly rate. The elderly rate

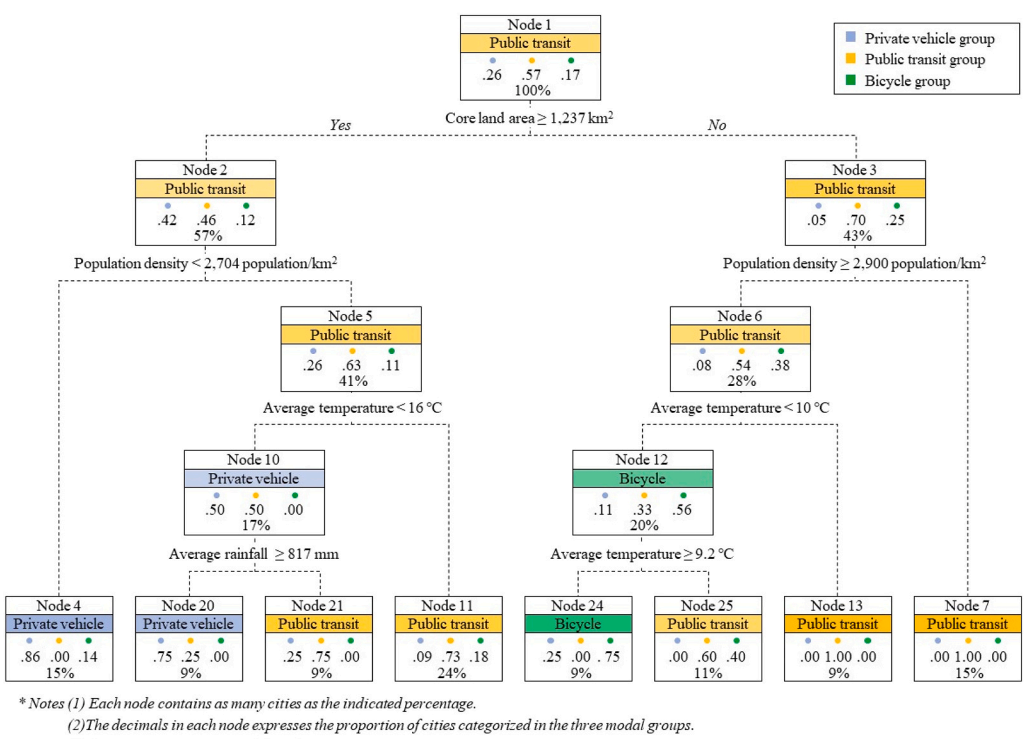
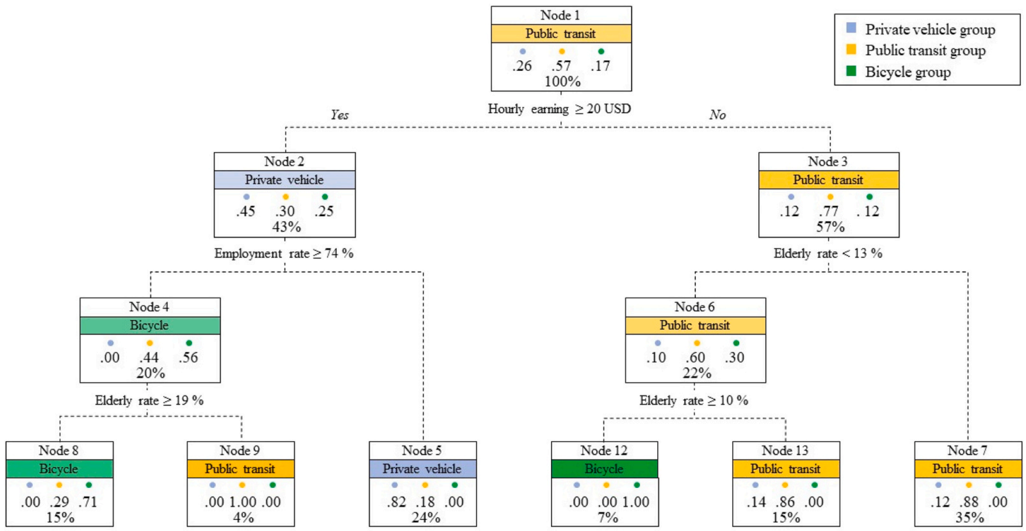


Fig. 7. Modal group classification tree with environmental variables.



** Notes (1) Each node contains as many cities as the indicated percentage.*
(2)The decimals in each node expresses the proportion of cities categorized in the three modal groups.
(3) Saturation means purity of classification.

Fig. 8. Modal group classification tree with socio-demographic variables.

frequently appears, but the young rate is not shown in the tree. Because the young rate negatively correlates with the elderly rate but is less critical, the tree primarily chooses the elderly rate when both variables compete for use in the split.

At the top of the tree, the category of hourly earning separates the private vehicle and public transit groups. Cities with more than 20 USD hourly earnings are likely to be in the private vehicle group. This implies that high income, representing purchasing power, affects car ownership and car usage.

Interestingly, the employment rate has a different effect, although it is similar to hourly earnings in that it is a labor-related variable. In branches from **Node 2**, we observe that cities with employment rates higher than 74 % are in the bicycle group or the public transit group; on the other hand, cities with low employment rates belong to the private vehicle group. There is also, however, a regional dependency in the city distributions: **Node 4** mainly consists of Northern and Western European cities, while half of the cities in **Node 5** are in the USA and Canada. Considering that socio-demographic factors can be similar for neighboring cities and countries, outcomes from **Node 2** can imply that commuters living in dense areas with good labor and job opportunities easily reach their workplaces by public transit or bicycle. On the other hand, commuters living in areas with low employment density seem to be forced to prefer private vehicles for long-distance travel to work.

The nonlinearity in the elderly rate, caused by the abovementioned conditions – hourly earning and employment rate – is observed across **Node 3**, **Node 4**, and **Node 6**. **Node 4**, including cities with high-earning and high employment rates, includes most cities with elderly rates higher than 19 %, such as Amsterdam and cities in Germany, into the bicycle group. However, in relatively low-earning cities, elderly rates of more than 13 % push cities into the public transit group in **Node 7**. On the other hand, cities in which the elderly rate is between 10 % and 13 % belong to the bicycle group in **Node 12**.

5.5. Effects of planning variables on modal groups

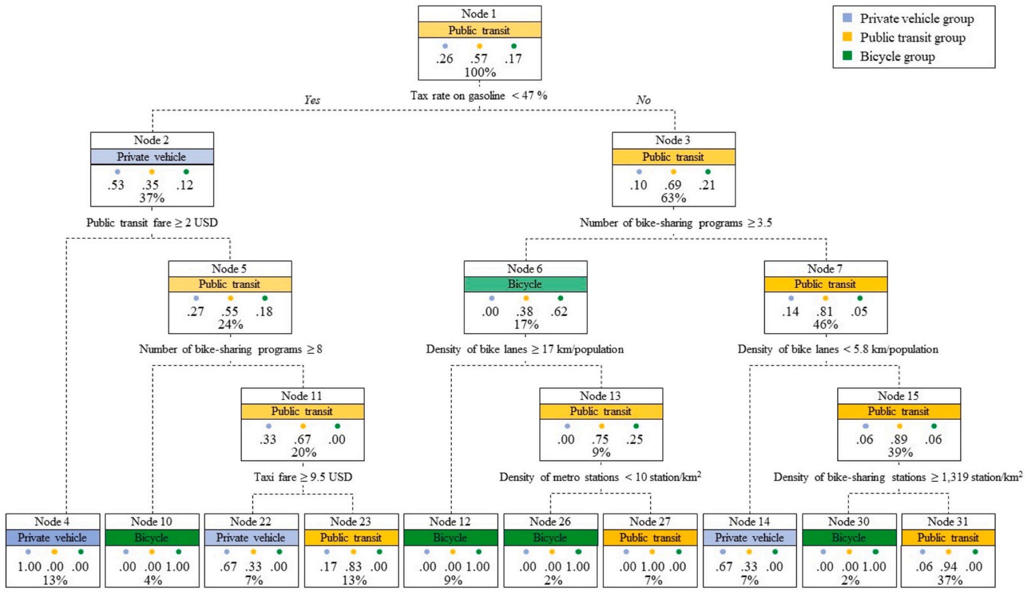
Among planning variables, gasoline taxes, number of bike-sharing programs, public transit fares, taxi fares, bike lane network density, bike-sharing station density, and metro station density are identified as meaningful to explain modal splits.

At the first split of the tree in Fig. 9, it is inferred that dependency on automobiles can increase with a lower gasoline tax, consistent with the tree in Fig. 6. In addition, **Node 4**, partitioned from **Node 2**, demonstrates that the high public transit fares can induce automobile dependence in cities with low gasoline taxes. Therefore, considering this finding and the modal group distribution in **Node 5**, we suggest that lower public transit fares can be a strategy to overcome lower gasoline taxes. In addition, in **Node 11**, the taxi fare, which has a positive correlation with public transit fares, shows a similar effect.

The relationship between transport facilities and modal group is mainly covered in the right subtree from **Node 1**. It can be seen that whether a city is in the bicycle group or the public transit group relies on the facilities corresponding to each mode. The number of bike-sharing programs is positively related to high demand for bicycle use, as observed over partitions from **Node 3** (and **Node 5** in the left subtree). Moreover, **Node 6** and **Node 15** indicate that developing bike-related infrastructure, such as bike lanes and bike-sharing stations, pushes cities into the bicycle group. Interestingly, **Node 13** shows that cycling is attractive even in areas with insufficient bike paths if there is low accessibility to the metro. This demonstrates that cycling, when it cannot act as a feeder for public transit, can be a competitor.

6. Discussion

Developing strategies or policies for modal shift from high-emission modes is vital for forming sustainable mobility infrastructure at the city level. However, modal splits result from complex inter-relations among environmental conditions, people, and resources in cities (Pucher, 1988), so it is essential to understand the city conditions and how they affect modal splits. Extending existing efforts, which are limited to travel mode- and city-specific cases, this paper aims to understand the hierarchical relationship among city attributes that determine the city-level modal split (private vehicle, public transit, and bicycle) by investigating 46 global city cases. Our research led us to answer the first two research questions in the Results section. This section answers the last question by



** Notes (1) Each node contains as many cities as the indicated percentage.*
(2) The decimals in each node expresses the proportion of cities categorized in the three modal groups.
(3) Saturation means purity of classification.

Fig. 9. Modal group classification tree with planning variables.

providing practical insight (recommendations) focusing on important variables from the results and addresses the need for future research.

6.1. Recommendations

As for general recommendations, decision-makers should first consider what a city has before planning a strategy to manage the modal split. Even a city with strategies or policies suitable for a particular trip mode may be induced into unintended modal groups by other critical variables. In Fig. 6, it can be seen that the gasoline tax rate significantly influences private vehicle dependency, as existing studies have already shown (Santos et al., 2013). Even with this push measure, private vehicles might be preferred to other modes due to direct and indirect effects of other variables such as the elderly rate. In addition, consistency is necessary when implementing modal shift strategies. Increasing the gasoline tax can be a meaningless strategy if the facilities for other modes – bicycles and public transit – are insufficient for any potential demand shifted from private vehicles.

With the nonlinearity in the elderly, the most important variable to determine the modal split, we can consider different options for each city condition to address the mobility of the elderly population to induce a modal shift.

Among cities with low gasoline taxes, those cities that have higher elderly rates are likely to be categorized into the private vehicle group (**Node 4** in Fig. 6). Existing studies observed that car use by the elderly gradually increases (Rosenbloom, 2001). Elderly car owners do not easily change their travel modes even when public transit services are improved by reducing walking distances and transfers (Ha et al., 2020). Also, it is known that daily mobility significantly influences aging people's overall quality of life and well-being (van Hoven & Meijering, 2019). Therefore, it may be practical to focus on improving such persons' driving safety on the road, for example by reinforcing good driving habits, instead of attempting to change their mode from private vehicles to low-emission modes. For example, developing policies and guidelines that regularly examine the driving capabilities of the aging population should be implemented, together with preventive measures for situations in which the elderly are prone to unintended accidents.

In cities with well-developed facilities for low-emission modes, a higher elderly rate can push a city to be included in the public transit group (**Node 31** in Fig. 6). Considering the current and potential demand for public transit by the elderly in such cities, it is necessary to improve those facilities by inclusive design and elderly-conscious implementation. We recommend enhancing serviceability for those elderly residents using public transit by reducing stairs during transfer and promoting reserved seat availability in vehicles (Cheng et al., 2019; Wong et al., 2018). Furthermore, if more demand-responsive transit services are to be introduced, it is important that these services offer a seamless travel experience and are made available for non-digital travelers (such as but not limited to the elderly) to enhance travel spontaneity and afford smooth and uninterrupted travel for all (Sampimon, 2020; van Hoven & Meijering, 2019).

Among planning-related variables, bicycle-related facilities, including bike-sharing programs, their stations, and public bicycle paths, are mainly used to identify modal groups. In Figs. 6 and 9, it can be seen that cities with more bicycle facilities belong to the bicycle or public transit groups. Installing bicycle-related facilities can be one strategy to increase public transit demand rates because bicycles can improve accessibility to public transit by assisting in first and last mile travels.

6.2. Future work

In our study, the consistency of temporal scope among the variables is of concern due to the hurdles involved in data gathering for worldwide cities. Considering inertia, as well as plans that current cities have, our results are explanatory, but it is necessary to clarify the causal effects of planning variables on the modal split. Therefore, as future research directions, we will first evaluate via longitudinal study practical policies and transportation plans to induce modal shifts from high-emission modes (private vehicles) to low-emission modes (public transit, cycling, and walking). This study does not employ absolute values to define target modal splits but three categories of it. To make this study more actionable, decision-makers can set their target modal splits from the current using exact values, such as from car 60 %, public transit 30 %, and bicycle 10 % to car 40 %, public transit 40 %, and bicycle 20 %, and we can recommend types and intensities of short-term (related to planning variables) and long-term (related to socio-demographic variables) policies that may be needed under given uncontrollable factors. Furthermore, to bring emerging transportation concepts, such as Mobility as a Service (MaaS), micro-mobility, and mobility-sharing programs to the marketplace, we need to more clearly define aspects that cause modal shifts by focusing on holistic travel behaviors and unique user experiences (Huang et al., 2020; Xiao et al., 2020; Xiao et al., 2021b). This will help us contrive strategies to consider a given city's conditions and attract target users to emerging modes that promote sustainability. As such, the revealed underlying factors that influence the modal shift will be used for the creation of city-level characteristics or personas.

7. Conclusions

Before planning a strategy for a target modal split, such as low automobile dependency or a switch to public transportation use, decision-makers should pay special attention to what a city has and how city attributes affect the modal split within its system boundary. In addition, for practical suggestions, the city attributes should be considered as to their controllability. In this context, this study elucidated the impact of city attributes, categorized into environmental, socio-demographic, and planning realms, on the modal split at the city level. With data drawn from 46 global city cases, we employed a two-step data mining framework, composed of K-Means and decision tree methods, to obtain an interpretable model for a city-level modal split.

As for findings for the environmental variables, it was demonstrated that a city with a dense population is highly associated with low-emission modes; among socio-demographic variables, a similar relationship was observed for employment rate. Cycling is the travel mode most vulnerable to hot weather conditions. This study found that the socio-demographic variables, especially the elderly and young rates, are most important in determining the modal split. In cities with high hourly earnings, the share of private vehicles is high, similar to the effect of income level on car ownership (Ding et al., 2017; Sun et al., 2017; Zhang et al., 2017). Among planning variables, bicycle-related facilities, especially the number of bike-sharing programs within a city, can be a solid indicator of cities with lower dependency on motorized modes. In addition, a high gasoline tax and low public transit fares are positively associated with low automobile dependency.

This study also revealed several important discussion points and allowed us to use them for recommendations for a target modal split. As a broad recommendation, because of the interlocking relationships among variables involved in the modal split, what cities have should be studied in terms of multiple factors before planning a modal shift strategy at the city level. Notably, facilities and policies for specific modes can, directly and indirectly, affect demand for other modes; consistency in the overall city plan to achieve the targeted modal split is necessary. In the practical discussion of the results, modal shift strategies for the elderly population should consider that population's travel behaviors and safety: depending on the city archetype, different plans for the elderly will be necessary. Lastly, improving bicycle-related infrastructure is recommended to reduce usage of high-emission modes (Shaheen et al., 2013) and increase general public transit demand.

CRediT authorship contribution statement

Sujin Lee: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization

Jinwoo Lee: Conceptualization, Methodology, Formal analysis, Resources, Writing – original draft, Funding acquisition

Suzanne Hiemstra-van Mastrigt: Methodology, Validation, Data curation, Writing – review & editing

Euiyoung Kim: Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) [2020R1C1C1005034 and 2021R1A4A1033486] and the KAIST-KU Joint Research Center, KAIST, Korea. We thank the Delft University of Technology for the funding of the Open Access Article Publishing Charge. The researchers acknowledge our scientific research staff in the Mobility program at Faculty of Industrial Design Engineering, Delft University of Technology, in the Netherlands for their insight and constructive feedback in developing the framework presented in the paper.

References

1 Alsnih, R., & Hensher, D. A. (2003). The mobility and accessibility expectations of seniors in an aging population. *Transportation Research Part A: Policy and Practice, 37*(10), 903–916. https://doi.org/10.1016/S0965-8564(03)00073-9

2 Antipova, A., Wang, F., & Wilmot, C. (2011). Urban land uses, socio-demographic attributes and commuting: A multilevel modeling approach. *Applied Geography, 31*(3), 1010–1018. https://doi.org/10.1016/j.apgeog.2011.02.001

3 Atkinson, E. J., & Therneau, T. M. (2000). An introduction to recursive partitioning using the RPART routines. *Mayo Clinic, 61*, 33.

4 Batty, P., Palacin, R., & González-Gil, A. (2015). Challenges and opportunities in developing urban modal shift. *Travel Behaviour and Society, 2*(2), 109–123. https://doi.org/10.1016/j.tbs.2014.12.001

5 BiTiBi.eu. (2017). BiTiBi - Final report. Retrieved from www.bitibi.eu/dox/BitiBi_Final%20Report_2017.pdf Accessed January 04, 2021.

6 Böcker, L., Dijst, M., & Prillwitz, J. (2013). Impact of everyday weather on individual daily travel behaviours in perspective: A literature review. *Transport Reviews, 33*(1), 71–91. https://doi.org/10.1080/01441647.2012.747114

7 Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees* (1st ed.). Abingdon: Routledge. https://doi.org/10.1201/9781315139470

8 Buehler, R. (2011). Determinants of transport mode choice: A comparison of Germany and the USA. *Journal of Transport Geography, 19*(4), 644–657. https://doi.org/10.1016/j.jtrangeo.2010.07.005

9 Buehler, R., Pucher, J., Gerike, R., & Götschi, T. (2017). Reducing car dependence in the heart of Europe: Lessons from Germany, Austria, and Switzerland. *Transport Reviews, 37*(1), 4–28. https://doi.org/10.1080/01441647.2016.1177799

10 C40 Cities. (2019). Defining carbon neutrality for cities & managing residual emissions—Cities' perspective & guidance. Retrieved from https://www.c40knowledgehub.org/s/article/Defining-carbon-neutrality-for-cities-and-managing-residual-emissions-Cities-perspective-and-guidance?language=en_US. (Accessed 4 May 2021). Retrieved from.

11 Cheng, L., Chen, X., Yang, S., Cao, Z., De Vos, J., & Witlox, F. (2019). Active travel for active ageing in China: The role of built environment. *Journal of Transport Geography, 76*(November 2018), 142–152. https://doi.org/10.1016/j.jtrangeo.2019.03.010

12 DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation, 12*(4), 41–56. https://doi.org/10.5038/2375-0901.12.4.3

13 Ding, C., Wang, D., Liu, C., Zhang, Y., & Yang, J. (2017). Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance. *Transportation Research Part A: Policy and Practice, 100*, 65–80. https://doi.org/10.1016/j.tra.2017.04.008

14 Dingil, A. E., Schweizer, J., Rupi, F., & Stasiskiene, Z. (2018). Transport indicator analysis and comparison of 151 urban areas, based on open source data. *European Transport Research Review, 10*(2). https://doi.org/10.1186/s12544-018-0334-4

15 Eren, E., & Uz, V. E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society, 54*(October 2019). https://doi.org/10.1016/j.scs.2019.101882

16 European Commission. (2013). *Promotion of Clean and Energy Efficient Road Transport Vehicles: Report From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brussels*.
17 European Union. (2019). Regulation (EU). 2019/631 of the European Parliament and of the Council of 17 April 2019 setting CO2 emission performance standards for new passenger cars and for new light commercial vehicles, and repealing regulations (EC) No 443/2009 and (EU) No 510/2011. *Official Journal of the European Union, 111*, 13–53. L 2019.
18 Foard, A. A., Liao, B., Carneiro, B., Mikhailov, D., Dumitrescu, E., Rorimer, J., Tatum, J., Simon, K., Lom, L., & Zhang, Y. (2013). *Assessing city economic performance: The city economic capacity index*.
19 Giuliano, G., & Narayan, D. (2003). Another look at travel patterns and urban form: The US and Great Britain. *Urban Studies, 40*(11), 2295–2312. https://doi.org/10.1080/0042098032000123303
20 Ha, J., Lee, S., & Ko, J. (2020). Unraveling the impact of travel time, cost, and transit burdens on commute mode choice for different income and age groups. *Transportation Research Part A: Policy and Practice, 141*(August 2019), 147–166. https://doi.org/10.1016/j.tra.2020.07.020
21 Haustein, S., & Nielsen, T. A. S. (2016). European mobility cultures: A survey-based cluster analysis across 28 European countries. *Journal of Transport Geography, 54*, 173–180. https://doi.org/10.1016/j.jtrangeo.2016.05.014
22 Hawley, A. H. (1972). Population density and the city. *Demography, 9*(4), 521–529. https://doi.org/10.2307/2060663
23 Huang, Y., Xiao, Z., Member, S., Wang, D., Jiang, H., Member, S., & Wu, D. (2020). Exploring individual travel patterns across private car trajectory data. *IEEE Transactions on Intelligent Transportation Systems, 21*(12), 5036–5050. https://doi.org/10.1109/TITS.2019.2948188
24 Jiang, H., Zhang, Y., Xiao, Z., Zhao, P., & Iyengar, A. (2021). An empirical study of travel behavior using private car trajectory data. *IEEE Transactions on Network Science and Engineering, 8*(1), 53–64. https://doi.org/10.1109/TNSE.2020.3025529
25 Kenworthy, J. R., & Laube, F. B. (1999). Patterns of automobile dependence in cities: An international overview of key physical and economic dimensions with some implications for urban policy. *Transportation Research Part A: Policy and Practice, 33* (7–8), 691–723. https://doi.org/10.1016/S0965-8564(99)00006-3
26 Kim, S., & Ulfarsson, G. F. (2004). Travel mode choice of the elderly: Effects of personal, household, neighborhood, and trip characteristics. *Transportation Research Record, 1894*, 117–126. https://doi.org/10.3141/1894-13
27 Klinger, T., Kenworthy, J. R., & Lanzendorf, M. (2013). Dimensions of urban mobility cultures - A comparison of German cities. *Journal of Transport Geography, 31*, 18–29. https://doi.org/10.1016/j.jtrangeo.2013.05.002
28 Lee, Y. Y., Md Din, M. F., Ponraj, M., Noor, Z. Z., Iwao, K., & Chelliapan, S. (2017). Overview of Urban Heat Island (UHI) phenomenon towards human thermal comfort. *Environmental Engineering and Management Journal, 16*(9), 2097–2112. https://doi.org/10.30638/eemj.2017.217
29 Li, J., Walker, J. L., Srinivasan, S., & Anderson, W. P. (2010). Modeling private car ownership in China: Investigation of urban form impact across megacities. *Transportation Research Record, 2193*, 76–84. https://doi.org/10.3141/2193-10
30 Litman, T. (2017). Evaluating accessibility for transportation planning: Measuring people's ability to reach desired goods and activities. *Transportation Research*, (January 2008), 62.
31 Liu, C., Susilo, Y. O., & Karlström, A. (2015). The influence of weather characteristics variability on individual's travel mode choice in different seasons and regions in Sweden. *Transport Policy, 41*, 147–158. https://doi.org/10.1016/j.tranpol.2015.01.001
32 Lloyd, S. P. (1982). Least-squares quantization in PCM. *IEEE Transactions on Information Theory, 28*, 129–137.
33 Ma, X., Yuan, Y., Van Oort, N., & Hoogendoorn, S. (2020). Bike-sharing systems'impact on modal shift: A case study in Delft, the Netherlands. *Journal of Cleaner Production, 259*, Article 120846. https://doi.org/10.1016/j.jclepro.2020.120846
34 May, A. D. (2015). Encouraging good practice in the development of sustainable urban mobility plans. *Case Studies on Transport Policy, 3*(1), 3–11. https://doi.org/10.1016/j.cstp.2014.09.001
35 McCarthy, L., Delbosc, A., Currie, G., & Molloy, A. (2017). Factors influencing travel mode choice among families with young children (aged 0–4): A review of the literature. *Transport Reviews, 37*(6), 767–781. https://doi.org/10.1080/01441647.2017.1354942
36 McIntosh, J., Trubka, R., Kenworthy, J., & Newman, P. (2014). The role of urban form and transit in city car dependence: Analysis of 26 global cities from 1960 to 2000. *Transportation Research Part D: Transport and Environment, 33*, 95–110. https://doi.org/10.1016/j.trd.2014.08.013
37 Mohring, H. (1972). Optimization and scale economies in urban bus transportation. *American Economic Review, 62*(4), 591–604. http://www.jstor.org/stable/1806101.
38 Moniruzzaman, M., Páez, A., Nurul Habib, K. M., & Morency, C. (2013). Mode use and trip length of seniors in Montreal. *Journal of Transport Geography, 30*, 89–99. https://doi.org/10.1016/j.jtrangeo.2013.03.007
39 OECD. (2012). *Redefining "urban": A new way to measure metropolitan areas*. OECD Publishing. https://doi.org/10.1787/9789264174108-en
40 OECD. (2020). *Consumption tax trends 2020: VAT/GST and excise rates, trends and policy issues*. OECD Publishing. https://doi.org/10.1787/8b65ce7e-en
41 [dataset] OECD. (2021). *Metropolitan areas*. Retrieved from. OECD Regional Statistics. https://doi.org/10.1787/data-00531-en Accessed February 16, 2021.
42 Olawole, M. O., & Aloba, O. (2014). Mobility characteristics of the elderly and their associated level of satisfaction with transport services in Osogbo, Southwestern Nigeria. *Transport Policy, 35*, 105–116. https://doi.org/10.1016/j.tranpol.2014.05.018
43 Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., & 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. https://doi.org/10.1016/j.tranpol.2005.12.004
44 Pucher, J. (1988). Urban travel behavior as the outcome of public policy: The example of modal-split in Western Europe and North America. *Journal of the American Planning Association, 54*(4), 509–520. https://doi.org/10.1080/01944368808976677
45 Rosenbloom, S. (2001). Sustainability and automobility among the elderly: An international assessment. *Transportation, 28*(4), 375–408. https://doi.org/10.1023/A:1011802707259
46 Sampimon, M. (2020). *The design of an demand responsive transport service for non digital travellers: Design report (Issue October)*. Delft University of Technology. http://resolver.tudelft.nl/uuid:d8368b57-7f8f-405d-9fee-39f76516a041.
47 Santos, G., Maoh, H., Potoglou, D., & von Brunn, T. (2013). Factors influencing modal split of commuting journeys in medium-size European cities. *Journal of Transport Geography, 30*, 127–137. https://doi.org/10.1016/j.jtrangeo.2013.04.005
48 Scheiner, J. (2010). Interrelations between travel mode choice and trip distance: Trends in Germany 1976–2002. *Journal of Transport Geography, 18*(1), 75–84. https://doi.org/10.1016/j.jtrangeo.2009.01.001
49 Schwanen, T. (2002). Urban form and commuting behaviour: A cross-European perspective. *Tijdschrift Voor Economische En Sociale Geografie, 93*(3), 336–343. https://doi.org/10.1111/1467-9663.00206
50 Shaheen, S., Martin, E., & Cohen, A. (2013). Public bikesharing and modal shift behavior: A comparative study of early bikesharing systems in North America. *International Journal of Transportation, 1*(1), 35–54. https://doi.org/10.14257/ijt.2013.1.1.03
51 Soltani, A., & Allan, A. (2006). Analyzing the impacts of microscale urban attributes on travel: Evidence from Suburban Adelaide, Australia. *Journal of Urban Planning and Development, 132*(3), 132–137. https://doi.org/10.1061/(asce)0733-9488(2006)132:3(132)
52 Sun, B., Ermagun, A., & Dan, B. (2017). Built environmental impacts on commuting mode choice and distance: Evidence from Shanghai. *Transportation Research Part D: Transport and Environment, 52*, 441–453. https://doi.org/10.1016/j.trd.2016.06.001
53 United Nations. (1997). Special session for the purpose of an overall review and appraisal of the implementation of agenda 21. In *Proceedings of the 19th special session of the general assembly to review and appraise the implementation of agenda 21, New York, NY, USA*.
54 United Nations. (2015). *Adoption of the Paris agreement. Framework convention on climate change; FCCC/CP/2015/L.9/Rev.1*. Paris, France: UN.
55 van Hoven, B., & Meijering, L. (2019). Mundane mobilities in later life - Exploring experiences of everyday trip-making by older adults in a Dutch urban neighbourhood. *Research in Transportation Business and Management, 30*(August), Article 100375. https://doi.org/10.1016/j.rtbm.2019.100375
56 Vanoutrive, T. (2015). The modal split of cities: A workplace-based mixed modelling perspective. *Tijdschrift Voor Economische En Sociale Geografie, 106*(5), 503–520. https://doi.org/10.1111/tesg.12113
57 Wong, R. C. P., Szeto, W. Y., Yang, L., Li, Y. C., & Wong, S. C. (2018). Public transport policy measures for improving elderly mobility. *Transport Policy, 63*(July 2017), 73–79. https://doi.org/10.1016/j.tranpol.2017.12.015
58 Xiao, Z., Xu, S., Li, T., Jiang, H., Zhang, R., Regan, A. C., & Chen, H. (2020). On extracting regular travel behavior of private cars based on trajectory data analysis. *IEEE Transactions on Vehicular Technology, 69*(12), 14537–14549. https://doi.org/10.1109/TVT.2020.3043434
59 Xiao, Z., Xiao, H., Chen, W., Chen, H., Regan, A., & Jiang, H. (2021). Exploring human mobility patterns and travel behavior: A focus on private cars. *IEEE Intelligent Transportation Systems Magazine*, (August 2021)https://doi.org/10.1109/MITS.2021.3098627
60 Xiao, Z., Fang, H., Jiang, H., Bai, J., Havyarimana, V., Chen, H., & Jiao, L. (2021). Understanding private car aggregation effect via spatio-temporal analysis of trajectory data. *IEEE Transactions on Cybernetics, 1–12*. https://doi.org/10.1109/TCYB.2021.3117705
61 Yang, W., & Zhou, S. (2020). Using decision tree analysis to identify the determinants of residents' CO2 emissions from different types of trips: A case study of Guangzhou, China. *Journal of Cleaner Production, 277*, Article 124071. https://doi.org/10.1016/j.jclepro.2020.124071
62 Zhang, Y., Zheng, S., Sun, C., & Wang, R. (2017). Does subway proximity discourage automobility? Evidence from Beijing. *Transportation Research Part D: Transport and Environment, 52*, 506–517. https://doi.org/10.1016/j.trd.2016.11.009