An Analysis of Seasonal Effects on Transport Mode Choices based on GPS Trajectories and a Multinomial Logistic Regression Model

Timothy Otim, Dina Bousdar Ahmed and Estefania Munoz Diaz
Institute of Communications and Navigation
German Aerospace Center
Oberpfaffenhofen, 82234, Germany
{timothy.otim, dina.bousdarahmed, estefania.munoz}@dlr.de

***Abstract*— To improve the planning of transportation modes, it is necessary to understand how passenger choose the transportation mode depending on factors such as temperature, precipitation, trip distance, trip duration. In this article, we build up on previous work to analyze the effect of seasonality in the choice of transportation modes. To that end, we use GPS trajectories in Beijing from the Geolife Trajectory data set, and analyze them with respect to the four seasons: Summer, Autumn, Winter, and Spring. Our analysis reveals how seasonality influences the choice of transportation mode, e.g. people tend to bike less in Summer as temperature increases, whereas rush hour traffic influences the car share in Autumn. We build a multinomial logit model, one per season, to predict the transportation mode based on the following factors: temperature, air quality, traffic during rush hour, time of day, trip distance and trip duration. The results show how changes in the factors effect the choice of transportation mode. For instance, long trips in Autumn decrease the probability of walking by 54% and increase the probability of car, bus and train use b 37.7%, 11.1% and 13.5%, respectively. We conclude that rush hours only influence the choice of transport mode during Autumn and Spring. Also, we have observed some similarities in the choice of transport modes throughout different seasons, e.g. the trip distance is a relevant factor for the choice of transport mode in all four seasons.**

Keywords— Air quality; Geolife; Traffic demand; Transport planning; Transport regulation;Social Transportation; Shared Mobility

I. Introduction

Mobility issues are increasingly important all over the world particularly in developing countries which are undergoing rapid economic and population growth, motorization, and urbanisation. Finding sustainable mobility solutions, which do not compromise economic growth and social inclusion, would require effective planning and regulation of the transport sector.

Fostering better transportation planning and policies necessitates understanding of human travel behaviours. Several factors that directly or implicitly influence human travel behaviour have been investigated in the state-of-the-art, including transportation supply [1], individual characteristics [2], trip purpose [3] and weather conditions [4]. However, more and more studies are paying attention to the influence of weather conditions on individual travel behaviour across various cities [4]–[6]. For instance Hagenauer and Helbich et al. [7] find that temperature is more important than

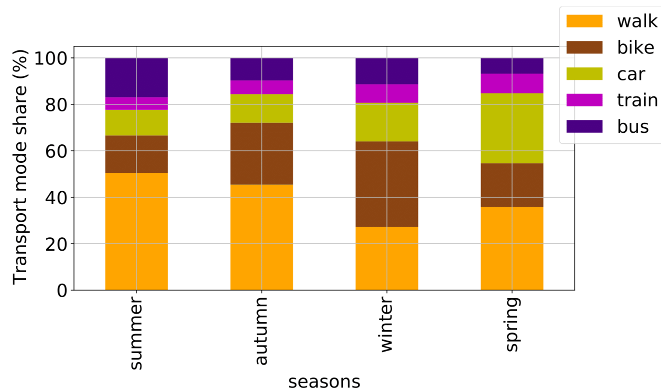


Fig. 1: Distribution of transport mode share per season using GPS trajectories in Beijing city

precipitation or wind speed in predicting bicycle and public transportation trips for Dutch travelers.

However, the authors in [8] reveal that precipitation produces more significant ridership fluctuations than temperature in Nanjing. Other than temperature and precipitation, existing literature on weather and daily mobility focus on humidity [9], wind speed [10], air quality [11] and seasons [12].

With respect to seasonality, the authors in [12] report that the car is the most chosen transport mode during bad weather for commuters in Chicago. Scholars in [13] investigate the relationship between air pollution and seasonality, and reveal that more trips are conducted during the Summer when the air quality is better than during the Winter in Taiyuan. Instead of only considering the four calendar seasons such as the Summer, Autumn, Spring, and Winter, the authors in [14] consider nine complex seasonality blocks such as the school/university period, Christmas season in influencing transport mode decisions in Brisbane. The authors reveal that there is a strong correlation between complex seasonality and bus ridership.

While studies of the effect of a calendar season on transport mode decisions has attracted the interest of some researchers, given its potential, it is surprising that this topic has not received broader attention in the literature. Moreover, the data from which these studies are conducted is based on survey or diary travel data. Findings generated from surveys are dependent on how the surveyee or participants give accurate and honest answers. In fact, Arana et al. [15] reveals that respondents who take part in these questionnaires can be subject to prejudices, memory lapses, and biases which can skew the answers. Therefore, one alternative to surveys is to take advantage of satellite positioning data which can now be recorded by information sources such as our smartphones. To the best of our knowledge, there is no study that models the seasonal effects on transport choices based on GPS trajectories.

The goal of this paper is show the possibility of modelling seasonal effects on transport mode choices in Beijing using GPS trajectories. In Fig. 1, the distribution of transport modes in each season is showed. We see that seasons have an impact on the transport mode choices. For instance, during the Summer and Autumn, walking is the most preferred transport mode while the bike and the car are most preferred means during the Winter and Spring seasons, respectively.

Therefore, this paper studies the impact of factors such as the temperature, air quality, traffic during rush hour, time of the day, trip distance, and trip duration on influencing transport mode decisions over the Summer, Autumn, Winter, and Spring seasons. The modelling is performed by the multinomial logit (MNL) technique, which is the most popular technique in developing traffic demand models. In fact, we also apply descriptive statistical analysis to identify relationships and patterns in the dataset. This paper aims to further extend our previous work in [16] in which we model the impact of weather conditions and context data on transport mode choices based on satellite positioning data.

The structure of the remainder of this paper is as follows. Section II presents the databases applied in this work from the case study city of Beijing. The statistical analysis and multinomial logistic regression model are described in Section III. The model implementation and results are in Section IV while the results are discussed in Section V. The conclusions and future work are presented in Section VI.

II. Case study

The city for our case study is Beijing in China, which is world's largest developing economy. Due to its high population density, Beijing has undergone rapid urbanization and motorization resulting in to severe congestion and air quality problems, hence making it an interesting case study for this work.

In this section, we describe the datasets from which the GPS trajectories and seasonal information are collected.

A. GPS trajectories

This dataset was collected during the Geolife Microsoft project by 182 users from April 2007 to August 2012. During the data collection, the users trajectories were recorded by GPS loggers and phones [17]. Hence, in the dataset, the GPS trajectories are described by a sequence of time-stamped points, containing longitude, latitude, altitude, and transportation mode labels.

A total of 2,754 labeled trajectories are collected specifically in Beijing. In Table I shows the description of the amount of data per transport mode in the dataset for Beijing. A total duration of 1409.36 h and a total distance of 23959 km were recorded.

Looking at Table I, it can be observed that over 100 h were recorded per transport mode and walking accounts for the majority of the hours recorded. The latter is expected given that walking is recommended as the best and most efficient commuting mode in Beijing.

TABLE I: Descriptive of the amount of data per transport mode in the dataset for Beijing

Mode	% Transport mode share	Duration (h)	Distance (km)
Walk	46.2	408.49	1890.95
Bike	20.48	238.57	2591.77
Car	14.19	331.19	11637.42
Bus	13.14	313.14	5948.30
Train	5.99	117.97	1890.95

It is worth mentioning that this dataset has been applied in different research fields such in transportation mode identification [18], user identification [19], trajectory completion [20].

B. Seasonal information

Beijing has four distinct seasons which are influenced by the temperate and continental monsoon climate. Summer lasts about 120 days from mid-May to mid-September and is quite hot and humid. Autumn season starts around mid-September to mid-November. According to [21], it is recommended as the best time to visit Beijing for the season is characterised by plenty of sunshine and cooler temperatures. The Winter season starts from mid-November to mid-March, while Spring starts mid-March to mid-May.

C. Matching Seasonal information to GPS trajectories

As mentioned earlier, the GPS trajectories from the Geolife dataset contain the location, date, time, and the transport mode chosen by the traveler.

In matching the seasonal data to the GPS trajectories, the date obtained from the GPS trajectories is used to link the trajectories to their corresponding seasons. The result generates season-related trajectories, where 1474, 679, 115, and 486 trajectories recorded in the Summer, Autumn, Winter, and Spring seasons, respectively. Interestingly, this result shows that more trips were conducted in Summer than during the other seasons, which is in line with the findings made by the authors in [13].

D. Seasonal effects of weather and context information on transport mode choices

Based on the season-related trajectories, historical weather condition data for the corresponding travel days in the different seasons are collected from a meteorological station located within the city of Beijing [22]. We specifically analyse the seasonal effects of hourly temperature and air pollution on the transport mode choices. Other weather conditions such as the hourly relative humidity, precipitation, wind speed are excluded from this study because they have a limited impact on influencing transport mode choices in Beijing as documented by our previous work in [16].

In our analysis, the weather conditions are classified in to different levels, which helps to track the weather changes. For instance Fig. 2 and Fig. 3 shows the changes in temperature and air quality are classified in different levels generated from published weather knowledge [23] and [24]. A summary of the factors with continuous observable variables is presented in Table II. Key results show that the average temperature was about 24°C, 14°C, -0.2°C, and 14°C during the Summer, Autumn, Winter, and Spring seasons for the period in which the data was collected.

TABLE II: Descriptive characteristics of continuous observable variables

	Summer	Autumn	Winter	Spring
Temperature (°C)				
Maximum	38.5	30.4	17.0	31.5
Minimum	10.6	-5.3	-9.2	-3.7
Mean	24.4	14.3	-0.2	14.1
Air quality ($\mu g/m^3$)				
Maximum	500	481	406	356
Minimum	27	13	37	24
Mean	155.4	146.4	162.4	152
Trip distance (km)				
Maximum	117.2	87.2	58.7	81.6
Minimum	0	10.86	0	1
Mean	7.9	6.7	10.6	10.1
Trip duration (min)				
Maximum	674	636	552	410
Minimum	2	4	3	1
Mean	28	31	53	29

In addition context information such as the rush hour, time of the day (day and night time), trip distance, and trip duration are also considered. The definitions of rush hour and time of day in Beijing are as follows: 7AM to 9AM and 4PM to 8PM is the morning and evening rush hours, respectively while night time is from 6PM to 6AM and day time from 6AM to 6PM.

III. Methodology

In this section we present our approach, including the statistical analysis and the multinomial logistic regression modal.

A. Statistical analysis

In this section, we apply descriptive statistical analysis, which is a common method in data analysis to analyse seasonal effects on transport mode choices. Note that descriptive analysis identifies relationships and patterns in the dataset using graphs. However, the findings cannot be used to draw conclusions about the population from which the sample data has been extracted.

Fig. 2, Fig. 3, Fig. 4, and Fig. 5 show the effects of temperature, air quality, trip duration, and trip distance on the transport modes in the Summer and Winter seasons, respectively. The graphs showing the results in the Autumn and Spring season and the results showing the effects of duration of the day, presence of traffic during rush hour are excluded because of space limitations.

During the Summer, Fig. 2a shows that as temperature increases, there is a decrease in the walking share while there is an increase in the bus share. In Fig. 2b, it shown that during Winter, whenever there is snow, the bike is the most preferred transport choice, however as the temperature increases, the bike share decreases and there is an increase in the bus share.

In Fig. 3a, we observe that as the air quality changes from good (0-100) µg/m3 to terrible to (201-500) µg/m3, there is a limited effect on the transport mode during the Summer. However, during the Winter, Fig. 3b shows that as the air quality turns from good to terrible, there is an increase in the walking share and a decrease in the car share. Perhaps a reason for the reduction in the usage of private cars could as a result of vehicle controls introduced by the government to reduce $C0_2$ emissions.

Fig. 4 shows how the trip duration influence our transport mode decisions. For instance for any trip duration less than 20 min, walking and biking have the largest shares while for trip durations beyond 20 mins, car, bus, train are the more preferred transport mode choices. In fact, we see a similar trend during the Summer and Winter seasons as showed in Fig. 4a and Fig. 4b, respectively.

The seasonal effects of trip distance on transport mode choices for the Summer and Winter seasons are shown in Fig. 5. We see that trip distances have a tremendous effect on the transport modes. This trend is similar in both Summer and Winter seasons. For instance, in both seasons, for short distance e.g. less than 6 km, walking and biking have the highest share, but for trip distances beyond 6 km, either bus, car, and train have the highest share, which is in line with the findings in [25].

B. Multinomial logistic regression model

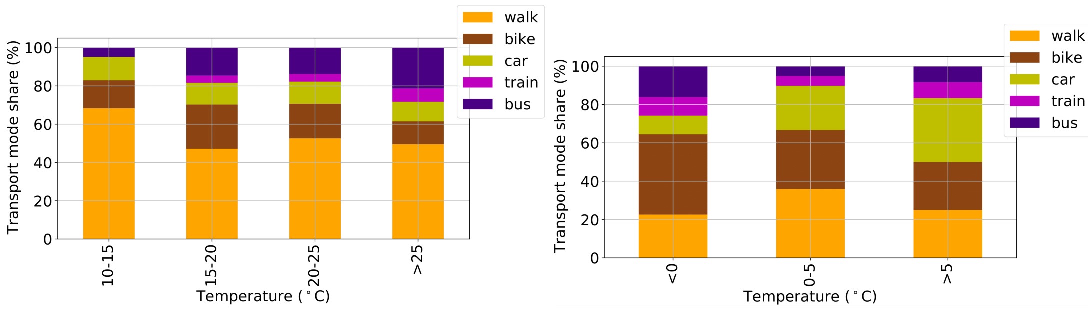
The MNL model is one of the most well-known statistical models for analysing the relationship between two or more variables. In fact, in analysing and predicting travel mode choices among mutual alternatives, it is the most widely used model.

In deriving this model, the decision rule, from which the decision maker evaluates the alternatives is based on maximizing the the Utility function, U_{in}, which is defined as follows:

$$U_{in} = V_{in} + \epsilon_{in}. \tag{1}$$

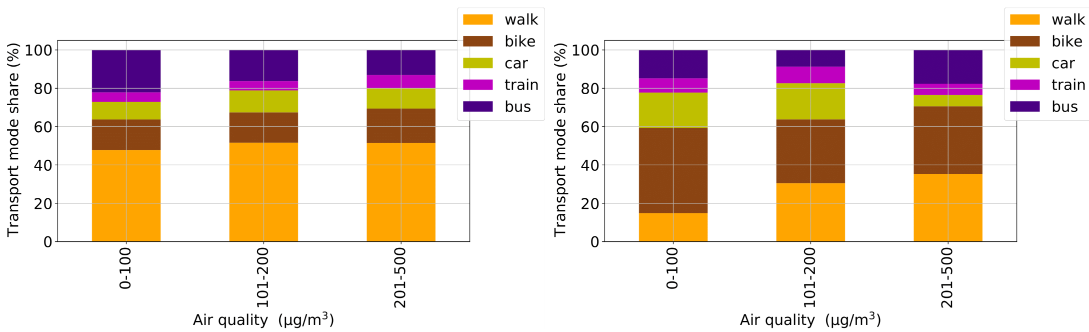
The probability that a transport port mode i is chosen by a traveler n from a set of transport mode choices $C_n = \{1, 2, ..., i, ..., j_n\}$ is given by:

$$\begin{aligned} P_{in} &= P[V_{in} + \epsilon_{in} \geqslant V_{jn} + \epsilon_{jn}, \forall j \in C_n] \\ &= P[(V_{in} + \epsilon_{in}) = max_{j \in C_n}(V_{jn} + \epsilon_{jn})] \\ &= P[\epsilon_{jn} - \epsilon_{in} \leq V_{in} - V_{jn}, \forall j \in C_n]. \end{aligned} \tag{2}$$



(a) (b)

Fig. 2: Seasonal effects of temperature on the transport mode choices : (a) Summer and (b) Winter

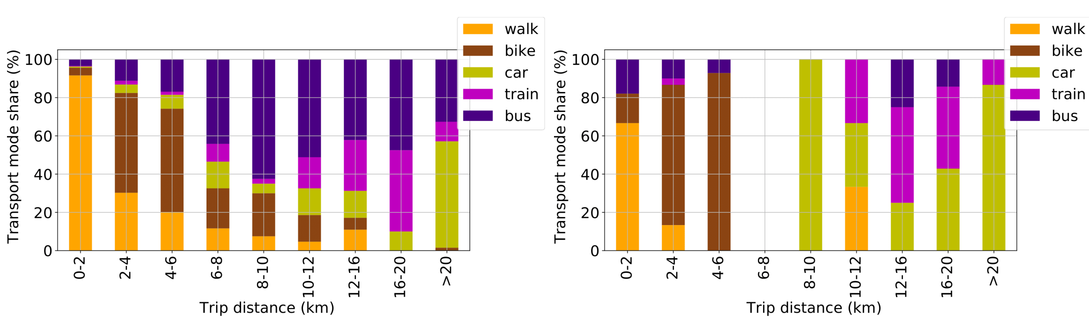


(a) (b)

Fig. 3: Seasonal effects of air quality on the transport mode choices : (a) Summer and (b) Winter. The levels 0-100 µg/m3, 101-200 µg/m3, and 201-500 µg/m3 mean that the air quality categorization is good, poor, and terrible, respectively.

(a) (b)

Fig. 4: Seasonal effects of trip duration on the transport mode choices : (a) Summer and (b) Winter



(a) (b)

Fig. 5: Seasonal effects of trip distance on the transport mode choices : (a) Summer and (b) Winter

where V_{in} is given in (3) by a vector of observable variables z_n and their coefficients γ_i:

$$V_{in} = \gamma_i^T z_n. \quad (3)$$

ϵ_{in} is the random term which expresses the errors of the utility function. The error is independent and identically distributed and takes on a Gumbel distribution. P_{in} is then defined as:

$$P_{in} = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}. \quad (4)$$

However, coefficients of the MNL are not intuitive. Therefore, marginal effects are often computed in (5) to predict the amount of change in the transport modes shares due to one unit change of the observatory variable in the model system.

$$\frac{\partial P_{in}}{\partial z_n} = P_{in}(\gamma_i - \overline{\gamma}_i). \quad (5)$$

IV. Model implementation and results

Four MNL models are developed on the basis of the data collected in the four seasons. The main objective is to quantify the magnitude by which these observatory variables: temperature, air quality, traffic during rush hour, time of the day, trip distance, and trip duration influence transport mode decisions over the Summer, Autumn, Winter, and Spring seasons.

The four MNL models are assigned the same observatory variables to compare the results. The observatory variables are considered as categorical except for the temperature, which has been considered as a continuous variable.

The categorization is according to the definitions of the historical weather and context information in Section II D. In fact, the categories for the trip duration and trip distances are developed based on relationships and patterns observed in Fig. 4 and Fig. 5. For instance, a short duration trip is defined as the trip whose duration is shorter than 20 min, otherwise it is regarded as a long duration trip. Similarly, a short distance trip is one where the distance traveled is shorter than 6 km, otherwise it is regarded as a long distance trip. This categorization is derived from the observation that short distance and short duration trips are mainly accomplished by walking or biking.

Five modes of transports whose statistics are presented in Table I are considered as the dependent variable. The modelling is done according to equations (1) to (5) and is implemented via the software package Stata [26].

V. Discussion

Table III reports the marginal effects for each variable during the four calendar seasons on transport mode choices. In general, the positive values show an increase in the probability of choosing a given transport mode while the negative indicate a decrease.

TABLE III: Marginal effects of the influence of seasons on transport mode choices.

observatory variables	Transport modes				
	Walk	Bike	Car	Bus	Train
Summer					
Temperature (°C)	-0.001	-0.010^{**}	-0.001	0.011^{**}	0.002^{**}
Air quality ($\mu g/m^3$)					
$0-100$ (good)	ref	ref	ref	ref	ref
$101-200$ (poor)	0.015	-0.016	0.025	-0.027	0.003
$201-500$ (terrible)	0.039	0.040	0.005	-0.085^{*}	0.0001
Long distance trip (km) (*ref. short distance trip*)	-0.631^{**}	-0.158^{**}	0.314^{**}	0.241^{*}	0.233^{**}
Long duration trip (min) (*ref. short duration trip*)	-0.162^{**}	0.097^{**}	-0.0004	0.087^{*}	-0.021^{**}
Rush hour (*ref. non rush hour*)	0.041	0.004	-0.027	-0.006	-0.012
Day (*ref. night*)	-0.069^{*}	0.021	-0.0009	0.056^{*}	-0.006
Autumn					
Temperature (°C)	0.006	-0.011^{**}	-0.0004	0.006^{**}	-0.0003
Air quality ($\mu g/m^3$)					
$0-100$ (good)	ref	ref	ref	ref	ref
$101-200$ (poor)	-0.072	0.091^{*}	-0.006	-0.018	0.006
$201-500$ (terrible)	-0.084	0.164^{**}	-0.027	-0.040	-0.012
Long distance trip (km) (*ref. short distance trip*)	-0.541^{**}	-0.083	0.377^{**}	0.111^{**}	0.135^{**}
Long duration trip (min) (*ref. short duration trip*)	0.126^{**}	-0.137^{**}	-0.027	0.039	-0.002
Rush hour (*ref. non rush hour*)	-0.054	0.126^{**}	-0.047^{**}	-0.002	-0.022
Day (*ref. night*)	0.106^{*}	-0.192^{**}	0.025	0.059^{*}	0.002
Winter					
Temperature (°C)	0.025	-0.0001	0.0001	-0.025	-0.0003
Air quality ($\mu g/m^3$)					
$0-100$ (good)	ref	ref	ref	ref	ref
$101-200$ (poor)	0.221	-0.006	-0.0069	-0.213	-0.002
$201-500$ (terrible)	0.110	-0.006	-0.0093	-0.102	-0.001
Long distance trip (km) (*ref. short distance trip*)	-0.371^{**}	-0.572^{**}	0.999^{**}	0.056	0.0005
Long duration trip (min) (*ref. short duration trip*)	0.268	0.124	-0.003	-0.027	-0.221
Rush hour (*ref. non rush hour*)	0.574	-0.011	-0.0005	-0.593	-0.0011
Day (*ref. night*)	0.574	0.004	-0.0001	-0.577	0.002
Spring					
Temperature (°C)	0.001	-0.005	0.010^{**}	-0.042	-0.0001
Air quality ($\mu g/m^3$)					
$0-100$ (good)	ref	ref	ref	ref	ref
$101-200$ (poor)	-0.055	0.022	-0.006	0.036	-0.002
$201-500$ (terrible)	-0.086	0.029	-0.102	-0.023	-0.011
Long distance trip (km) (*ref. short distance trip*)	-0.500^{**}	-0.392^{**}	0.796^{**}	-0.019	0.115^{**}
Long duration trip (min) (*ref. short duration trip*)	-0.027	0.300^{**}	-0.354^{*}	0.133^{*}	-0.051
Rush hour (*ref. non rush hour*)	0.026	0.227^{**}	-0.150^{**}	-0.593	-0.089^{**}
Day (*ref. night*)	-0.075	-0.037	-0.007	0.011	0.002

$**$ Significant at $\alpha < 0.01$ $\quad$ $*$ Significant at $\alpha < 0.05$

A. Summer

During the Summer, it noticeable that as the temperature increases, the probabilities of biking is expected to decrease by 10% while the probabilities of choosing the bus and train are expected to increase by 11% and 2%, respectively. This result is in line with the findings by the authors in [27] and [7] where too high temperature is found to favour motorised transport over biking. Note that the reference for the temperature is the average temperature for the corresponding season [see Table II].

When the air quality becomes terrible, we observe that there is a 8.5% decrease in probability of choosing the bus perhaps due to vehicle controls put in place by the government to reduce C02 emissions.

Long distance trips decrease the probability of walking and biking by 63.1% and 15.8%, respectively, while increase the car, bus, and train use by 31.4%, 24.1%, and 23.3%.

Long duration trips are expected to decrease the probability of walking and taking the train by 16.2% and 2.1%, respectively, while increase the probability of taking the bike and bus by 9.7% and 8.7%, respectively.

During the day time, we see a reduction in the likelihood of walking by 6.9% and an increase in the use of the bus by 5.6%. A possible reason is that during the day there are more buses operating than at night.

B. Autumn

Similar to the Summer season, when temperature increases, the probabilities of biking is expected to decrease by 11% while the probabilities of choosing the bus is expected to increase by 6%.

We also see that when the air quality turns to poor and terrible categories, the probabilities of choosing a more environmentally friendly travel choice such as the bike increases by 9.1% and 16.4%, respectively.

Long distance trips decrease the probability of walking by 54.1% and increase the use of the car, bus, and train by 37.7%, 11.1%, and 13.5%, respectively.

Long duration trips increase the probability of walking by 12.6% while decrease the probability of biking by 13.7%.

During the day time, we see an increase in the likelihood of walking and the use of the bus by 10.6% and 5.9%, respectively, while there is a decrease of 19.2% in biking.

Unlike the Summer were rush hour traffic has no impact on the choice of transport mode, during Autumn, a 4.7% decrease in the use of the car and a 12.6% increase in the use of the bike are expected.

C. Winter

In contrast to the Summer, Autumn and Spring seasons were a number of factors influencing the transport mode are observed, during the Winter, only the trip distance is found to be significant. This result is also inconsistent with the patterns observed in Fig. 2b, Fig. 3b, Fig. 4b and could be explained by the few number of trajectories which are recorded during the Winter from which the modelling has been performed.

However, the model shows that during Winter, long distance trips decrease the probability of walking and biking by 37.1% and 57.2%, respectively, while increasing the probability of choosing the car by 99%.

D. Spring

During this season, the model predicts a 10% increment in the car use when the temperature increases.

Long distance trips are expected to decrease walking ad biking by 50% and 39.2%, respectively while increasing the use of the car and train by 79.6% and 11.5%, respectively.

Long duration trips are expected to increase the use of the bike and bus by 30% and 13.3%, respectively while decrease the use of the car by 35.4%.

Rush hour increases the probability of using the bike by 22.7% while decreasing the probability of car and train use by 15% and 8.9%, respectively. The result could possibly be explained by the need to reduce or avoid congestion i.e. traffic jams and congestion of people in trains.

VI. Conclusions

This study looked at seasonal factors that might affect the transport mode choice in the city of Beijing, China. We studied the effect of the following factors: temperature, air quality, traffic during rush hour, time of the day, trip distance, and trip duration.

Our results show firstly that seasons impacted transport mode choices. Secondly, there were noticeable differences on how factors influenced the transport modes depending on the season. For instance, temperature, air quality, and time of day mostly influenced the choice of transport modes during the Summer and Autumn seasons.

Additionally, the traffic during rush hour only influenced the choice of transport mode during the Autumn and Spring seasons. Some similarities among the factors influencing the transport mode choices were also noticeable among the seasons. For instance, trip distance was found to impact the transport mode in all seasons, while the trip duration influenced the transport modes in all seasons except for the Winter season.

There were limitations in this study, which introduced a bias in the results. Particularly, the reduced number of GPS trajectories recorded during Winter.

References

[1] Ö. Şimşekoğlu, T. Nordfjærn, and T. Rundmo, "The role of attitudes, transport priorities, and car use habit for travel mode use and intentions to use public transportation in an urban norwegian public," *Transport Policy*, vol. 42, pp. 113–120, 2015.

[2] M. Dixit and A. Sivakumar, "Capturing the impact of individual characteristics on transport accessibility and equity analysis," *Transportation Research. Part D, Transport and Environment*, vol. 87, 2020.

[3] L. A. Guzman, D. Oviedo, and C. M. R. Rivera, "Assessing equity in transport accessibility to work and study: The bogotá region," *Journal of Transport Geography*, vol. 58, pp. 236–246, 2017.

[4] F. Ahmed, G. Rose, M. A. Figliozzi, and C. Jakob, "Commuter cyclist's sensitivity to changes in weather: Insight from two cities with different climatic conditions," 2012.

[5] L. Böcker, J. Prillwitz, and M. Dijst, "Climate change impacts on mode choices and travelled distances: a comparison of present with 2050 weather conditions for the Randstad Holland," *Journal of Transport Geography*, vol. 28, no. C, pp. 176–185, 2013. [Online]. Available: https://ideas.repec.org/a/eee/jotrge/v28y2013icp176-185.html

[6] L. Böcker, M. Dijst, and J. Prillwitz, "Impact of everyday weather on individual daily travel behaviours in perspective: A literature review," *Transport Reviews*, vol. 33, pp. 71 – 91, 2013.

[7] J. Hagenauer and M. Helbich, "A comparative study of machine learning classifiers for modeling travel mode choice," *Expert Syst. Appl.*, vol. 78, pp. 273–282, 2017.

[8] J. Li, X. Li, D. wei Chen, and L. Godding, "Assessment of metro ridership fluctuation caused by weather conditions in asian context: Using archived weather and ridership data in nanjing," *Journal of Transport Geography*, vol. 66, pp. 356–368, 2017.

[9] L. Li, J. Wang, Z. Song, Z. Dong, and B. J. Wu, "Analysing the impact of weather on bus ridership using smart card data," *Iet Intelligent Transport Systems*, vol. 9, pp. 221–229, 2015.

[10] X. Li, Q. Ma, W. Wang, and B. Wang, "Influence of weather conditions on the intercity travel mode choice: A case of xi'an," *Computational Intelligence and Neuroscience*, vol. 2021, 2021.

[11] L. Ma, H. Xiong, Z. Wang, and K. Xie, "Impact of weather conditions on middle school students' commute mode choices: Empirical findings from beijing, china," *Transportation Research Part D: Transport and Environment*, vol. 68, pp. 39–51, mar 2019.

[12] M. F. Hyland, C. Frei, A. Frei, and H. S. Mahmassani, "Riders on the storm: Exploring weather and seasonality effects on commute mode choice in chicago," *Travel Behaviour and Society*, 2018.

[13] W. Li and M. Kamargianni, "Air pollution and seasonality effects on mode choice in china," *Transportation Research Record*, vol. 2634, pp. 101 – 109, 2017.

[14] S. A. Kashfi, J. M. Bunker, and T. Yigitcanlar, "Understanding the effects of complex seasonality on suburban daily transit ridership," *Science & Engineering Faculty*, 2015.

[15] P. Arana, S. Cabezudo, and M. Peñalba, "Influence of weather conditions on transit ridership: A statistical study using data from smartcards," *Transportation Research Part A-policy and Practice*, vol. 59, pp. 1–12, 2014.

[16] "Modelling the impact of weather and context data on transport mode choices: A case study of gps trajectories from beijing." preprint on webpage at https://www.preprints.org/manuscript/202204.0131/v1.

[17] "GeoLife: Building Social Networks Using Human Location History," https://www.microsoft.com/en-us/research/project/geolife, accessed: 2021-09-10.

[18] C. Markos and J. J. Yu, "Unsupervised deep learning for gps-based transportation mode identification," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 2020, pp. 1–6.

[19] M. Qi, Z. Wang, Z. He, and Z. Shao, "User identification across asynchronous mobility trajectories," *Sensors*, vol. 19, no. 9, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/9/2102

[20] A. Nawaz, Z. Huang, S. Wang, A. Akbar, H. AlSalman, and A. Gumaei, "Gps trajectory completion using end-to-end bidirectional convolutional recurrent encoder-decoder architecture with attention mechanism," *Sensors*, vol. 20, no. 18, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/18/5143

[21] "StataCorp, Stata base reference manual," https://www.chinadiscovery.com/beijing-tours/weather-seasons.html, accessed: 2021-12-12.

[22] "POWER Data Access Viewe," https://power.larc.nasa.gov/data-access-viewer/, accessed: 2021-11-12.

[23] C. Cavallaro and J. Vitrià, "Corridor detection from large gps trajectories datasets," *Applied Sciences*, vol. 10, no. 14, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/14/5003

[24] L. Aultman-Hall, D. Lane, and R. R. Lambert, "Assessing impact of weather and season on pedestrian traffic volumes," *Transportation Research Record*, vol. 2140, no. 1, pp. 35–43, 2009. [Online]. Available: https://doi.org/10.3141/2140-04

[25] T. H. de Sá, D. C. Parra, and C. A. Monteiro, "Impact of travel mode shift and trip distance on active and non-active transportation in the são paulo metropolitan area in brazil," *Preventive Medicine Reports*, vol. 2, pp. 183 – 188, 2015.

[26] "StataCorp, Stata base reference manual," https://www.stata.com/manuals/r.pdf, accessed: 2021-12-12.

[27] L. Creemers, G. Wets, and M. Cools, "Meteorological variation in daily travel behaviour: evidence from revealed preference data from the netherlands," *Theoretical and Applied Climatology*, vol. 120, pp. 183–194, 2014.