

26th Euro Working Group on Transportation Meeting (EWGT 2024)

Non-routine travel in multi-day trip diary data

Miriam Magdolen^{a*}, Peter Vortisch^a

^a*Institute for Transport Studies Karlsruhe Institute of Technology (KIT), Kaiserstr. 12, 76131 Karlsruhe, Germany*

Abstract

This study addresses the need for more knowledge in long-distance travel behaviour by developing a method to identify non-routine travel behaviour in multi-day travel surveys. Following the definition of tourism, this study uses trip diary data to approximate an individual's usual environment and identify travel events leaving this usual environment as non-routine travel. In contrast to commonly used definitions of long-distance travel, e.g. based on a distance criterion, this study provides an approach that identifies non-routine travel on an individual level depending on the reported behaviour. After calculating and comparing ten indicators to measure the size of the usual environment based on tour distances reported in three weeks, one indicator considering the interquartile range is selected. A following analysis shows that the developed approach performs equally sufficient if only two weeks of trip diaries are available. Findings reveal that, on average, about 9% of all reported tours of a sample from a large-scale German household travel survey are non-routine travel.

© 2024 The Authors. Published by ELSEVIER B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 26th Euro Working Group on Transportation Meeting

Keywords: Non-routine travel; trip diary; usual environment; long-distance travel;

1. Introduction and background

Studying long-distance travel as a relevant component of overall travel behaviour is crucial. In recent years, two aspects have become focuses in travel behaviour research that apply to long-distance travel specifically: Promoting more sustainable and equitable travel. Many studies have demonstrated that long-distance travel accounts for a large part of travel-related emissions (Mattioli, 2023). This is, on the one hand, due to the long distances travelled within only a few travel events. For example, in Germany, although less than 2% of all trips exceed a one-way distance of 100 km, they account for about 46% of the kilometres travelled (Schulz et al., 2024). On the other hand, air travel

* Corresponding author. Tel.: +49-721-608-47735.

E-mail address: miriam.magdolen@kit.edu

plays a significant role in long-distance travel, which additionally involves high emissions. The second relevant aspect is equity in transport. Long-distance travel demand is unequally distributed in the population. While many people never or seldom travel to distant places, there is a small group of frequent travellers who are responsible for a large part of travel. Factors such as education, income, gender and place of residence influence long-distance travel demand (Czepkiewicz et al., 2018). Considering these two relevant aspects, sustainability and equity, long-distance travel should receive more attention in both research and political discussions.

One of the main challenges lies in the absence of a standardised and uniform definition of long-distance travel, with many studies employing their own definitions. Most of them are distance-based, such as a one-way distance of at least 100 km, while others focus on specific travel purposes, such as holiday travel, or the duration of stay, such as one or more overnights (Mattioli, 2023). Despite this variety of definitions, there is an underlying idea shared by these definitions: Focusing on travel events that are extraordinary in an individual's behaviour. Thus, researchers try to analyse the part of travel that is non-habitual or non-routine and in which travel decisions are different from everyday decisions regarding destination choice, mode choice, or route choice. In this context, the definition of 'tourism' by the United Nations World Tourism Organization (UNWTO) becomes pertinent. Tourism is defined as the movement of people outside their individual usual environment for private or business purposes, except for commuting (World Tourism Organization, 2007). There is no minimum distance or duration. The usual environment refers to the geographical area where an individual's routines occur. Thus, tourism encompasses non-routine travel beyond an individual's routine daily life. For each individual, the size and shape of the usual environment are different.

Applying the definition of tourism to travel data faces the challenge of information about routines seldom being collected. However, extended reporting periods allow for identifying behavioural rhythms and routines in daily life (Axhausen et al., 2002). Different from other approaches that identify tourists at destinations, e.g. based on social media data (Hasnat & Hasan, 2018; Mor et al., 2023), this study presents an approach to identify non-routine travel in longitudinal data from national household travel surveys. The aim is not to quantify the volume of tourism in a specific area, but to identify what proportion of the travel reported by an individual is non-routine. For this, the information on individuals' multi-day travel behaviour is analysed to distinguish between routine and non-routine travel.

This study addresses the following research questions: First, how can the definition of tourism be operationalized to approximate the usual environment of individuals and identify non-routine travel in longitudinal trip diary data? Second, what observation period is required to allow for the differentiation between routine and non-routine travel?

In the following, the data used for the analysis are presented, and different statistical methods that approximate the usual environment of each respondent on the individual level are developed. A comparison of the methods and the selection of the most promising approach follows. A second analysis examines how long the observation period must be to distinguish between the routines and non-routines in individual behaviour. This is followed by a short description of the characteristics of the identified non-routine travel. The paper ends with a conclusion and an outlook for further research.

2. Data and methods

This study is based on data from the German Mobility Panel (MOP). The MOP is an annual household travel survey conducted between 1994 and 2023 and collects data from about 2,000 to 3,000 respondents each year (Vallée et al., 2022). It is a rotating panel, meaning that some respondents leave the panel each year, and others enter it. The respondents are asked to participate in up to three survey waves. The survey collects data on everyday mobility with a one-week trip diary. The MOP focuses on capturing everyday travel and the survey therefore is conducted outside from school breaks. To reduce the influence of weather conditions, the survey takes place at the same time each year in autumn. It is essential to highlight that the MOP does not collect data on the locations of destinations in the trip diary for privacy reasons. Thus, the places visited during the weeks of the report cannot be tracked on a map. However, the MOP is still the only large-scale travel survey in Germany that collects longitudinal data, which allows for identifying routines in behaviour. If the respondents participate in all three years, three weeks of travel diaries are available to identify the routines, e.g. typical trip chains, typical distances and durations and typical mode use. This is used as a basis for determining a method to identify the usual environment of each individual and to differentiate between routine and non-routine travel behaviour. While a previous study focused on demonstrating the potential of identifying non-routine behavioural patterns in three reporting weeks (Magdolen et al., 2020), this paper focuses on

selecting the best method to do so. For this purpose, a much larger sample is analysed and more advanced approaches are tested. The following steps are conducted:

- 1) Preparing survey data and filtering respondents that reported three weeks and had no significant changes in their circumstances, e.g. residential relocation, which would impact the daily routines between the reports
- 2) Building of tours (starting from home and ending at home), e.g. to determine activity sequences, tour distances, tour durations, etc.
- 3) Developing and calculating indicators to determine an individual threshold for the size of the usual environment and for the identification of non-routine tours that exceed this individual threshold
- 4) Comparing the indicators regarding
 - the distribution of thresholds to define the usual environment
 - the volume of identified non-routine tours
 - the sensitivity for low-mobile and high-mobile respondents
- 5) Selection of the indicator that works as the best approximation to define the usual environment and to identify non-routine tours in longitudinal behaviour

Ten indicators are developed based on common statistical measures for central tendency and dispersion that are often used for outlier detection (Chikodili et al., 2021; Leys et al., 2013; Rousseeuw & Hubert, 2011; Zijlstra et al., 2007). The idea is that the threshold for outlier detection can be transferred to the threshold for the size of the usual environment. Thus, the indicators are applied to the tour distances of all reported tours by an individual that start and end at the home location. By purpose, no socio-demographic characteristics or other assumptions are included in the methodology in order to avoid generalizations. The determination of the criterion should be based solely on the reported travel behaviour. Two indicators each are based on the mean (1.5x mean, 2.0x mean) and the median (1.5x median, 2.0x median) of the tour distances. Further, two indicators are based on the z-score (z-score 2SD, z-score 3SD). These two indicators are calculated as the mean tour distance plus 2.0 times, respectively 3.0 times, the standard deviation (Rousseeuw & Hubert, 2011). Another two indicators that are useful for outlier detection are calculated as the 3.0 and 5.0 times median deviation from the median (Med+3MAD, Med+5MAD) (Leys et al., 2013). Lastly, the interquartile range ($Q3+1.0 \times IQR$, $Q3+1.5 \times IQR$) is considered. This outlier detection method is also known as ‘Tukey’s fences’ (Zijlstra et al., 2007). For this, the third quartile of the tour distances is calculated and added by 1.0 or 1.5 times the difference between the first and third quartile. All these statistical measures are calculated on an individual level based on the individually reported travel behaviour in the three weeks of trip diaries.

3. Preparation of the study sample

Table 1 provides the sociodemographic characteristics of the sample after checking if the respondents reported three trip diaries and for stable circumstances (step 1). The relevant characteristics are residential relocation, work relocation, change in occupation (student, full- or part-time job, retired, (currently) not working), change in car ownership (no car vs. at least one car in household) and driver’s license. If there is a change in these characteristics, impacts on travel behaviour are likely. Thus, the trip diaries from the three years are not useful to define the usual environment of this individual. Further, the number of days a respondent reported illness or holidays is checked. Respondents stay in the sample if such peculiarities influence a maximum of five days during the three weeks.

After filtering the survey years 2011 to 2019 (before COVID-19) for respondents reporting three weeks of trip diaries and excluding those with significant changes impacting their travel routines, a sample of 2,437 respondents remained. The comparison of the sociodemographic characteristics of the sample with the overall German population in Table 1 reveals some differences. Regarding the household characteristics, the remaining sample has a lower number of 1-person-households and a lower number of households with no cars. However, the most striking difference can be seen within the distribution of age groups. Especially the three youngest age groups are underrepresented in the sample. The main reason for this lies in the checks applied during the first step of data preparation. With the MOP as the data source, the three weeks of trip diaries are inevitably from three different years. However, the younger age groups are likely to experience essential changes between three years. Changing from being a student to working or moving out of the parent’s house is likely to happen between 18 and 25 years. This is reflected in the sample by the small number of respondents from the young age groups and needs to be considered when interpreting the results.

Table 1. Sociodemographic characteristics

		Survey sample		Official statistics*
		n	in %	in %
<i>Household level</i>		1,820		
Household size	1-person-household	548	30.1%	41.8%
	2-person-household	860	47.3%	33.5%
	3-person-household	212	11.6%	12.0%
	>3- person-household	200	11.0%	12.6%
Municipality size of residence (number of inhabitants)	Less than 20,000	703	38.6%	37.8%
	20,000 – 50,000	347	19.1%	18.2%
	50,000 – 100,000	183	10.1%	9.2%
	100,000 – 500,000	283	15.5%	16.3%
	500,000 and more	304	16.7%	18.5%
Car ownership	No car in household	224	12.31%	22.60%
	>= 1 car in household	1,596	87.69%	77.40%
<i>Person level</i>		2,437		
Gender	Female	1,153	47.3%	50.9%
	Male	1,284	52.7%	49.1%
Age groups	10 to 17 years	88	3.6%	6.9%
	18 to 25 years	31	1.3%	9.5%
	26 to 35 years	83	3.4%	14.1%
	36 to 50 years	446	18.3%	21.1%
	51 to 60 years	616	25.3%	17.8%
	61 to 70 years	603	24.7%	13.4%
	> 70 years	570	23.4%	17.3%

*The official statistics are based on the microcensus 2017 (DESTATIS - Statistisches Bundesamt, 2018) and the sample survey of income and expenditure 2018 (DESTATIS - Statistisches Bundesamt, 2021)

4. Selection of the indicator describing the size of the usual environment

The remaining respondents reported more than 175,000 trips, aggregating to 77,222 tours (step 2). By calculating the ten described indicators, a threshold for the usual environment for each person and each indicator was found in step 3. Figure 1 shows the threshold distribution for the size of the usual environment of each person.

On the left side, the indicators with smaller factors that result in smaller thresholds are shown. On the right side, the higher factor distributions are shown leading to higher tour distance thresholds and, thus, larger usual environments. Overall, the figure illustrates a high variation in the thresholds among the respondents. Some of the indicators seem to underestimate the usual environment of individuals with low thresholds, e.g. 1.5x median, and others likely overestimate the size of the usual environment, e.g. z-score 3SD, with many respondents whose usual environment has a tour distance greater than 350 km.

All tours with distances exceeding the individual threshold for the usual environment are labelled as non-routine tours. All other tours are labelled as routine tours of the individual. Table 2 provides an overview of the identified non-routine tours for each of the ten indicators. The mean values are given, but it should be emphasised that all values were calculated at an individual level and can, therefore, deviate considerably between respondents. In line with the thresholds, the share of identified non-routine tours varies strongly between 3.6 % and 28.9 %.

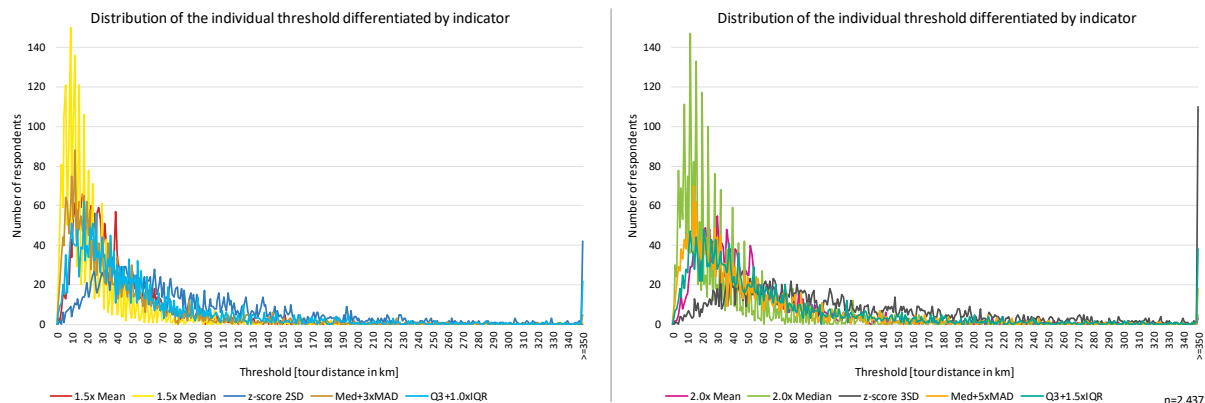


Fig. 1. Distribution of thresholds for the size of the individual usual environment – differentiated by each indicator with lower factors on the left side and higher factors on the right side

Table 2 also provides information on the mean durations and mean distances for non-routine and routine tours. The results show that the two types of tours differ strongly. In addition, the indicators differ regarding the distribution of non-routine tours among the respondents. The percentage of respondents without non-routine tours in three weeks of trip diaries varies between 1.1 % and 32.9 %, depending on the indicator.

Table 2. Comparison of the applied indicators

	Method for identification									
	1.5 x Mean	2.0 x Mean	1.5 x Median	2.0 x Median	z-score 2SD	z-score 3SD	Med + 3xMAD	Med + 5xMAD	Q3+ 1.0xIQR	Q3+ 1.5xIQR
Total number of NRT	15,226	9,667	22,352	17,167	4,712	2,784	14,135	10,736	8,716	7,285
Mean threshold (tour distance in km)	39.20	52.27	26.95	35.94	92.64	125.89	40.16	54.95	56.21	68.35
% NRT of all tours	19.7%	12.5%	28.9%	22.2%	6.1%	3.6%	18.3%	13.9%	11.3%	9.4%
Mean distance of NRT	95.5	122.2	68.9	82.2	179.9	235.2	91.5	107.7	120.8	134.3
Mean duration of NRT	380.7	378.2	313.2	328.2	402.3	425.0	336.3	347.6	349.6	358.1
Mean distance of routine tours	12.5	15.5	12.6	13.6	19.1	21.2	14.9	16.2	17.2	17.9
Mean duration of routine tours	193.1	208.9	196.3	202.1	218.9	222.8	206.3	211.1	214.9	216.0
% of respondents without NRT	1.1%	6.7%	4.5%	8.5%	11.6%	32.9%	8.1%	15.0%	13.5%	20.2%
Mean number of NRT per respondent	6.3	4.0	9.2	7.0	1.9	1.1	5.8	4.4	3.6	3.0

NRT = non-routine tours

Based on this comparison, it is difficult to directly select the best indicator. However, it can be concluded that some of the indicators perform better than others. The indicators 1.5x median, 2.0x median, 1.5x mean, and Med+3MAD seem to underestimate the size of the usual environment and thus overestimate the non-routine travel events compared to the other indicators. In contrast, both indicators based on the z-score may be too conservative for the identification of non-routine tours. An additional check is carried out to determine whether the indicators react sensitively to particularly low-mobile or high-mobile respondents. It turns out that the indicators 1.5x Mean, 2.0x Mean, Q3+1.0IQR, and Q3+1.5IQR are least sensitive to the number of tours available for each individual.

A final comparison is visualized in Fig. 2. For this, the coefficient of variation (CV) is used, defined as the ratio of the standard deviation to the mean. The idea is that by labelling the non-routine tours, the CV of the distances of the remaining routine tours should decrease. The objective is to find an indicator that considerably reduces the CV for the routine tours but, at the same time, does not overestimate the number of non-routine tours.

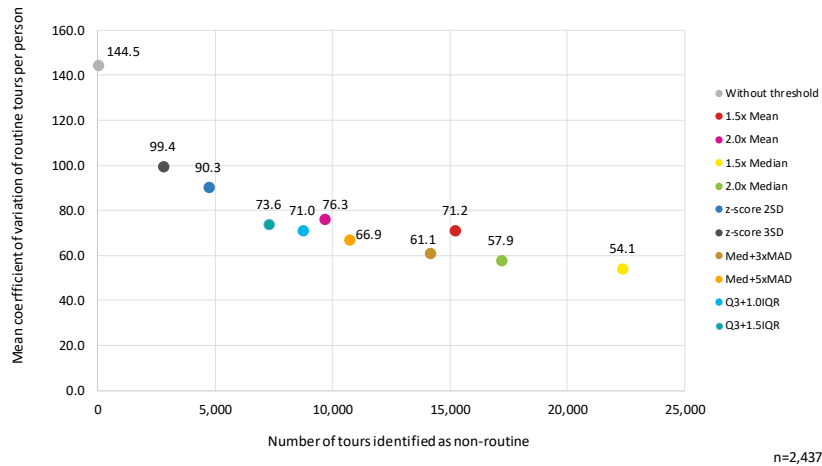


Fig. 2. Comparison of the number of identified non-routine tours and the remaining coefficient of variation of the distances in routine tours

Fig. 2 shows that the CV of the distances of all tours is across all respondents on average 144.5. The method 1.5x median reduces the CV of the remaining routine tours the most; however, it also labels most tours as non-routine. Hence, the ratio between the two aspects is not sufficient. By actually calculating the ratio between the reduction of the CV and the number of tours labelled as non-routine, Q3+1.5IQR is revealed as the best-performing indicator with a remaining CV of 73.6. After evaluating all aspects described, the indicator Q3+1.5xIQR is selected as the most promising approach to define the usual environment and to differentiate between routine and non-routine travel behaviour. On this basis, the first research question of how to identify an individual's usual environment in multi-day trip diary data can be answered: By applying the Q3+1.5IQR method on the reported tour distances to calculate the individual threshold and by labelling reported travel behaviour that exceeds this individual distance threshold. In this way, an average of 9 % of the tours are identified as non-routine travel in the present data.

5. Application to shorter observation periods

Based on the idea that the longer the observation period, the better is the identification of the size of the usual environment, the selection of the most promising approach was based on three weeks of reports. In this section, it is evaluated how well the chosen approach works for shorter observation periods. As described above, filtering for people with the same circumstances in all three years results in a comparatively small sample, especially in the young age groups. This could be mitigated to some extent if a shorter observation period is sufficient to identify non-routine travel, and in the case of the MOP, only two instead of three survey years need to be without changes.

To check the performance for shorter observation periods, the method developed is applied on one day, two days, three days, etc. and the difference between each threshold of a day and the previous day is calculated for each person. Fig. 3 shows the mean differences and the mean absolute differences between the thresholds of each day and the previous day. It can be seen that after about 6 days, the mean differences between the thresholds is close to 0. However, the mean absolute difference is still around 10 km. After about 14 days the mean absolute difference is less than 5 km and seems to converge slowly, which indicates the answer to the second research question about the minimum observation period. The results indicate for the MOP that not only persons with three weeks of report, but also with two weeks of trip diaries are suitable for the determining the individual threshold as an approximation for the individual usual environment.

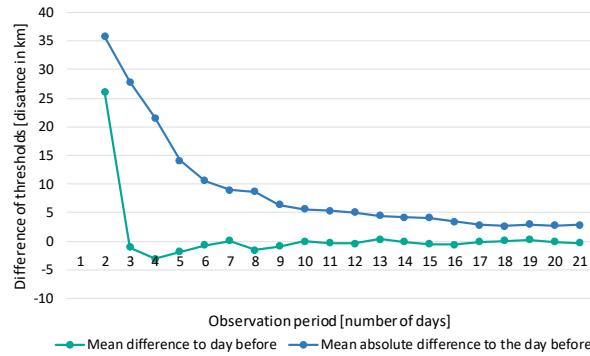


Fig. 3 Changes in thresholds depending on the length of the observation period

In a next step, the differences between the threshold value determined on the basis of two weeks and that of three weeks are evaluated. For this, only persons without any special days in their report (sick days or vacation days) are analysed ($n=1,678$ respondents). Two out of the three weeks of trip diaries are selected by a simple random sampling approach. For the comparison, the same two weeks per person are analysed. The identification of whether a tour in these two weeks is routine or non-routine is based on the $Q3+1.5 \times IQR$ method. The threshold of this method is calculated on the one hand based on the three weeks of trip diaries (threshold based on 3 weeks) and on the other hand based on the selected two weeks of trip diaries (threshold based on 2 weeks). These two thresholds, using the same method but a different data basis, are then applied to the same two weeks of trip diaries.

90 % of all tours were classified as routine tours in both approaches, and 8 % were identified as non-routine tours in both. Thus, 98 % of all tours were labelled in the same way. There is no difference for these tours, if the threshold for the usual environment is calculated based on two weeks or on three weeks. Only 2 % of all tours have a different classification depending on the reporting length used to determine the threshold. To assess the impact of the partly different compositions of the identified non-routine tours, some main characteristics are analysed in Table 3.

Table 3. Characteristics of identified non-routine tours differentiated by thresholds based on two and three weeks

Characteristics	With threshold based on 3 weeks	With threshold based on 2 weeks
<i>Tour level</i>	$n=3,207$ NRT	$n=3,121$ NRT
% NRT of all tours	9%	8%
% non-routine trips of all trips	11%	10%
% of distances travelled within NRT	38%	37%
Mean distance of NRT	110 km	109 km
Mean duration of NRT	350 min	351 min
Mean number of trips per NRT	2.7	2.7
<i>Person level</i>	$n=1,678$	$n=1,678$
Mean threshold	67 km	67 km
% of respondents without NRT	29%	29%
Mean number of NRT per respondent (in two weeks)	1.9	1.9

NRT= non-routine tours;

Again, it can be seen that the differences between the thresholds based on 2 weeks compared to 3 weeks are negligible. The average non-routine tour is about 110 km long and lasts about six hours. For 70 % of the respondents at least one non-routine tour can be detected in the two weeks of report. Although only 9 % of the tours are labelled as non-routine, they account for more than 37 % of the travelled distances. One central aspect should be considered when investigating the non-routine behaviour of respondents: The longer the observation period, the better. This can be seen when comparing the share of respondents without non-routine tours in the observation period of two weeks

(29 %) compared to three weeks (20 %). This illustrates that the longer the reporting period is, the more likely non-routine travel is captured. While a shorter period already allows for describing the characteristics of non-routine travel across the respondents, longitudinal data is needed to capture and describe non-routine travel on an individual level.

6. Conclusion

In this paper, a method was developed to identify the size of an individual's usual environment based on longitudinal trip diary data. The individual's threshold representing the size of the usual environment was then used to differentiate between routine and non-routine travel in the sense of the definition of tourism by the UNWTO. Ten indicators based on common statistical measures were calculated based on three weeks of trip diaries. The resulting thresholds for differentiating between routine and non-routine travel were then compared to each other, and one indicator based on the interquartile range was selected as the best approximation for describing the size of the usual environment. An additional analysis revealed that the method chosen performs equally sufficient if only two weeks of trip diaries are available. In the sample, about 70% of the respondents had at least one non-routine tour during two weeks of reported trip diaries. However, the study sample cannot be considered representative, as young age groups are underrepresented in this study due to the survey design. The possibility of applying the developed method to only two reporting weeks opens the door for future analyses with larger samples, in which the population as a whole is better represented. Further studies will then focus on the characteristics of the respondents related to non-routine behaviour. The presented method offers an approach to identify non-routine travel within the reported behaviour of an individual and is thus independent from external assumptions, e.g. by specifying a minimum distance of 100 km, or assumptions based on sociodemographic characteristics of the participants. The approach can therefore also be applied to other contexts, e.g. longitudinal national household travel surveys of other countries, as the delimitation of non-routine travel is derived individually and data-based from the reported behaviour of each individual.

References

- Axhausen, K. W., Zimmermann, A., Schönfelder, S., Rindsfuser, G., & Haupt, T. (2002). Observing the rhythms of daily life: A six-week travel diary. *Transportation*, 29, 95–124. <https://doi.org/10.1023/A:1014247822322>
- Chikodili, N. B., Abdulmalik, M. D., Abisoye, O. A., & Bashir, S. A. (2021). Outlier Detection in Multivariate Time Series Data Using a Fusion of K-Medoid, Standardized Euclidean Distance and Z-Score. In S. Misra & B. Muhammad-Bello (Hrsg.), *Information and Communication Technology and Applications* (Bd. 1350, S. 259–271). Springer International Publishing. https://doi.org/10.1007/978-3-030-69143-1_21
- Czepkiewicz, M., Heinonen, J., & Ottelin, J. (2018). Why do urbanites travel more than do others? A review of associations between urban form and long-distance leisure travel. *Environmental Research Letters*, 13(7). <https://doi.org/10.1088/1748-9326/aac9d2>
- DESTATIS - Statistisches Bundesamt. (2018). *Mikrozenus 2017—Qualitätsbericht*.
- DESTATIS - Statistisches Bundesamt. (2021). *Einkommens- und Verbrauchsstichprobe (EVS) 2018—Qualitätsbericht*.
- Hasnat, M. M., & Hasan, S. (2018). Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. *Transportation Research Part C: Emerging Technologies*, 96, 38–54. <https://doi.org/10.1016/j.trc.2018.09.006>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Magdolen, M., Ecke, L., Hilgert, T., Chlond, B., & Vortisch, P. (2020). Identification of Non-Routine Tours in Everyday Travel Behavior. In *Transportation Research Board (Hrsg.), TRB 99th Annual Meeting*. <https://doi.org/10.5445/IR/1000105265>
- Mattioli, G. (2023). Long-distance travel. *Préparer la transition mobilité*. <https://forumviesmobiles.org/en/dictionary/15897/long-distance-travel>
- Mor, M., Dalyot, S., & Ram, Y. (2023). Who is a tourist? Classifying international urban tourists using machine learning. *Tourism Management*, 95, 104689. <https://doi.org/10.1016/j.tourman.2022.104689>
- Rousseeuw, P. J., & Hubert, M. (2011). Robust statistics for outlier detection. *WIREs Data Mining and Knowledge Discovery*, 1(1), 73–79. <https://doi.org/10.1002/widm.2>
- Schulz, A., Magdolen, M., Chlond, B., Kuhnimhof, T., & Köhler, K. (2024). Identifying and understanding long-distance travel demand by combining official transport statistics and survey data. *Transportation Research Procedia*, 76, 458–469. <https://doi.org/10.1016/j.trpro.2023.12.068>
- Vallée, J., Ecke, L., Chlond, B., & Vortisch, P. (2022). Deutsches Mobilitätspanel (MOP) – Wissenschaftliche Begleitung und Auswertungen Bericht 2021/2022: Alltagsmobilität und Fahrleistung. Bundesministerium für Verkehr und digitale Infrastruktur (BMVI). <https://doi.org/10.5445/IR/1000154338>
- World Tourism Organization. (2007). *Understanding Tourism: Basic Glossary*. <http://cf.cdn.unwto.org/sites/all/files/docpdf/glossaryenrev.pdf>
- Zijlstra, W. P., Van Der Ark, L. A., & Sijtsma, K. (2007). Outlier Detection in Test and Questionnaire Data. *Multivariate Behavioral Research*, 42(3), 531–555. <https://doi.org/10.1080/00273170701384340>