

# Analysing gender equality in Barcelona through (spatiotemporal) segmentation

Lidia Montero <sup>1</sup>, Lucía Mejía-Dorantes <sup>2</sup>, and Jaume Barceló<sup>1</sup>

<sup>1</sup> Universitat Politècnica de Catalunya (UPC); lidia.montero@upc.edu, jaume.barcelo@upc.edu,

<sup>2</sup> Consultant; mejia-dorantes@web.de

## Abstract

Citizens take part in different activities to satisfy their needs, to invest in their socio-economic progress, participate in social and health activities that improve their well-being. However, activity participation is influenced by many factors in the built environment, but also individual's attributes. Herein we analyze activity participation and travel through sequence analysis. This method explores sequences of daily activity and travel employing techniques from the sequencing of events in the life course of individuals. Studying sequences of daily episodes (each activity and each trip) considers the entire trajectory of a person's activity during a day while at the same time considering the number of activities, order of activities in a day, and their durations jointly. We applied this method to a sample of residents in the Metropolitan Area of Barcelona (RMB) in the 2018, 2019 and 2020 EMEF Travel Surveys. The EMEF2020 deserves a particular analysis since activity patterns are expected to vary compared to pre-COVID19 spread. We have focused on that fragmentation in activity participation over the mean among persons in specific gender, age, activity and transportation mode.

**Keywords:** Travel behaviour, fragmentation, gender, classification analysis, life course, activity participation

## 1. Introduction/Motivation

In general terms, the spatial and temporal distribution of activities in the urban environment determine how people move to reach places in a timely and secure manner. However, other issues define how we access different services, for example, the transport systems available in the area, along with the topographic and weather characteristics of the site. Furthermore, the intrinsic characteristics of the subject who needs to access different places, such as the socioeconomic level, profession, gender, disabilities, family status, religion, ethnicity, among others, play a key role in shaping mobility.

In this regard, different studies observe that lifestyles differ notably in terms of gender. The gender gap influences gendered mobility patterns (EIGE 2019). For example, the high share of women in certain professions, like healthcare or education, influences how they work, their income, and their working schedules. Furthermore, the unpaid work and family responsibilities (care work) also shape their participation in economic activities, and hence it influence how, when and with whom they move. This is translated into women having more complicated commuter patterns than men, based on shorter, chained trips (Cresswell 2008; EIGE 2022), and preferring certain transport modes (Cubells, Marquet, and Miralles-Guasch 2020; Hanson 2010), and different working schedules (Hyde et al. 2019).

In addition, years before the pandemic, it was widely accepted that the coming Information and Communication Technologies (ICTs) would bring changes in the way people move, for example to work (Alexander, Ettema, and Dijst 2010). This would be reflected in the organization of daily activities which would lead to a higher flexibility in scheduling daily activities (Couclelis 2000, 2006). As a result of this flexibility, the schedule and sequence of daily activities would be transformed into a multiple switching between different activities in a day leading to increased transport demand because many activities are no longer bound to specific times and specific places. Similarly, a change of transport demand would be expected as a result of a higher flexibility. Authors such as (Alexander, Ettema, and Dijst 2010) also comment that it may even impact dwellings' requirements. The former could be verified through the analysis of fragmentation of individual daily activities. This concept of fragmentation is related to how activities are spatiotemporally reorganized, by subdividing activities into smaller components that are then performed at different times and/or locations.

Indeed, after the COVID-19 pandemic outbreak, we have seen a remarkable increase of ICTs in daily activities. According to the EU statistics (Eurostat 2020) in the EU before the pandemic only 5.2% of the labour force usually worked from home, whereas in Spain this share was of about 4.3%. A recent survey by the National Statistics Institute in Spain (INE 2021), observes that people that have teleworked partially and totally in Spain and Catalonia has increased to 17.6% and 23.4% respectively. However, as Fana et al. (Fana et al. 2020) observe, generally people that telework tend to be high-skilled workers with better wages and contract conditions. This means, that in Barcelona 76.6 % may not work from home, many of them employed as frontline workers.

Notwithstanding this increase in ICT in daily activities, there are many activities that need to be carried out physically, which cannot be transferred to virtual services: visiting schools, doctors, and many escorting activities. Furthermore, other activities are preferably carried out personally, for example those that make us feel part of a community or the society, that contribute to people's well-being. In all these cases, citizens depend on the transport alternatives available.

In this study we carry out a sequence analysis to measure the fragmentation in activity participation in a Transport Oriented Development (TOD) area to discuss the similarities or differences between results. The aim of this paper is to gain a better understanding of the spatiotemporal patterns and travel behaviour in Barcelona, to analyze gendered mobility patterns and discuss which characteristics play a significant role in shaping mobility. Our approach is based on sequence analysis, making use of three recent and consecutive mobility surveys in the Metropolitan Region of Barcelona (RMB). This paper is divided into 7 sections and organized as follows. After this introduction, we present the literature review. Afterwards, we present the case study, datasets and the methodological approach. Section 5 presents the results. Finally, we present in section 6 and 7 the discussion and the main conclusions and implications of this study.

## 2. Literature Review

Sequence analysis has been developed in social sciences to understand the occurrence of social events in a structured manner. Andrew Abbott is recognized as a trailblazer in sequence analysis, developing the concepts and methodology to move beyond on how historians order events and how quantitative analysis deals with sequence in social processes, with a wide variety of publications exposing the development of these concepts and methodology (Abbott 1983, 1984; Abbott and Forrest 1986; Abbott and Tsay 2000; Leszczyc and Timmermans 2002). As (Abbott and Forrest 1986) states: *"Social reality happens in sequences of actions with constraining or*

*enabling structures (...). It is a matter of particular social actors, in particular social places, at particular social times”.*

A sequence is a list of events, actions, performed in a ordered manner. As Abbott and Forrest (Abbott and Forrest 1986) describe, a sequence dataset may describe two different patterns. One may describe activities that occur only once and the other may describe activities that occur several times in a certain sequence. Whereas the first idea may be solved with permutation, the second analyzes recurrent events.

This approach is an alternative to other methods, such as time-series methods, which intends to highlight historical, unilinear information which, according to Abbott, in the later case is restricted and not the main point of analysis. In this respect, authors such as Abbott and Tsay (Abbott and Tsay 2000), or Levy and Widmer (Levy and Widmer 2013) explain that many states or events in individual life courses are unilinear, such as family, occupational or residential social trajectories.

Many authors in the sociology field have studied life course events such as marriage, childbearing, and employment through activity sequence analysis (Elzinga and Studer 2019; Giele and Elder 1998). This analysis has been recently applied in mobility behaviour. In this respect, Bhat and Pnjari (Bhat and Pinjari 2007) observe that sequences of activities and the daily transitioning from one activity to another as well as the amount of time spent in each activity represents an important direction of travel behavior analysis.

In fact, analysis of activities' sequences and travel is critical in formulating econometric models embedded in activity-based daily simulations of household activity-travel patterns for large-scale travel demand analysis as many authors observe (Bhat et al. 2013; Burchell, Reuschke, and Zhang 2020; Paleti et al. 2017; Rasouli and Timmermans 2014).

Authors, such as Studer and Ritschard (Studer and Ritschard 2016), identify the following important intertwined characteristics:

- Experienced states: the distinct alternatives present in the sequence
- Distribution: the total time or state distribution within a sequence
- Timing: the moment of time in which each state appears
- Duration: the time span in the different successive states
- Sequencing: the order in which the distinct successive states take place

As stated by Studer and Ritschard (Studer and Ritschard 2016), two sequences are compared to quantify the level of mismatch between these sequences, which is a measure of dissimilarity. These authors observe that there are different approaches to define the dissimilarity of sequences. Recently, McBride et al. (McBride, Davis, and Goulias 2020), applied this approach to analyze mobility behaviour in California. They explore events with regards to time and their duration.

On the other hand, the fragmentation concept has been defined by different authors, such as (Alexander, Ettema, and Dijst 2010; Couclelis 2006; Hubers, Schwanen, and Dijst 2008; McBride, Davis, and Goulias 2020). They define the fragmentation of travel and activities as the sequence of many short trips that take place during the daily schedule of a person. Whereas the temporal fragmentation is related to the different times that activities are carried out, the spatial fragmentation is related to the locations where activities are performed. Together with other activities and movements that occur in a larger timeframe, they build up a string of activities with different durations and purposes. This string may have different complexities depending on different extrinsic and intrinsic characteristics of the individual and the urban environment which shape mobility behaviour.

Furthermore, despite of not providing an explanation of how and why individuals engage in activity-travel fragmentation, the classification of activity-travel fragmentation into clusters makes possible to understand different groups. For instance, it allows us, to establish the relationship with socioeconomic characteristics of the segments. Some authors, such as McBride et al. (McBride, Davis, and Goulias 2020), have pointed the need to understand gender roles that it is considered to be related to time allocation to activities and thus, activity-travel.

Furthermore, there are some researchers that observe gender differentiated patterns in segmentation analyses. For example, Leszczyc and Timmermans (Leszczyc and Timmermans 2002) analyze the Dutch diary and concluded that gender and age are important determinants of moving from one activity type to another. Burchell et al., (Burchell, Reuschke, and Zhang 2020) analyze the gender differences in the segmentation of workplace patterns using the 2015 European Working Conditions Survey (EWCS 2015), which presents information before the pandemic. The authors observe that there are clear differences when analyzing gender, for example women were more likely than men to work at the employer's offices. On the other hand, von Behren et al. (von Behren et al. 2020) carried out an image-based clustering analysis of the individuals' pattern segmentation using the German Mobility Panel of activity (BMDV 2020). The authors identify two clusters with children in the household, one is predominantly characterised by women and part time workers.

In the next sections we will apply this theoretic approach to analyze the gendered mobility patterns in the Metropolitan Region of Barcelona. To our knowledge, this is the first attempt to apply this technique in the Spanish context, and making use of a longitudinal dataset.

### 3. Case Study

We apply this theory in the Metropolitan Region of Barcelona (RMB) depicted in Figure 1. It is composed of 36 municipalities in the Metropolitan Area of Barcelona (AMB) subarea, which accounts with 3,239,337 inhabitants. It has a well-scattered public transportation network with more than 200 bus lines, 4,000 stops, 10 metro lines, 15 railways lines, and two tramway lines. More than 9 million trips are carried out every day. The rest of the RMB area consists of 164 municipalities and 1,848,514 inhabitants. A detailed travel demand modelling is regularly conducted in the AMB subarea (see green zone Figure 1).

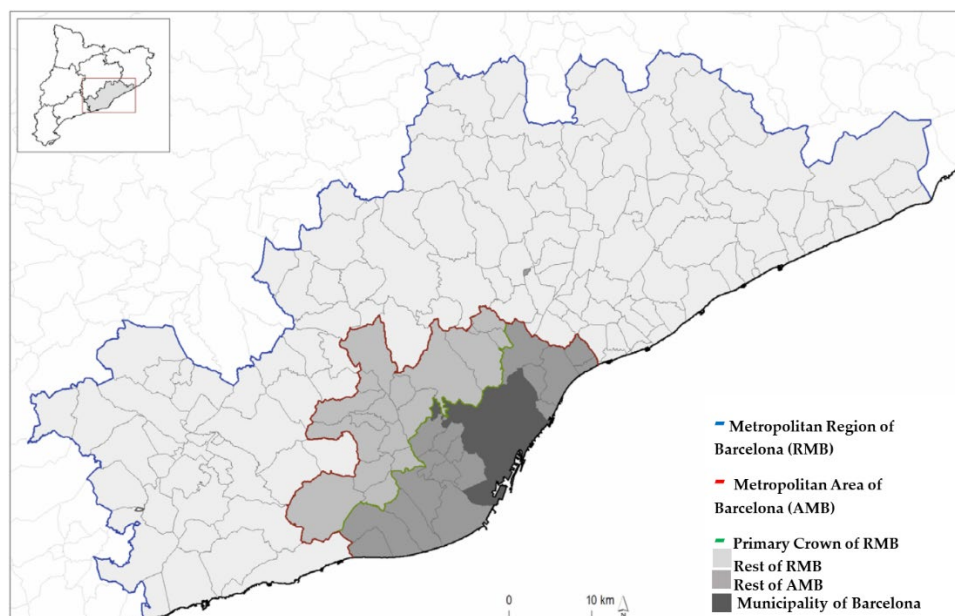


Figure 1. RMB Study Area: Transportation Analysis Zones. AMB subarea in green.

A case of special interest is the Primary Crown (Primary in what follows) of the Metropolitan area that includes the 18 most populated municipalities. More information on the Metropolitan Region of Barcelona may be found in (Mejía-Dorantes, Montero, and Barceló 2021).

## 4. Data Description and Methodology

The EMEF (Weekday Mobility Survey, in Catalan) surveys of 2018, 2019 and 2020 are examined to define the sequences of visited places by a person during a day jointly with the duration of activities at each place and the travel times to reach these places.

The entire daily sequences of activities and travel patterns are quantified by three indicators, defined in McBride et al. (McBride, Davis, and Goulias 2020):

- Normalized Entropy, which explores the variety in daily schedules;
- Turbulence, which shows the complexity in daily schedules, and
- Complexity, a normalized [0,1] score based on entropy, and different sequences of the individual's schedule.
- Travel Time Ratio (TTR): the total travel time in a day divided by the sum of the total time outside the home plus the total travel time in a day'

As explained by McBride et al. (McBride, Davis, and Goulias 2020), these summary indicators are correlated to each other. They quantify daily activity-travel patterns for each individual in a numerical way.

The statistical analysis is carried out by the TraMineR package in R (Gabadinho et al. 2011; R Development Core Team 2021). It has been widely used for the analysis of biographical longitudinal data in social sciences, but other approaches have been also described.

After examining the sequence of activities and travel patterns, a clustering technique was used to evaluate the results.

### 4.1. Data description

The mobility survey on working days (EMEF) from 2018 to 2020 were used to carry out this research. They are traditional mobility surveys that analyze the mobility of residents in the Metropolitan Region of Barcelona (RMB) for individuals aged 16 and over. The spatial granularity is at municipality level, but as Barcelona is divided into ten districts, it leads to a total of 296 macrozones, where only 45 of them in the AMB area. The EMEF 2020 survey was launched during the Fall 2020, when some mobility restrictions and the prevalence of some online activities were still present due to COVID-19 situation. The EMEF2020 deserves a particular analysis since activity patterns are expected to vary compared to pre-COVID19 spread.

The data collected for each journey refers to trips for the day before: origin and destination (macrozones), purpose, mode (a very detailed list of possibilities), travel start time and duration (min), vehicle use, parking use, etc. The sample units are individual residents, not households. The sample size for each year after removing the category "professional drivers" are of: 9,930; 9,934; and 10,024, respectively for 2018 to 2020 in the AMB area. The total trips are 36,368; 37,463; and, 30,591, respectively for 2018 to 2020 in the RMB area. After filtering (professional travellers, residential area missings, etc.), the final number of residents included in the total sample was found to be of 26,860.

We also make use of a demographic and land use data. This information is defined with the same spatial granularity defined by the EMEF surveys. It consists on population segmented by gender,

age group (5 groups), education level, educational places, services, land use, residential morphology, average per capita rent and number of stops in public network.

## 4.2. Methodology

A sequence is defined as a series of time points at which a subject can move from one discrete “state” to another. People with many states in their daily schedule have fragmented schedules. In this research we used sequence analysis to statistically analyze the fragmentation of respondents’ days using a minute-by-minute time series, in which every minute of the day contains a specific state for each person in the study. These states are based on types of places which individuals visit during their diary day. Activities initially considered are: home (H); work (W); casual (C) for not frequently visited places; other (O) for frequently visited places that are not the working place; and travel (T).

Among the many techniques in the travel behavior field that can be used to measure the duration of activities and transition rates from one activity to the next, we make use of entropy, turbulence and complexity. Entropy is the proportion of total time spent in each state and the number of state transitions is not taken into account (Gabadinho et al. 2011):

$$h(x) = h(\pi_1 \dots \pi_S) = - \sum_{i=1}^S \pi_i \log(\pi_i)$$

Where  $\pi_i$  is the proportion of occurrences of the  $i$ th state in the considered sequence.  $S$  is the number of potential states and  $x$  is the sequence defined from minute to minute day activities.

The proportion of minutes allocated to each state during a day defines the entropy indicator. For this measure the number of state changes is irrelevant. If a person has no state change during the entire day, his/her entropy would be 0. In contrast, visiting several states makes entropy to increase. The range of possible values depends on the number of states and the maximum is allocated at sequences showing equal amount of time in each state. In our case maximum entropy is 1.61. A normalized entropy score is often used consisting on dividing entropy into maximum entropy, thus a 0 to 1 range is obtained.

The sequence turbulence is a measure proposed by Elzinga and Liefbroer (Elzinga and Liefbroer 2007) for measuring schedule complexity in daily activities. It is based on sequence permanence and uses the number of distinct subsequences that can be extracted from the distinct state sequence and the variance of consecutive time points spent in a distinct state. For a sequence  $x$ , the formula for  $T(x)$  turbulence is (McBride, Davis, and Goulias 2019):

$$T(x) = \log_2 \left( \phi(x) \frac{s_{max}^2 + 1}{s^2 + 1} \right)$$

Where

$\phi(x)$  is the of distinct subsequences that can be extracted from the distinct state sequence accounting on time precedence.

$s^2$  is the variance for the state duration

$s_{max}^2$  is the maximum variance to be given based on the duration of the sequence and it is computed as  $s_{max}^2 = (n - 1)(1 - \bar{t})^2$ , where  $n-1$  is the number of transitions in the sequence and  $\bar{t}$  is the sequence duration divided by the number of distinct states in the sequence.

$$C(x) = \sqrt{\left( \frac{nt(x)}{l(x) - 1} \frac{h(x)}{h_{max}} \right)}$$

Where  $nt(x)$  is the number of distinct transitions within a sequence,  $l(x)$  is the length of the sequence,  $h(x)$  is the entropy indicator and  $h_{max}$  is the maximum entropy in the sample. This indicator will have a value between 0 and 1, with zero corresponding to entropy zero and no transitions (e.g., staying at a single place for the entire day of observation).

The travel time ratio (TTR), defined as McBride et al. (McBride, Davis, and Goulias 2020), is as a compact indicator that represents the trade-offs of people between travel and activity time. In this paper, TTR is defined as ‘the total travel time in a day divided by the sum of the total time at home plus the total travel time in a day’. Thus, TTR ranges from 0.5 (no trip-makers) to 1.0 (whole day out of home).

To achieve a better understanding of the behaviour of the turbulence and its dependency with the individual features like gender, age group, activity, handicapped status, telework possibility, day of the week and macrozone in the area of Barcelona, along with different descriptive variables including percentage of households over 100m2, number of bus stops, number of metro, tram and train stops, a linear model of their logarithms has been applied to the 2018-2020 subsample of trip makers.

Later we compare all sequences with each other to address sequence dissimilarity and compute pairwise mean fragmentation indicators by year using the Tukey’s Honestly Significant Difference (HSD) Test (Tukey et al. 1984).

Finally, we use a clustering technique to group sequences of activities with similar dissimilarity scores. The final number of clusters is optimized to represent the data using a criterion of within group similarity and across groups dissimilarity.

## 5. Results

A first analysis was carried out by simplifying the kind of activities that people carry out, or places that people visit. The activities initially considered are home (H); work (W); casual (C) for not frequently visited places; other (O) for frequently visited places that are not the working place; and travel (T). Table 1 shows as an example some sequences of activities and durations for the 3 first elements of our sample and some defined scores.

Unit No.	Daily activity sequence	Used time by activity (min)	Total duration (min)	Entropy	Normalized Entropy	Turbulence	Complexity
1	H-T-W-T-H	360-20-600-20-440	1440	0.707	0.439	6.998	0.768
2	H-T-W-T-H-T-O-T-H	450-10-740-10-110-20-280-40-0	1660	1.137	0.706	10.975	0.924
3	H-T-O-T-H-T-O-T-H	450-30-390-30-150-10-170-30-180	1440	0.781	0.485	10.984	0.766

Table 1 . Daily-travel pattern for 3 units of the working sample.

The first two patterns belong to foreign young men living in the first district of Barcelona city having neither a car, nor a motorbike. The third pattern belongs to an student who normally uses public transport or walking. A graphical representation of a daily sequence is depicted in Figure 2. The activity set is defined as the activity alphabet and, herein, for illustration purposes, set up as casual (C), home (H), other (O), travel (T), and work (W).



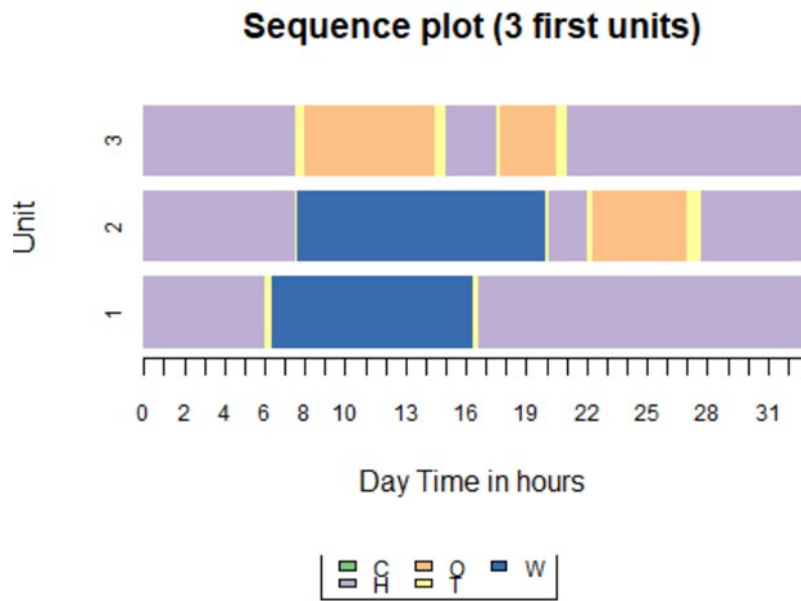


Figure 2. Graphical representation of daily-travel pattern for 3 units of the working sample.

For illustration purposes, the activity distribution across the day in a subset of the working sample is shown in Figure 3. An interpretation of this figure indicates at 13:00h, 3% of the sample units show a casual activity, 32% stays at home, 14% are involved in other activity, 9% is traveling and the rest, 58%, is working.

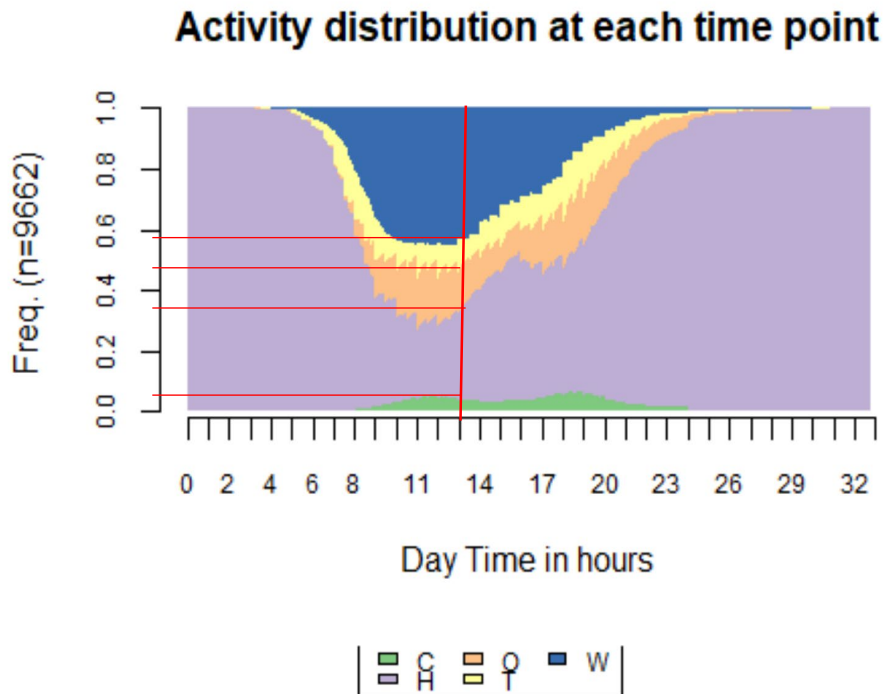


Figure 3. Activity Distribution by time

The relationship among numeric indicators relationships are shown in Figure 4. Herein a direct association can be seen between normalized entropy and turbulence as well as complexity vs turbulence, having both a correlation around 0.7.



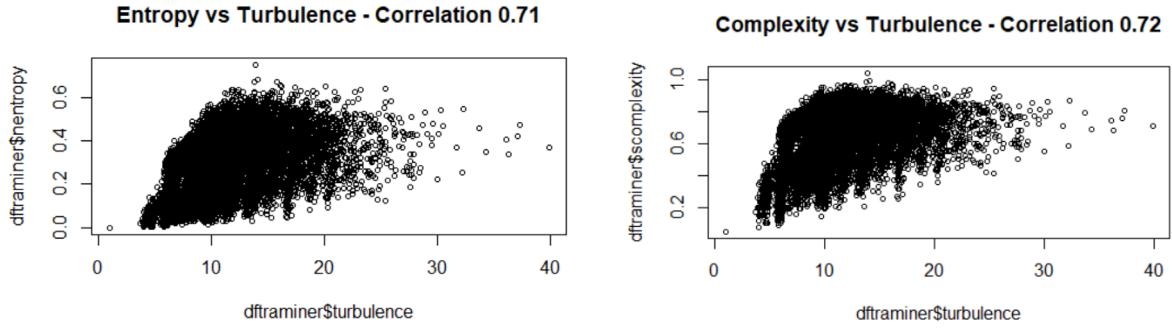


Figure 4. Normalized entropy vs Turbulence indicators (left) and Complexity vs Turbulence indicators (right). EMEF 2018 to 2020 data

Figure 5 shows normalized entropy, turbulence, and complexity distributions useful for the generation of synthetic populations (anonymized population that is representative of real population numbers), where some questions arise. For example, why has the normalized entropy a maximum at 0.5 corresponding to daily pattern (H-T-W-T-H-T-O-T-H 450-10-740-10-110-20-280-40), and the 0 value corresponds to non-trip-makers, the same high bar at zero is seen for turbulence and complexity histograms.

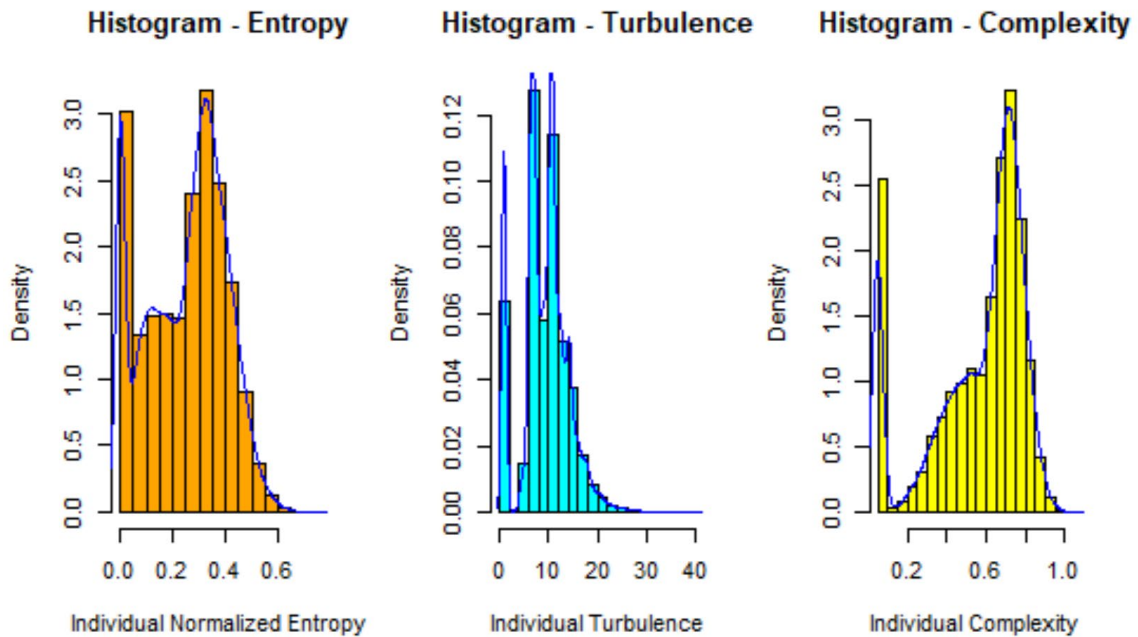


Figure 5. Normalized entropy, Turbulence and Complexity distributions. EMEF 2018 to 2020 data

To achieve a better understanding of the behaviour of the turbulence and its dependency of the individual features, a linear model of their logarithms has been applied to the 2018-2020 EMEF *subsample of trip makers*. These individual characteristics, include gender, age group, activity, disability status, telework possibility (with regards to year 2020), day of the week and macrozone descriptive variables including percentage of households over 100m<sup>2</sup>, number of bus stops, number of metro, tram and train stops. The marginal effect of the variables having significant net-effects is shown in Figure 6. The results show that the fragmentation of female out-of-home activities is significantly greater than those for men in the 30-44 age group. A remarkable effect of year 2020 on the logarithm of the turbulence is seen once controlled by the other considered factors.

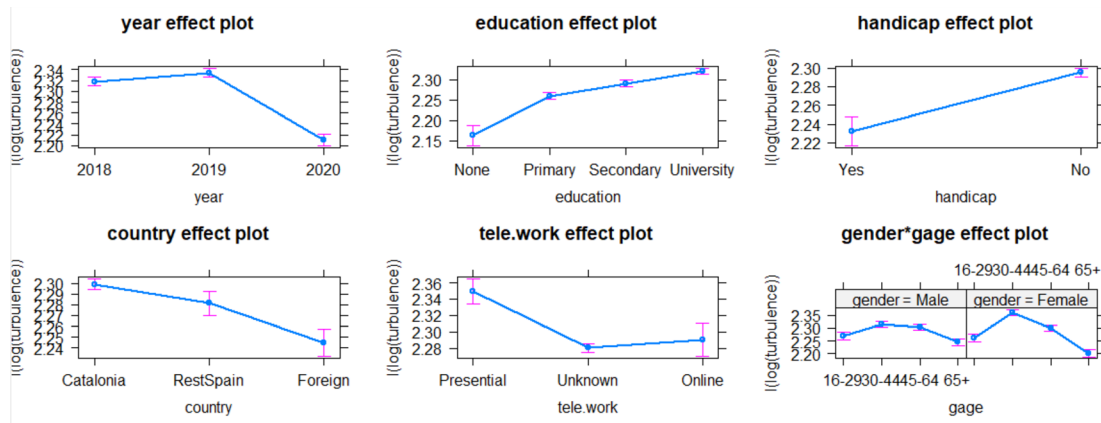


Figure 6. Turbulence transformation marginal effect for 2018-2020 EMEF (only trip makers)

Furthermore, Figure 7 shows the logarithm of the turbulence according to gender, age group and year, once education, birthplace, disability status and telework possibilities accounted in the model. Women in the 30-44 age group show the highest turbulence in any year; while men turbulences seem stable from 16 to 64 age groups in 2018 and 2019, the youngest groups, either men or women are affected by a severe decrease in fragmentation in 2020. In general terms, the 2020 turbulences are reduced after COVID-19 breakout.

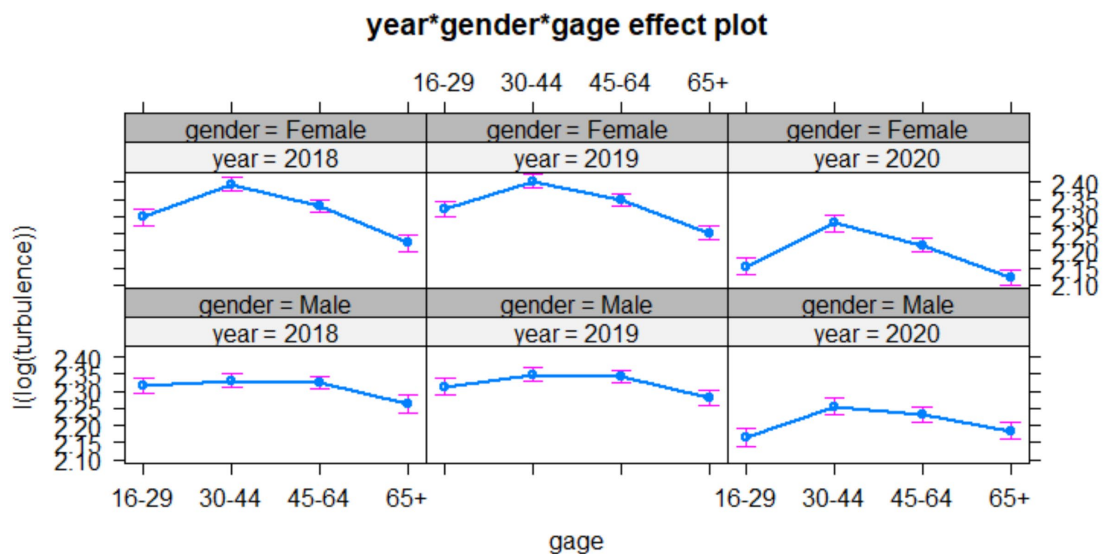


Figure 7. Marginal effect of the turbulence transformation with regards to year, gender, and age-group interaction in the total 2018-2020 EMEF dataset (only trip makers) after controlling by education, origin country, handicapped and telework effects.

Given the nature of the results and the no justification of any normality hypothesis, further analysis to get insights on their behaviour were conducted resorting to nonparametric statistical methods, as for example Kruskal-Wallis test for homogeneity in means across groups. The Tukey HSD test has been employed to address pairwise comparisons between means in groups.

Considering the (normalized) entropy indicator, the gross effect of gender is significant according to Kruskal-Wallis non-parametric test for means depending on year at 95% confidence level. The male entropy is greater than the female entropy being the highest value at the youngest age group (see Figure 8).

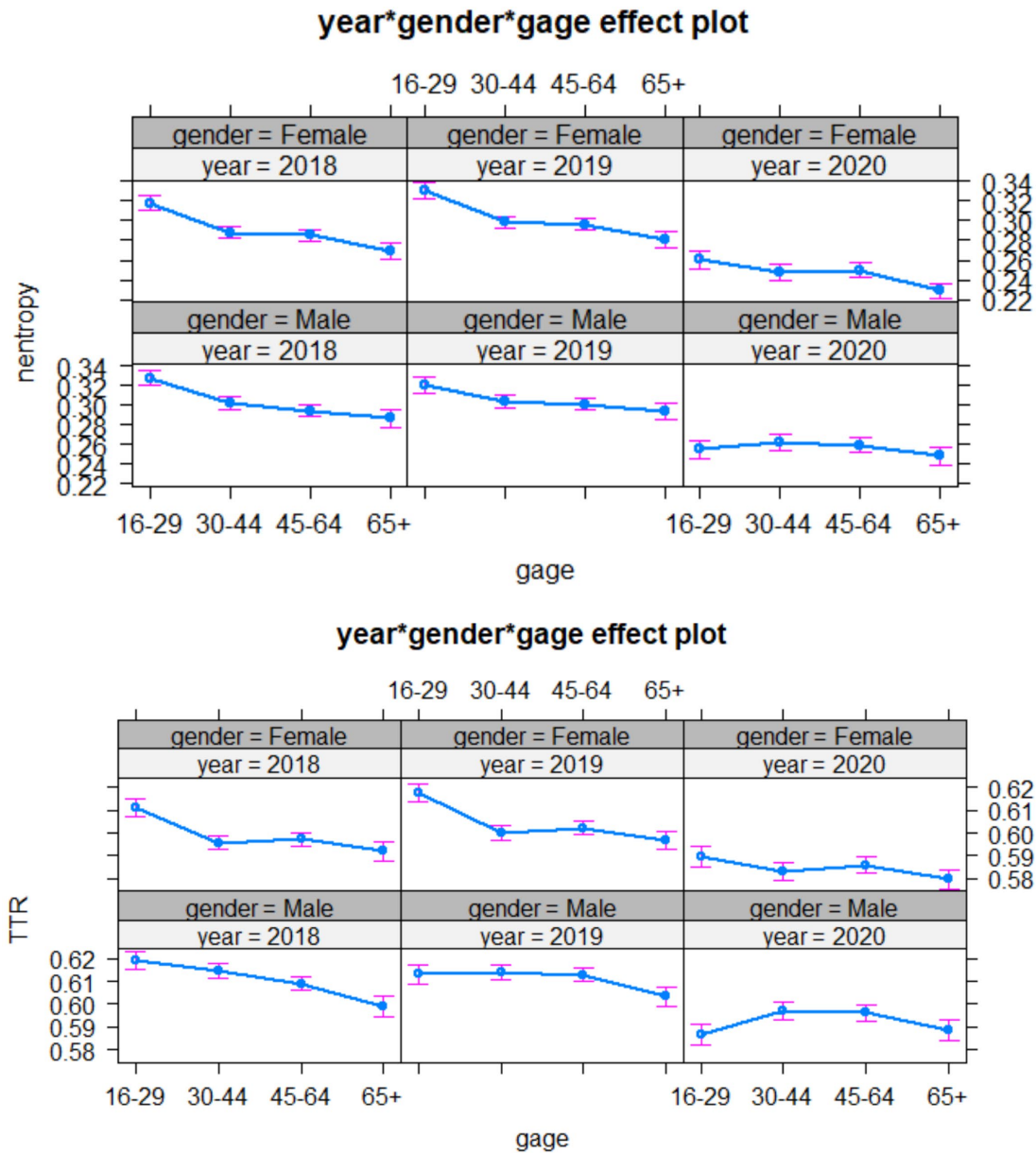


Figure 8. Normalized Entropy marginal effects (upper figure) and Travel Time Rate (TTR, lower figure) for gender and age-group and year interaction in 2018-2020 EMEF subsample of trip-makers after controlling by year, education, activity, country of origin, disability, and telework effects.

As previously shown, primary results were found very promising, therefore a redefined set of activities based on types of places individuals visit during their diary day were used for further analyses: Home (H), Work (W), School (only for students, S), Casual (for not frequently visited places, C), and Others (frequently visited places that are not the working place, O). Furthermore, the alphabet should and must be extended and refined since activities like Escorting/Accompanying (A) can be considered, as well as the differentiation of activities by travel mode since their data are available. Travel activities are divided into TW for active modes (walking, cycling), TC for private transport (car, van, motorcycle), TP for public transport (bus, metro, tram and train) and TM for other transport modes, such as van and truck. An alphabet of 10 activities is considered. Either trip makers, or non-trip makers are included and a sequence of 1440 minutes (or more) with each minute classified as a category of the alphabet. Figure 8 shows the refined daily activity pattern distribution across years. Evidently, the 2020 state activity pattern distribution shows an overrepresentation of whole day home activities.

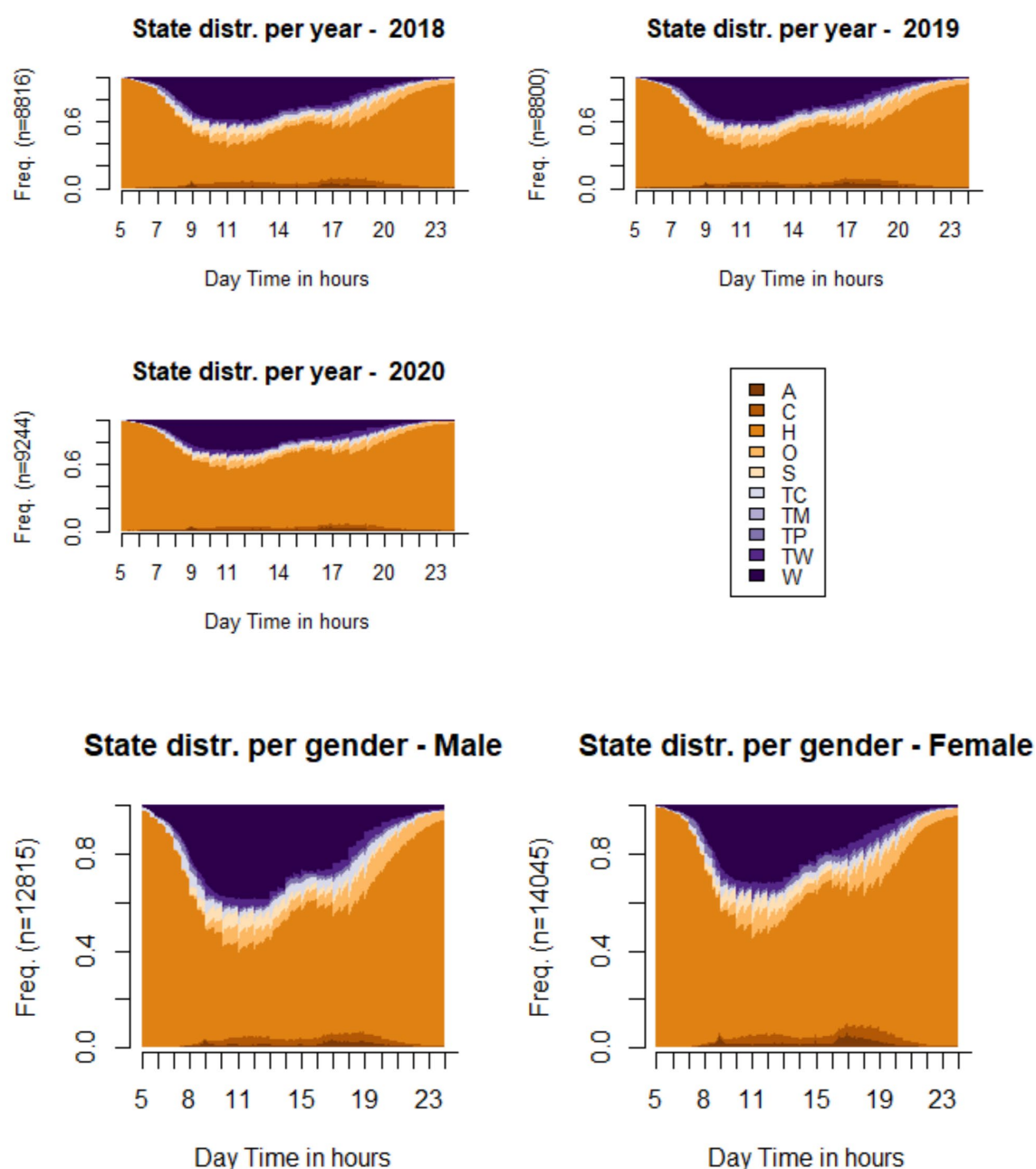


Figure 9. State distribution by year (top) and by gender (down). EMEF 2018 to 2020

The percentages of activity distribution across the day for men and women (Figure 9) and by year are shown in Table 2. Work activities out of home is greater for men than women, school activity is similar in both. Home activity in 2020 (73.5%) increases by almost 13 points compared to previous years, while school and work activities outside home are clearly reduced. The percentage of private transport use is severely affected by residential area (analysed here as crowns), showing that in non-central crowns there is an increase of private transport activity (3.4%) compared to Barcelona city (1.5%); consequently, public transport in the central crown as Barcelona city shows 2.8% incidence, while external crowns lie between 1.2-1.4% incidence. The spatial effects must be addressed in activity analysis (Table 2).

Gender	Activity									
	A	C	H	O	S	TC	TM	TP	TW	W

	Escort.	Casual	Home	Other	School	Private Tr.	Other Tr.mode	Public Tr.	Active modes	Work
Male	1.0%	1.9%	62.6%	4.9%	2.5%	3.0%	0.2%	1.4%	2.2%	20.2%
Female	1.3%	2.3%	68.1%	4.6%	2.4%	2.0%	0.1%	1.8%	2.2%	15.2%

Year	Activity									
	A	C	H	O	S	TC	TM	TP	TW	W
2018	1.0%	2.7%	61.7%	5.1%	2.8%	2.9%	0.0%	2.0%	2.1%	19.6%
2019	1.5%	2.1%	60.8%	5.9%	3.0%	2.8%	0.1%	1.9%	2.4%	19.5%
2020	0.9%	1.6%	73.5%	3.5%	1.5%	1.8%	0.2%	0.9%	2.2%	13.8%

Crown	Activity									
	A	C	H	O	S	TC	TM	TP	TW	W
BCN	1.2%	2.7%	61.9%	5.8%	2.8%	1.5%	0.1%	2.8%	2.9%	18.4%
PRIMARY	1.4%	2.4%	63.8%	4.6%	2.5%	2.2%	0.1%	2.0%	2.5%	18.4%
AMB	1.3%	2.0%	63.3%	4.5%	2.8%	3.1%	0.1%	1.4%	2.2%	19.1%
RMB	1.2%	2.2%	62.1%	5.1%	2.6%	3.4%	0.2%	1.2%	2.3%	19.9%

Table 2. Percentages of activity distribution across the day for men and women and year (Figure 9). The percentage of activities is according to residential area (See Fig. 1). From 5:00 to 24:00 h using the EMEF 2018 to 2020.

The Kruskal-Wallis test confirms that a non-homogeneous normalized entropy is present across years ( $pvalue = 0$ ), while homogeneity of variances cannot be rejected. Tukey multiple comparisons of means confirms entropy mean in 2020 is less than those in 2018 and 2019 at 95% confidence. Accounting for the subset of trip makers the same conclusions hold and entropy in 2020 is reduced by 18.57%. The Kruskal-Wallis test is a non-parametric test for addressing mean homogeneity in groups (3 years) and the Tukey HSD can be applied to pairwise comparisons of means defined by groups.

The non-parametric Kruskal-Wallis test also confirms that a non-homogeneous travel time ratio (TTR) is present across years ( $pvalue = 0$ ). The Tukey multiple comparisons of means confirms a TTR mean in 2020 less than those in 2018 and 2019 at 95% confidence. Accounting for the subset of trip makers the same conclusion holds and TTR in 2020 is reduced by 3.5% compared to 2018 and 2019 aggregated TTR (See Figure 10 and Table 3).

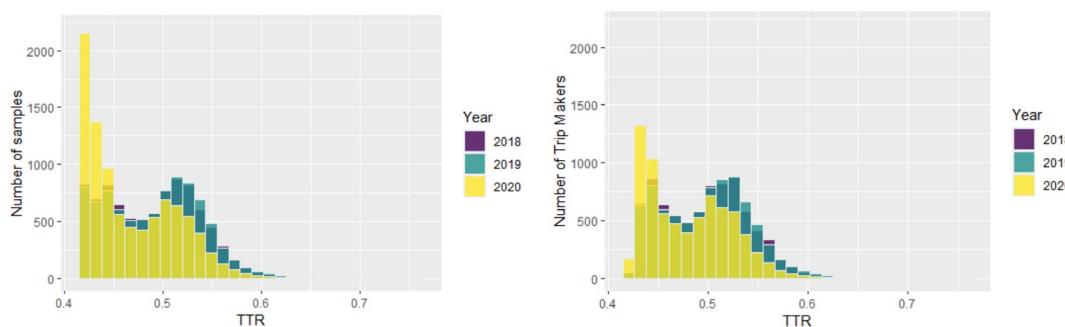


Figure 10. Travel time rate per year. All samples (left). Only trip makers (right)

Furthermore, the non-parametric Kruskal-Wallis test also confirms a non-homogeneous turbulence across years ( $pvalue = 0$ ). The Tukey multiple comparisons of means confirms a turbulence mean in 2020 is less than those in 2018 and 2019 at 95% confidence. Accounting for the subset of trip makers the same conclusions hold and turbulence in 2020 is reduced by 7.57%.

No differences are found between 2018 and 2019 at 99% confidence, as shown in Table 3 and Figure 11.

Turbulence						
Levels	Mean	Difference of Levels	Difference of means	lower bound diff. 95 % CI	upper bound diff. 95 % CI	Adjusted P-value
2018	10.62	2019-2018	0.18	0.04	0.31	0.01
2019	10.80	2020-2018	-0.73	-0.86	-0.59	0.00
2020	9.90	2020-2019	-0.90	-1.04	-0.76	0.00

TTR (Travel Time Ratio)						
Levels	Mean	Difference of Levels	Difference of means	lower bound diff. 95 % CI	upper bound diff. 95 % CI	Adjusted P-value
2018	0.4968	2019-2018	0.0018	0.0002	0.0034	0.0203
2019	0.4987	2020-2018	-0.0162	-0.0179	-0.0146	0.0000
2020	0.4806	2020-2019	-0.0180	-0.0197	-0.0164	0.0000

Table 3. Tukey HSD pairwise comparison for mean turbulence and TTR across years. EMEF 2018 to 2020.

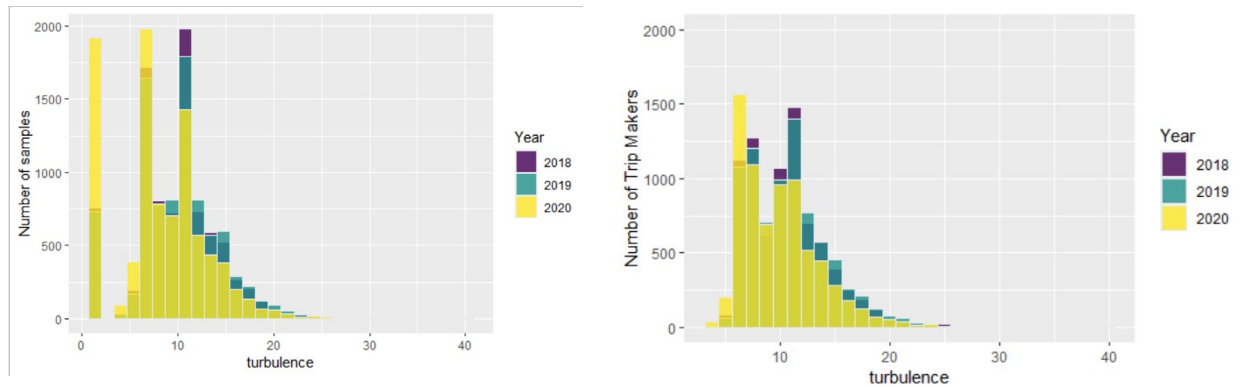


Figure 11. Turbulence distribution depending on year

The analysis is further developed through the following steps:

1. It considers every sequence of 1,140 minutes (from 5am to midnight) with each minute classified as home, work, casual, other, school, escorting and the four Travel categories: TW, TP, TC and TM.
2. A comparison in all sequences with each other and compute pairwise dissimilarity indicators is carried out.
3. A clustering technique that groups sequences by similar dissimilarity scores in such a way that low dissimilarity sequences are grouped together is applied.
4. The number of clusters that optimally represents the data using a criterion of within group similarity and across groups dissimilarity is evaluated.

To identify similar sequences, a rule for sequence comparison is needed. For example, we can perform different operations to reproduce one sequence departing from another and assign penalties to each operation (Wilson 1998). Measuring the difference between two sequences depends on the number of operations and sum of penalties accumulated in the comparison. The operations applied to this comparison are replacement, insertion, and deletion. In the sequence alignment literature, the measurement of dissimilarity and the number of operations needed to



make two sequences the same is called a distance or a dissimilarity score. The distance between two sequences is the minimum combination of operations (Abbott and Tsay 2000). The output of an algorithm that does these operations among all the sequences is a matrix of dissimilarity scores.

Measures of similarity between sequences found in literature are Longest Common Prefix (LCP, it is based on the length of the longest common prefix), Longest Common Subsequence (LCS, it is based on the length of the longest common subsequence) and Optimal Matching distances (OM, it is based on edit distances that are the minimal cost, in terms of insertions, deletions and substitutions, for transforming one sequence into another; there are two possibilities one assuming that all substitution costs are set equal to a constant and an alternative using the transition rates between states observed in the sequence data). An Optimal Matching distance based on substitution costs using transition rates found in the data would allow to obtain a representative distance matrix between sequences. This matrix has millions of cells ( $26,860 \times 26,860 = 721,459,600$ ) containing the dissimilarity scores among sequences for each person in the working sample. Method `seqdist()` from TraMineR package in R allows to calculate the dissimilarity matrix based on optimal matching or Hamming distances (Studer and Ritschard 2016) since day duration in minutes is the same for all units in the working sample. Nevertheless, it cannot be applied due to large memory requirements.

The dimensionality problem can be addressed in different ways. For example, McBride et al. (McBride, Davis, and Goulias 2020) chose to select a subsample, while another alternative consists on a reduction of dimensionality that is recommendable to suitable cope with this large dataset composed by daily sequences from 5 to 24h. The later was used herein, resulting in a  $26,860 \times 1,140$  matrix containing an activity from the selected alphabet (10 options) in which each column behaves as a categorical factor having 10 possible levels (activities). Consequently, the Multiple Correspondence Analysis (MCA) has to be applied instead of the Principal Component Analysis (PCA). Given the large memory requirements, 2.3 Gb, the proper MCA method in FactoMineR package (Husson et al. 2008) had to be chosen. All the available possibilities in R were checked (5 packages provide a MCA tool). Finally, the `mca()` method in MASS library was able to deal with this large dimensionality dataset. According to Kaiser criteria, 500 axes for minute-activity composed by trip makers (non-trip makers are grouped into one specific cluster), 93.47% of data variability is retained and sample projections in the new space can be computed. The dimension reduction is 95%.

After the data reduction technique, the next step consisted in conducting a Hierarchical Clustering (HC) (Husson, Lê, and Pages 2010) to the daily travel patterns included in the minute activity matrix based on distances between daily travel sequences, grouping in each cluster the points that are more similar among them than with the points in other groups. The Hierarchical Clustering method in FactoMineR package allows to reduce the computational burden by starting the agglomerative process on a heuristic partition being 10% of the original length and cutting the hierarchical agglomerative tree using well-known techniques.

The representativeness of 100 clusters accounts for 43% of total data variability. The cluster outcome has been transferred to the dataset containing descriptive characteristics of samples. It is remarkable the low number of samples for some clusters (see Figure 12).



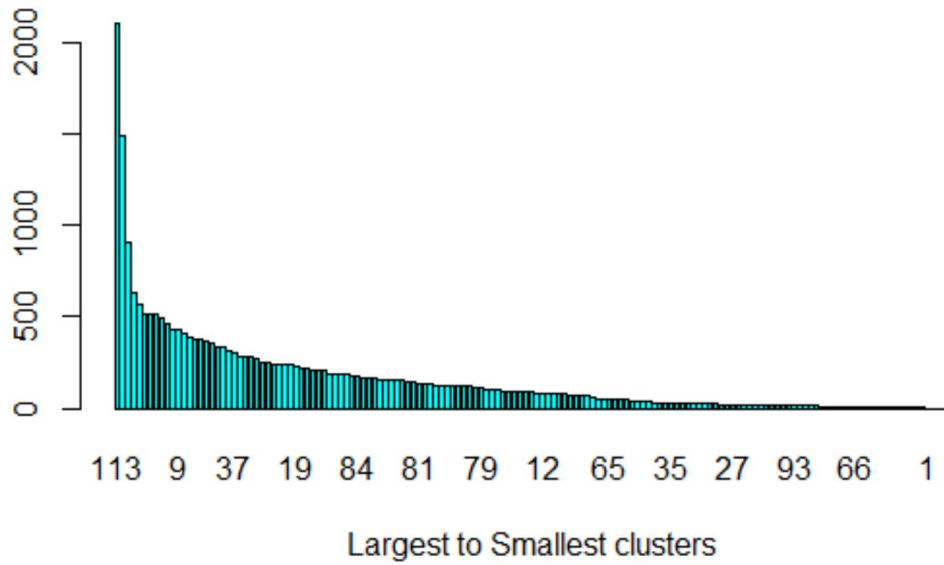


Figure 12. Cluster size (decreasing order in samples) for the whole EMEF 2018 to 2020 sample

The five largest clusters have been selected to show entropy, turbulence, complexity, and travel time ratio distribution. This is depicted in Figure 13 showing outliers and different statistical values for fragmentation indicators supporting the useful role of these indicators in cluster characterization.

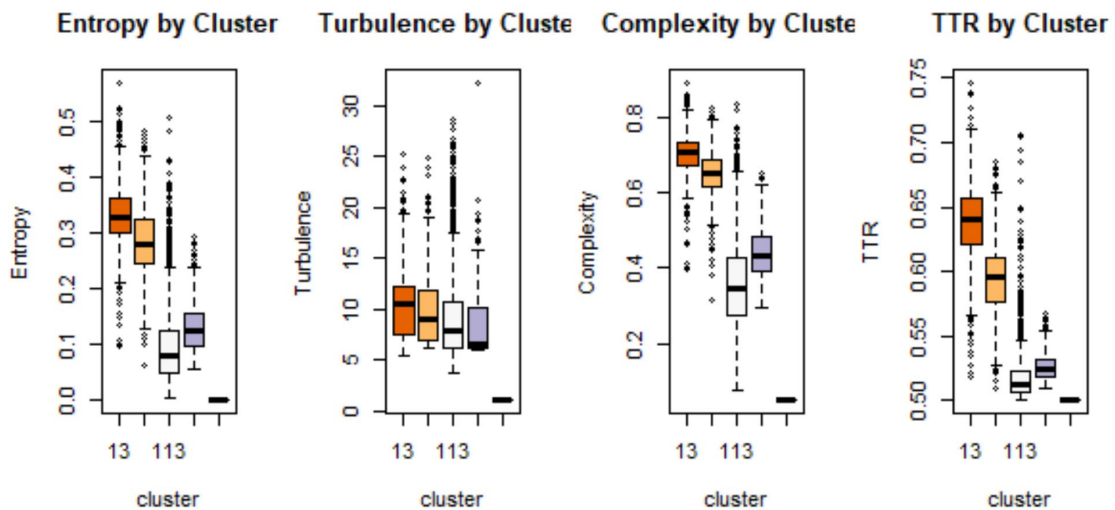


Figure 13. Normalized entropy, turbulence, complexity and TTR for largest clusters. EMEF 2018 to 2020.

Table 4 shows the profile for one of the clusters, based on categorical variables (cluster number 13). It gathers the profiles of car drivers and occasionally public transport users, working men in age groups between 30-44 and 45-64. They have higher education and come from Catalonia. Sample units (trip makers) from 2018 and 2019 are over-represented. The characterization of cluster 13 (profiling based on categorical variables) using the EMEF 2018-2020 is shown in Table 4. Cla/Mod indicates among the whole sample the percentage of units showing the category that have been allocated in cluster 13; Mod/Cla is the percentage of the category in cluster 13; Global is the percentage of the category in the sample; p.value refers to hypothesis testing of Mod/Cla percentage in cluster 13 being the same as global percentage; and v.test is the statistic

supporting the test (Husson, Lê, and Pages 2010). It is worth highlighting that this exercise may be carried out to any other clusters.

Category in variable	Cla/Mod	Mod/Cla	Global	p.value	v.test
activity=Active	33.44	99.80	52.12	0.00E+00	Inf
education=University	27.93	56.43	35.29	1.01E-75	18.41
gage=30-44	28.11	43.77	27.19	1.26E-52	15.27
handicap=No	19.23	97.79	88.80	6.50E-45	14.06
fbike=Occasionally	33.52	24.40	12.71	1.50E-43	13.84
gender=Male	23.70	59.72	44.00	4.85E-41	13.42
degree.hand=[0%,32%]	19.02	97.79	89.80	2.41E-38	12.95
fcard=Always	35.89	18.50	9.00	3.65E-38	12.92
fcard=Often	34.13	19.03	9.74	1.01E-34	12.29
country=Catalonia	20.41	80.90	69.21	7.50E-29	11.15
ftpub=Occasionally	24.72	37.40	26.42	7.68E-25	10.29
fcard=Occasionally	26.93	24.60	15.95	9.93E-22	9.58
year=2018	24.09	37.13	26.92	1.50E-21	9.54
gage=45-64	23.07	44.03	33.34	2.39E-21	9.49
fwalk=Often	22.15	49.33	38.90	2.05E-19	9.01
year=2019	22.76	34.79	26.69	2.62E-14	7.62
fcarnd=Occasionally	21.07	40.48	33.55	6.74E-10	6.17
tele.work=Presential	23.82	18.30	13.41	4.01E-09	5.88
fwalk=Occasionally	22.01	21.72	17.23	8.20E-07	4.93
day=Monday	20.49	23.73	20.23	2.64E-04	3.65
flexhorari=No	21.74	13.87	11.14	3.18E-04	3.60
day=Tuesday	20.16	19.84	17.18	3.16E-03	2.95
fmicro=Occasionally	25.00	2.41	1.69	2.15E-02	2.30
fbike=Often	21.26	5.90	4.85	4.14E-02	2.04

Table 4. Cluster 13 characterization. EMEF 2018-2020

An analysis in greater depth has been applied to:

- Clusters significantly overrepresented by men samples compared to those overrepresenting by women.
- Clusters significantly overrepresented by year 2020 samples compared to those overrepresenting 2018 and 2019 samples.

For example, clusters 19, 22 and 62 show a significant higher percentage of women than men as depicted in Figure 14. The activity pattern distribution for each of these clusters can be seen in Figure 15. A significant daily activity from 15 to 21 hours is devoted to escorting activities by women combined with staying at home. On the other hand, working and other regular activities are shown from 7am to 3pm, and the highest transport activity is found at lunch time. The age-group, gender, activity, education level and transport modal preferences can be obtained by analysing its socio-economic composition.

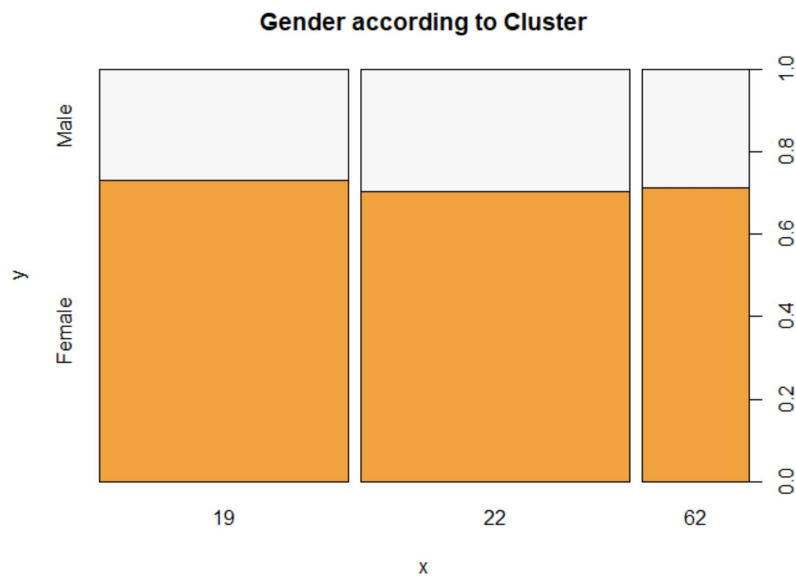


Figure 14. Women percentage in Clusters 19, 22 and 62. Daily activity pattern classification. EMEF2018 to EMEF2020 data

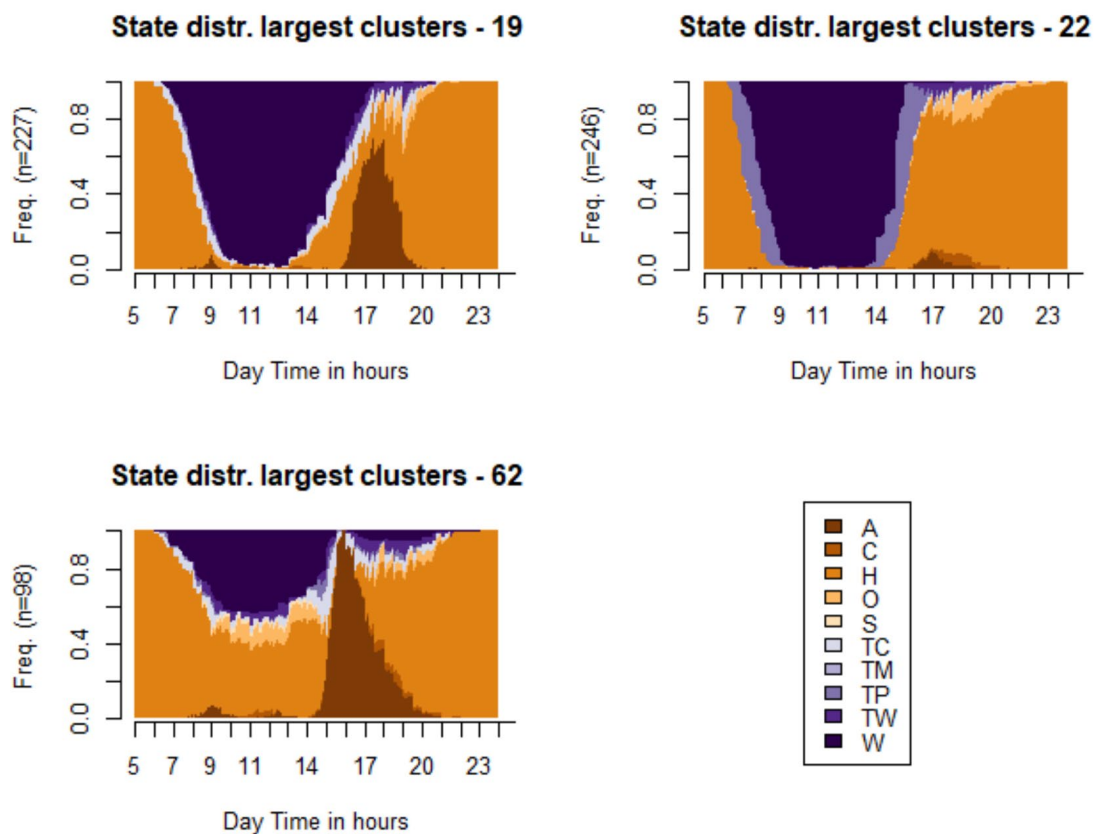


Figure 15. Classification of EMEF 2018 to 2020 daily activity patterns. The three most representative clusters over-representing women

Clusters 3, 8 and 59 show a significantly higher percentage of men than women and their activity pattern distribution, as may be seen in Figure 17. In particular, cluster 3 shows a great incidence of work activity in the first part of the day using private transport or walking as modal choices and staying at home in the afternoon with a short escorting activity and other regular activities. Age-group, gender, activity, education level and transport modal preferences can be obtained by profiling its socio-economic composition.

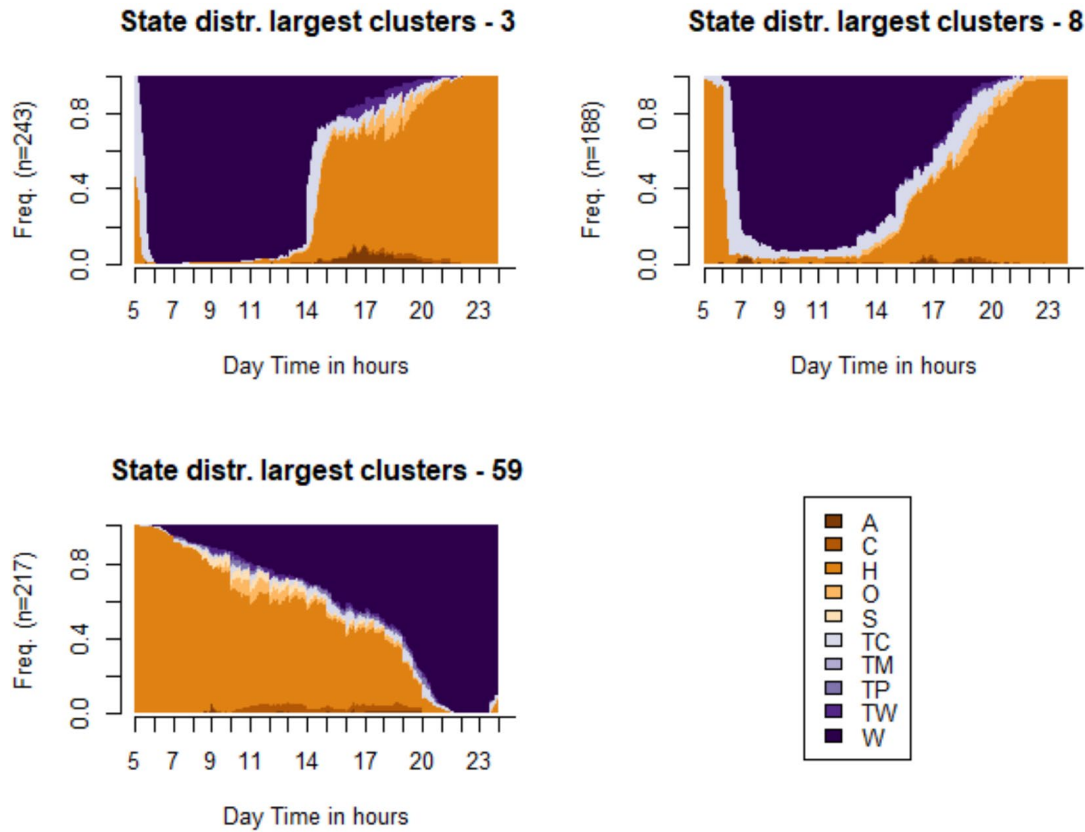


Figure 16. Classification of EMEF 2018 to 2020 daily activity patterns. Three most representative clusters over-representing men

The activity patterns and distribution in year 2020 are found in clusters 113, 123 and 151 (Figure 17). No trip-makers are the most frequent cluster (151), as it contains 60% of 2020-year samples staying the whole day at home (teleworking information is not available). The Cluster 113, groups those people walking or going by car to other frequent activities along the day. It includes escorting as no working activity, but a great part of the day is spent at home. Finally, cluster 125 represents those walking at late afternoon and visiting or other frequent activities in the morning, using neither car, nor public transport, basically to short-distance destinations.

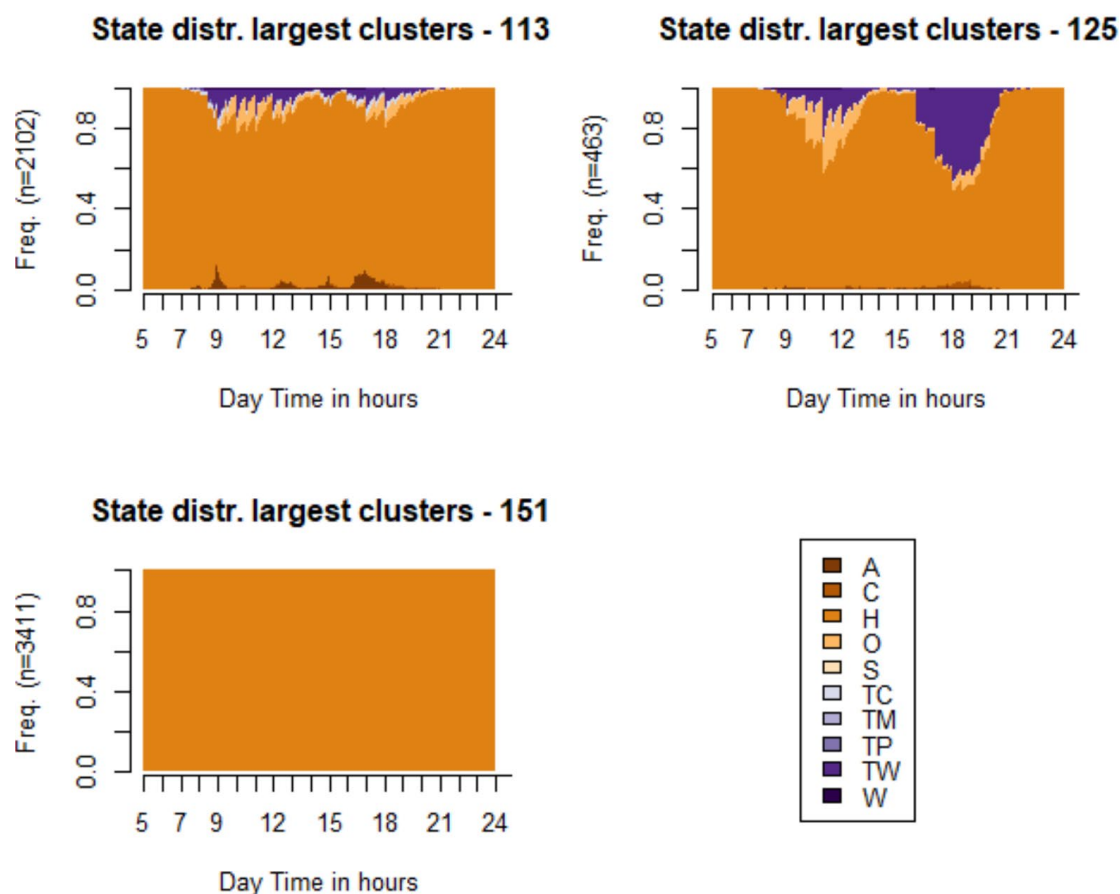


Figure 17. Classification of EMEF 2018 to 2020 daily activity patterns. The three most large clusters over-representing 2020 samples

Table 5 summarizes fragmentation indicators according to modal preferences and Figure 17 highlights a descriptive analysis for modal preferences, residential crown and year based on Multiple Correspondence Analysis.

Modal preference	Levels	nentropy	turbulence	scomplexity	TTR
fwalk	Never	0.29	9.53	0.640	0.497
fwalk	Occasionally	0.30	10.25	0.648	0.499
fwalk	Often	0.30	10.79	0.650	0.497
fwalk	Always	0.25	10.23	0.583	0.480
fbike	Never	0.27	10.31	0.619	0.489
fbike	Occasionally	0.32	11.10	0.679	0.507
fbike	Often	0.30	10.87	0.645	0.498
fbike	Always	0.29	10.38	0.637	0.494
fmicro	Never	0.28	10.44	0.628	0.492
fmicro	Occasionally	0.33	10.98	0.684	0.508
fmicro	Often	0.34	11.49	0.702	0.512
fmicro	Always	0.36	11.71	0.706	0.522
ftpub	Never	0.26	10.30	0.603	0.487
ftpub	Occasionally	0.30	10.80	0.647	0.496
ftpub	Often	0.31	10.35	0.668	0.500
ftpub	Always	0.32	10.24	0.672	0.501
fcard	Never	0.26	9.98	0.600	0.484

<b>fcard</b>	Occasionally	0.31	11.04	0.656	0.499
<b>fcard</b>	Often	0.32	11.13	0.678	0.507
<b>fcard</b>	Always	0.31	10.88	0.668	0.505
<b>fcarnd</b>	Never	0.28	10.24	0.624	0.492
<b>fcarnd</b>	Occasionally	0.30	10.77	0.651	0.497
<b>fcarnd</b>	Often	0.25	10.05	0.587	0.482
<b>fcarnd</b>	Always	0.27	10.38	0.615	0.489

Table 5. Means of fragmentation indicators according to transportation mode preference. EMEF 2018-20. Modes: fwalk stands for walking, fbike – cycling, ftpub – public transport, fcard - car as driver and fcarnd - car as non-driver.

Figure 18 presents the multiple correspondence analysis plot (Husson and Lê 2020) representing the projection in the first factorial plane for categories taking as active variables modal preferences (described in Table 5) , year and residential crown. Some remarkable associations are found:

- Modal preferences for 2020 are basically related to frequent walking, cycling, or using car as a non-driver.
- The frequent use of micromobility and public transport is usually related to Barcelona city residents.
- Frequent car use is related to external crowns to the metropolitan area.

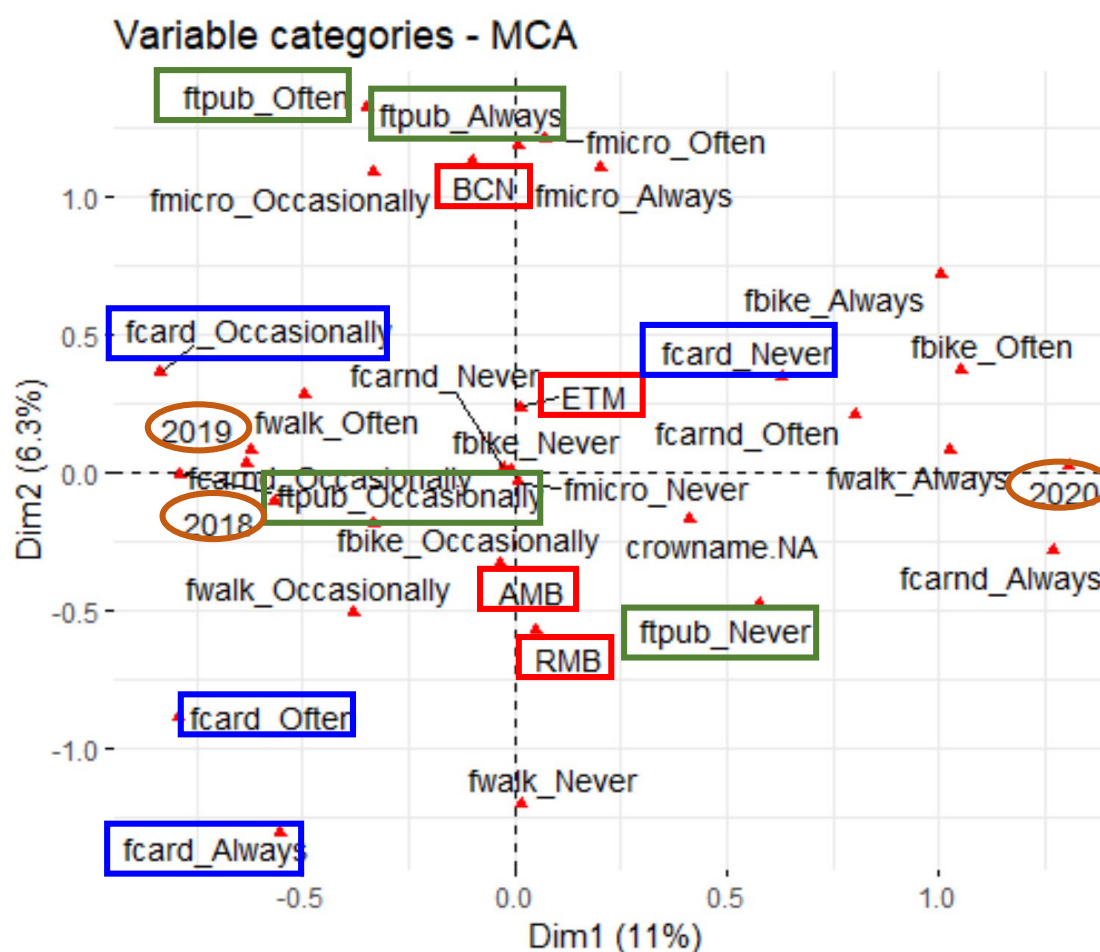


Figure 18. Modal preferences according to daily activity patterns by year and residential crown in Barcelona. EMEF 2018 to 2020. BCN – Barcelona city, ETM – Primary Crown (excluding BCN), AMB (Metropolitan Area excluding ETM) and RMB (Metropolitan Region excluding AMB).

## 6. Discussion

Comparing our results with those reported by McBride et al. (McBride, Davis, and Goulias 2019) (McBride, Davis, and Goulias 2020) may provide interesting inputs. However, there are differences in the available datasets that limit the comparison possibilities. While Barcelona's dataset for classification purposes is larger than their dataset, California's dataset is for a larger sample of households and individuals, better suited for exploratory data analysis and multivariate analysis involving fragmentation indicators, land use in residential area, individual and household characteristics. On the other hand, our activity set, defined by an alphabet consisting of 10 activities is also larger, as it contains data for 2-month for each of the three consecutive years (what captures year variation for working days) we only have data for individual, not for households, and therefore we cannot explore differences and commonalities in couples with and without children in the same household. California's data also include weekend activity patterns, whereas our dataset is restricted to working days. However, a relevant difference of Barcelona's data set with respect of California's one, is that it explicitly includes a rich information on travel modes, namely Public Transport, especially relevant when addressing the gender issues.

Barcelona's analysis, conducted with the officially available datasets (no ad hoc data collection has been possible) has been able to successfully address most of the points except for the household composition, as it is not available in the EMEF surveys.

California's results can be summarized as follows: People aged 25–65 had the most fragmented schedules (especially as measured by turbulence). Significant differences among people of different incomes were found, with key findings being the impact of poverty inhibiting activity variety for a person, and ethnicity/nativity playing a role. Gender also emerged as a major covariate for Entropy, but not for Turbulence (McBride, Davis, and Goulias 2019). The analysis also showed that specific age groups tend to have very long sequences of short activities and trips. Urban and suburban environments, however, they tend to have more fragmented schedules, most likely because of the mixing of short and long activities in their schedule. Their state thar their analysis, however, "it is not sufficient to conclusively identify people who suffer from social exclusion" (McBride, Davis, and Goulias 2019). A deeper analysis is shown in (K. Goulias 2020; K. G. Goulias, McBride, and Su 2020; McBride, Davis, and Goulias 2020) where social exclusion is claimed to happen in two ways: stay at home with little access to opportunities, which would be a measure of immobility, or to be extremely-active for the benefit of others with no personal time.

## 7. Conclusions

According to Couclelis' statement (Couclelis 2000) "the fragmentation of activities is one of the reasons for the widely observed increases in travel demand in the industrialized world", then, we believe that addressing fragmentation indicators and the potential connection to modal preferences in an European city such as Barcelona, and comparing it with a city with a totally different structure is notably relevant.

Activity analysis across days in the RMB area has been addressed in this paper. Our research is based in the work of McBride et al. (McBride, Davis, and Goulias 2020), in a totally different context, like that of the Metropolitan Area of Barcelona, where the spatial and temporal behavior of transport demand is rather diferent in terms of the underlying socioeconomic reality, the transport oriented development (TOD) which can be seen through the rich public transport network, the urban structure, and different activity alternatives and services.



Sequence analysis is used in this paper to measure fragmentation in activity participation and travel.

Studying sequences of daily activities (each activity at a place and each trip) includes the entire trajectory of a person's activity during a day while jointly considering the number of activities and trips, their ordering, and their durations. The complexity of the data resulting from fragmentation in that case led us to resort to make use of an OD dimensionality reduction and clustering techniques to conduct the analysis. The clustering analysis shows clear differences between clusters overrepresented by women and by men.

This study has also revealed some behaviours that to be properly understood require a deeper analysis, and likely more detailed data not available in this case. An example would be that of the bimodality of the turbulence.

This research leaves many doors open for future analysis. For example, a deeper analysis for all clusters.

Furthermore, as it has already been highlighted in Table 2, activity distributions by year, gender and geographic residential area show significative differences, as for instance the work activity proportion in men is higher than in women, and also the use of private transport increases in the external crowns of the Metropolitan area.

Finally, the results also show the influence of the urban environment, which clearly deserve a more detailed analysis at finer spatial scale.

## 8. Acknowledgements

This research was funded by PID2020-112967GB-C31 Spanish R+D Programs and by Secretaria d'Universitats-i-Recerca-Generalitat de Catalunya- 2017-SGR- 1749. The datasets were kindly shared by the Autoritat del Transport Metropolità (ATM). Their contribution to our research is gratefully acknowledged.

## 9. References

- Abbott, A. 1983. "Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 16(4): 129–47.  
<https://www.tandfonline.com/doi/abs/10.1080/01615440.1983.10594107> (June 6, 2022).
- . 1984. "Event Sequence and Event Duration: Colligation and Measurement." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 17(4): 192–204.  
<https://www.tandfonline.com/doi/abs/10.1080/01615440.1984.10594134> (June 6, 2022).
- Abbott, A., and J. Forrest. 1986. "Optimal Matching Methods for Historical Sequences." *Journal of Interdisciplinary History* 16(3): 471.
- Abbott, A., and A. Tsay. 2000. "Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect." *Sociological Methods & Research* 29(1): 3–33. <https://journals-sagepub-com.recursos.biblioteca.upc.edu/doi/10.1177/0049124100029001001> (February 15, 2022).
- Alexander, Bayarma, Dick Ettema, and Martin Dijst. 2010. "Fragmentation of Work Activity as a Multi-Dimensional Construct and Its Association with ICT, Employment and Sociodemographic Characteristics." *Journal of Transport Geography* 18(1): 55–64.

- von Behren, Sascha et al. 2020. "Image-Based Activity Pattern Segmentation Using Longitudinal Data of the German Mobility Panel." *Transportation Research Interdisciplinary Perspectives* 8: 100264.
- Bhat, Chandra R., and Abdul Rawoof Pinjari. 2007. "Duration Modeling." In *Handbook of Transport Modelling*, ed. Emerald Group Publishing. , 105–31. <https://doi.org/10.1108/9780857245670-006> (February 13, 2022).
- Bhat, Chandra R et al. 2013. "A Household-Level Activity Pattern Generation Model with an Application for Southern California." *Transportation* 2013 40:5 40(5): 1063–86. <https://link.springer.com/article/10.1007/s11116-013-9452-y> (February 13, 2022).
- BMDV. 2020. "Mobility in Germany." *Mobility in Germany*. <https://www.bmvi.de/EN/Services/Statistics/Mobility-in-Germany/mobility-in-germany.html> (July 15, 2022).
- Burchell, Brendan, Darja Reuschke, and Mary Zhang. 2020. "Spatial and Temporal Segmenting of Urban Workplaces: The Gendering of Multi-Locational Working:" <https://doi.org/10.1177/0042098020903248> 58(11): 2207–32. <https://journals.sagepub.com/doi/full/10.1177/0042098020903248> (May 1, 2022).
- Couclelis, Helen. 2000. "From Sustainable Transportation to Sustainable Accessibility: Can We Avoid a New." : 341–56. [https://link.springer.com/chapter/10.1007/978-3-662-04027-0\\_20](https://link.springer.com/chapter/10.1007/978-3-662-04027-0_20) (June 5, 2022).
- . 2006. "Pizza over the Internet: E-Commerce, the Fragmentation of Activity and the Tyranny of the Region." *Entrepreneurship & Regional Development* 16(1): 41–54. <https://www.tandfonline.com/doi/abs/10.1080/0898562042000205027> (February 13, 2022).
- Cresswell, Tim. 2008. "Gendering Mobility: Insights into the Construction of Spatial Concepts." In *Gendered Mobilities*, eds. T. Uteng and Tim Cresswell. London, UK: Routledge Taylor and Francis Group, 97–112.
- Cubells, Jerònia, Oriol Marquet, and Carme Miralles-Guasch. 2020. "Gender and Age Differences in Metropolitan Car Use. Recent Gender Gap Trends in Private Transport." *Sustainability (Switzerland)* 12(18).
- EIGE. 2019. "Browse Gender Statistics | Gender Statistics Database | European Institute for Gender Equality." *What lies behind the gender pay gap?* <https://eige.europa.eu/gender-statistics/dgs/data-talks/what-lies-behind-gender-pay-gap> (May 21, 2021).
- . 2022. "European Institute for Gender Equality." *European Institute for Gender Equality*. <https://eige.europa.eu/> (June 6, 2022).
- Elzinga, Cees H., and Aart C. Liefbroer. 2007. "De-Standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis." *European Journal of Population* 23(3–4): 225–50. <https://link.springer.com/article/10.1007/s10680-007-9133-7> (February 14, 2022).
- Elzinga, Cees H., and Matthias Studer. 2019. "Normalization of Distance and Similarity in Sequence Analysis." *Sociological Methods & Research* 48(4): 877–904. <https://journals-sagepub-com.recursos.biblioteca.upc.edu/doi/abs/10.1177/0049124119867849> (February 13, 2022).
- Eurostat. 2020. "How Usual Is It to Work from Home?" *EUROSTAT*. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20200206-1> (May 22, 2021).

- EWCS. 2015. "Sixth European Working Conditions Survey: 2015." *Eurofound*.  
<https://www.eurofound.europa.eu/surveys/european-working-conditions-surveys/sixth-european-working-conditions-survey-2015> (July 15, 2022).
- Fana, Marta et al. 2020. JCR Technical Reports *The COVID Confinement Measures and EU Labour Markets*. <https://ec.europa.eu/jrc> (July 8, 2022).
- Gabadinho, A, G Ritschard, M Studer, and N.S Müller. 2011. "Mining Sequence Data in R with the TraMineR Package: A User s Guide." : 129. <https://docplayer.net/7853982-Mining-sequence-data-in-r-with-the-traminer-package-a-user-s-guide-1.html> (February 13, 2022).
- Giele, Janet, and Glen Elder. 1998. Methods of Life Course Research: Qualitative and Quantitative Approaches *Methods of Life Course Research: Qualitative and Quantitative Approaches*. SAGE Publications, Inc.
- Goulias, K. 2020. *An Analysis of Accessibility, Social Interaction, and Activity-Travel Fragmentation in California in METRANS | Research Projects*.  
<https://www.metrans.org/research/an-analysis-of-accessibility-social-interaction-and-activity-travel-fragmentation-in-california> (June 10, 2022).
- Goulias, Konstadinos G., Elizabeth C. McBride, and Rongxiang Su. 2020. "Life Cycle Stages, Daily Contacts, and Activity-Travel Time Allocation for the Benefit of Self and Others." *Mobility and Travel Behaviour Across the Life Course*: 206–20.
- Hanson, Susan. 2010. "Gender and Mobility: New Approaches for Informing Sustainability." *Gender, Place and Culture* 17(1): 5–23.  
<https://www.tandfonline.com/action/journalInformation?journalCode=cgpc20> (June 25, 2021).
- Hubers, C., T. Schwanen, and M. Dijst. 2008. "ICT and Temporal Fragmentation of Activities: An Analytical Framework and Initial Empirical Findings." *Tijdschrift voor economische en sociale geografie* 99(5): 528–46. <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9663.2008.00490.x> (July 8, 2022).
- Husson, F., J. Josse, S. Lê, and J. Mazet. 2008. "FactoMineR: Exploratory Multivariate Data Analysis with R." <http://factominer.free.fr/> (June 10, 2022).
- Husson, F., and S. Lê. 2020. *EXPLORATORY MULTIVARIATE ANALYSIS BY EXAMPLE USING R*. CRC PRESS. <https://www.routledge.com/Exploratory-Multivariate-Analysis-by-Example-Using-R/Husson-Le-Pages/p/book/9780367658021> (June 10, 2022).
- Husson, F, S Lê, and J Pages. 2010. CRC Press *Exploratory Multivariate Analysis by Example Using R Analysis*. Chapman &.
- Hyde, Janet Shibley et al. 2019. "The Future of Sex and Gender in Psychology: Five Challenges to the Gender Binary." *The American psychologist* 74(2): 171–93.  
<https://pubmed.ncbi.nlm.nih.gov/30024214/> (May 28, 2022).
- INE. 2021. "Posibilidad de Teletrabajar y Teletrabajo Efectivo, de Forma Total o Parcial, Por Características Demográficas." *Instituto Nacional de Estadística*.  
[https://datos.gob.es/es/catalogo/ea0010587-posibilidad-de-teletrabajar-y-teletrabajo-efectivo-de-forma-total-o-parcial-por-caracteristicas-demograficas-identificador-api-t25-p450-base\\_2011-a2021-l0-04055-px](https://datos.gob.es/es/catalogo/ea0010587-posibilidad-de-teletrabajar-y-teletrabajo-efectivo-de-forma-total-o-parcial-por-caracteristicas-demograficas-identificador-api-t25-p450-base_2011-a2021-l0-04055-px) (July 15, 2022).
- Leszczyc, Peter T.L.Popkowski, and Harry Timmermans. 2002. "Unconditional and Conditional Competing Risk Models of Activity Duration and Activity Sequencing Decisions: An Empirical Comparison." *Journal of Geographical Systems* 4(2): 157–70.

- Levy, René, and Eric D (eds) Widmer. 2013. *Gendered Life Courses between Individualization and Standardization. A European Approach Applied to Switzerland*. Wien, Lit Verlag.
- McBride, E., A. Davis, and K. Goulias. 2020. "Exploration of Statewide Fragmentation of Activity and Travel and a Taxonomy of Daily Time Use Patterns Using Sequence Analysis in California." *Transportation Research Record* 2674(12): 38–51.
- McBride, E., A. Davis, and K. Goulias. 2019. "Fragmentation in Daily Schedule of Activities Using Activity Sequences:" *Transportation Research Record* 2673(4): 844–54. <https://journals-sagepub-com.recursos.biblioteca.upc.edu/doi/full/10.1177/0361198119837501> (February 13, 2022).
- Mejía-Dorantes, Lucía, Lidia Montero, and Jaume Barceló. 2021. "Mobility Trends before and after the Pandemic Outbreak: Analyzing the Metropolitan Area of Barcelona through the Lens of Equality and Sustainability." <https://doi.org/10.3390/su13147908> (December 12, 2021).
- Paleti, Rajesh et al. 2017. "Activity Sequencing, Location, and Formation of Individual Non-Mandatory Tours: Application to the Activity-Based Models for Columbus, Cincinnati, and Cleveland, OH." *Transportation* 44(3): 615–40. <https://link-springer-com.recursos.biblioteca.upc.edu/article/10.1007/s11116-015-9671-5> (February 13, 2022).
- R Development Core Team. 2021. "R: The R Project for Statistical Computing." <https://www.r-project.org/> (February 13, 2022).
- Rasouli, Soora, and Harry Timmermans. 2014. "Activity-Based Models of Travel Demand: Promises, Progress and Prospects." *International Journal of Urban Sciences* 18(1): 31–60. <https://www.tandfonline.com/doi/abs/10.1080/12265934.2013.835118> (February 13, 2022).
- Studer, M., and G. Ritschard. 2016. "What Matters in Differences between Life Trajectories: A Comparative Review of Sequence Dissimilarity Measures on JSTOR." *Journal of the Royal Statistical Society A* 179(Part 2): 481–511. <https://www.jstor.org/stable/43965553> (May 1, 2022).
- Tukey, John W. (John Wilder) et al. 1984. *The Collected Works of John W. Tukey*. Belmont Calif.: Wadsworth Advanced Books & Software.
- Wilson, W. C. 1998. "Activity Pattern Analysis by Means of Sequence-Alignment Methods:" *Environment and Planning A: Economy and Space* 30(6): 1017–38. <https://journals-sagepub-com.recursos.biblioteca.upc.edu/doi/10.1068/a301017> (February 15, 2022).