Case Studies on Transport Policy 18 (2024) 101287

Contents lists available at ScienceDirect

Case Studies on Transport Policy

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

Application of cell phone data to monitor attendance during motor racing major event. The case of Formula One Gran Prix in Imola

Alessandro Nalin [a,*], Andrea Simone [a], Claudio Lantieri [a], Denis Cappellari [b], Glauco Mantegari [b], Valeria Vignali [a]

[a] DICAM Department, School of Engineering, University of Bologna, Italy
[b] Motion Analytica, 33100 Udine, Italy

ARTICLE INFO

Keywords:
Big data
Major sport event
Transport-related policies

ABSTRACT

The advent of connected devices, such as smartphones, has had a transformative impact on the landscape of recent years. Once privacy concerns have been addressed, data can be handled and analysed in a proficient manner to gain insights into patterns and movements, thereby influencing urban policies. It is likely that mobility and transport-related topics have been the subject of the most extensive investigation in the field of cell phone big data. While the topic of commuting patterns has been extensively researched by numerous authors, there is a paucity of literature on the monitoring of attendance during major motorsport events. Despite the predictability of crowding (tickets are sold in advance and the schedule is fixed and rigid), multiday motorsport events are disruptive in terms of traffic, overcrowding and uneasiness for hosting cities. This paper aims to address the aforementioned gap by presenting a case study of monitoring attendance during the Formula One Emilia-Romagna and Made in Italy Grand Prix, held in Imola, Italy, from 22nd to 24th April 2022. The results demonstrated the potential of data to inform the prediction of mobility choices and the planning of appropriate mobility-related policies, with the aim of reducing the impact of future events. This represents a significant challenge for public administrations and stakeholders.

1. Introduction

The pervasive presence of connected IoT sensors and devices in urban environments affords private and public authorities the opportunity to collect and process data, thereby enhancing policies in a range of fields (Balduini et al., 2019), such as road safety (Battistini et al., 2023; Ghasemi et al., 2020; Kontaxi et al., 2021; Mohammadi et al., 2019; Warren et al., 2019) and transport-related issues, e.g., reduction of congestion and parking policies (Neilson et al., 2019; Xiao & Lou, 2018). The term big data is typically used to describe this source of information. The term can be applied to a variety of data types, including GPS traces (Battistini et al., 2022; Liguori et al., 2017; Shoman et al., 2023), traffic loop records, naturalistic driving data (NDD) (Carneiro Pereira & Prevedouros, 2022), Bluetooth traces, and social media contents. A specific typology of big data is related to mobile phone records. Such data is automatically recorded by telecommunications operators for the purposes of billing, management and network maintenance (H. Huang et al., 2021; Steenbruggen et al., 2015). Data from mobile phone network can be categorized into two typologies, depending on their generative process (Calabrese et al., 2015; Chin et al., 2019; H. Huang et al., 2021; Sadeghvaziri et al., 2016; Wang & Chen, 2018a, 2018b) and their spatiotemporal resolution (Bonnetain et al., 2019). Event-driven data are generated in response to specific events, such as a call, a SMS, or an Internet access, while network-driven data result from the continuous interaction between a connected device (e.g., a smartphone) and the telecommunications network, independently from any user-conscious actions. It can be argued that the latter typology is denser than the former, both in terms of space and time, although the coverage area may be affected by the surrounding environment. As observed by Chin et al. (2019), in non-urban contexts, the coverage area is typically wider, which may result in a less precise spatial resolution and subsequent localization process. In general, the production of ubiquitous, geographically and temporally referenced information amounts to billions of records per individual user on a daily basis (Chin et al., 2019). The potential drawbacks and limitations of this pervasive technology have been matter of research (M. Chen et al., 2014; Nalin et al., 2024;

* Corresponding author.
E-mail address: alessandro.nalin2@unibo.it (A. Nalin).

https://doi.org/10.1016/j.cstp.2024.101287
Received 25 October 2023; Received in revised form 19 August 2024; Accepted 27 August 2024
Available online 28 August 2024

Neilson et al., 2019; Parsafard et al., 2019; Pozdnukov & Waddell, 2016), and such factors may impact the analysis and interpretation of results if the validation process is not properly addressed (Fekih et al., 2021). Among others, the main discussions are related to the form of data, specifically that big data is typically unstructured (Gandomi & Haider, 2015), the integration with other source of data (Yang et al., 2020), either official (Ricciato et al., 2017) or not, the presence of data noise (Scannapieco et al., 2013) and the ping-pong effect (Fiadino et al., 2012; Zahedi & Shafahi, 2018). Nevertheless, mobile phone data is widely acknowledged as a valuable resource for gaining insight into how individuals behave in terms of mobility, consumption and environmental impacts. It offers a more comprehensive understanding of urban phenomena than traditional survey methods, such as travel surveys, which often have limited scope (Caceres et al., 2020; Calabrese et al., 2015; Cik & Fellendorf, 2019). Furthermore, major events offer case studies due to their capacity to attract considerable numbers of people. While some of these, such as the Olympic Games and FIFA World Cup venues, have been subjected to detailed analysis in terms of spectator flows, recurring multi-day motorsport events such as MotoGP or Formula One (F1) venues have not been analysed yet. Formula One (F1) is the most renowned single-seater car motorsport category (Gerard-Reimer, 2020), and each team competes with two drivers racing two cars (Aversa et al., 2015), in a growing environmental-aware championship (Mourao, 2018). From the spectators' point of view, as demonstrated elsewhere (Mastromarco & Runkel, 2009), technological changes affect the competitive balance. The mentioned changes can therefore influence the satisfaction of fans (Schneiders & Rocha, 2022) and the number of attendants at the venues.

In light of the aforementioned premises, the objective of this paper is to illustrate the capacity of mobile phone data to identify attendees at a multi-day, recurring event held annually, such as a F1 race weekend, and to discern their behaviours. The results will be of interest to those responsible for the planning and organisation of transport services for participants, with the aim of minimising disruption to the transport network and services. Given that the majority of F1 venues are situated in urban or *peri*-urban settings, this is a matter of particular significance. In contrast, F1 venues situated in non-urban or rural locations have been subject to criticism for lacking convenient access to essential facilities, leading to their replacement by alternative venues in urban areas. A comprehensive examination of the behaviours exhibited during the 2022 F1 Grand Prix held in Imola, Italy, will be conducted through the utilisation of a set of cell phone data from one of three principal mobile phone operators within the Italian market. The paper is organized as follows: Section 2 will provide a literature review about mobile phone data applications. In Section 3 a brief overview of case study context (i.e., Imola circuit and the schedule of the F1 weekend) is drawn. In Section 4, applied methodology will be described. In Section 5, a detailed overview of results will be presented, and it will be discussed in Section 6. Section 7 contains conclusions with possible future research streams.

2. Literature review

Several applications of mobile phone data can be identified across a diverse range of fields. In general, mobile phone data was able to trigger new research streams in several fields, as it enabled new methodological approaches in studying the behaviours of people, either visualized as the space–time prism or the space–time cube (Thevenin & Vuidel, 2021). Such applications include disaster response and recovery (Yabe et al., 2022), urban vibrancy detection (Zhang et al., 2021), the analysis of mobility patterns (Wang et al., 2017; Yang et al., 2020) with different levels of regularity (Wang & Chen, 2018a), the distribution of people in relation to urban points of interest (H. Chen et al., 2021) and activity patterns (Widhalm et al., 2015; Zahedi & Shafahi, 2018), and the correlation analysis of tourism and transportation facilities (Qian et al., 2021). It can be argued that the field of mobility-related research represents the most extensively investigated area of cell phone data applications, considering the ability of predicting and modelling mobility patterns (Pozdnukov & Waddell, 2016). COVID-19 outbreak was a turning point, as mobile phone data has been considered a valid tool to trace people behaviours and hence the virus spread (Hasselwander et al., 2021). The research interest on mobility has concentrated on commuting patterns at both the macro level (Caceres et al., 2007; Doyle et al., 2014; Fekih et al., 2021; Hadachi et al., 2020) and the urban level (Sevtsuk & Ratti, 2010), with some focuses on route choice behaviour modelling (Bwambale et al., 2019). Furthermore, the topics of identifying the mode of transport and the multimodal itineraries have been introduced as a subject for investigation (Berlingerio et al., 2013; Bonnetain et al., 2019; H. Huang et al., 2019; Wischer et al., 2023), with a focus on collective means of transportation (Demissie et al., 2016; Ma et al., 2018). Moreover, as Z. Huang et al. (2018) observed, some research has focused on the integration of disparate data sources to facilitate the study of human space–time activity patterns, such as GPS phone data (Afandizadeh Zargari et al., 2021; Sadeghvaziri et al., 2016).

The use of mobile phone data records in monitoring pre- and post-event flows, is evidenced by a body of existing literature including cultural, musical, and sporting events (Altin et al., 2022; Bao et al., 2023; Calabrese et al., 2010; Cik et al., 2017; Mamei & Colonna, 2016; Marques-Neto et al., 2018). With specific regards to sport events, the majority of attention has been focused on recurring events, such as soccer matches during the regular season of championships, or major events, such as the Olympics (González et al., 2017; Pintér & Felde, 2021; Ponieman et al., 2015; Xavier et al., 2012). This is due to the fact that the legacies and impacts of similar events are typically considered by analysts and administrators. A further area of potential interest is the examination of other major sporting events, such as motor racing Grand Prixes (GPs), with a particular focus on Formula One (F1) venues. It is beyond doubt that they can be classified as major sport events (Chamberlain et al., 2019; Kim et al., 2017; Lefebvre & Roult, 2013; McCartney, 2005; Müller, 2015; Roult et al., 2020; Smith, 2012), and that the organisational phase of a GP is a matter of concern for public authorities and administrators (Gezici & Er, 2014; Lefebvre & Roult, 2013; Liu & Gratton, 2010; Lowes, 2018; Mourão, 2017), as they seek to be competitive against competing candidate venues (Lefebvre & Roult, 2011). Analyses about economic legacies have been conducted in both European and non-European countries (Fairley et al., 2011; Kim et al., 2017; Storm et al., 2019). In addition, it is important to consider other aspects of major motorsport events that have been identified in the literature. These include the rules and procedures that privatise the urban space several days before the race (Roult et al., 2020) and the number of non-resident individuals involved, either directly or indirectly. These include spectators, visitors, members of the press, marketing personnel and members of the competing teams.

A general observation arising from the literature review is that research about F1 GPs has identified social and economic aspects of venues, whereas the analysis of mobility patterns of spectators remains an untapped area of investigation. The involvement of thousands of people is the most tangible and intuitive evidence of a racing weekend, and this is the only significant similarity with other major sporting events. However, while events such as the Olympics and the FIFA World Cup are regarded as 'once-in-a-lifetime events' (Ramasamy et al., 2020) and the transportation supply at hosting locations is usually addressed (Pereira, 2018), F1 GPs are recurrent events. Consequently, F1 venues represent a notable case study in terms of crowd monitoring, given that they are typically held on an annual basis. This implies that the event itself can have a disruptive effect on the transportation network (Kwoczek et al., 2014; Leilei et al., 2012), even if for a short period (typically less than a full week). However, externalities can be adequately assessed in preparation for the subsequent event (McCullough et al., 2020; Preuss, 2006).

3. Context description

A typical Formula One (F1) weekend is characterised by several competitive events. Some free practice sessions are held on Fridays and Saturdays. During the free practice sessions, the drivers are permitted to test the set-ups and additional mechanical or aerodynamic variations. On Saturday afternoons, drivers compete to obtain the fastest laps and the pole position during the qualifying session, which determines the race starting grid. In 2021, an additional short race (lasting 100 km) was introduced on some weekends, called 'Sprint' race. The main race takes place on Sunday afternoons. Collateral categories, such as Formula 2 (F2), Formula 3 (F3) and Porsche SuperCup, are held during F1 weekends. The schedule of the agonistic events is reported in Table 1. For the first time, the 'Sprint' race was held in Imola, so the traditional agonistic activity was rearranged. Consequently, Friday became an appealing agonistic day due to the modified programme. Furthermore, the schedule for the qualifying session on Friday (17:00–18:00 local time) was modified to optimise television audience and attendance by spectators on weekdays. With regard to weather conditions, Friday and Saturday were characterised by rainfall, while Sunday was characterised by sunshine. The abovementioned details are useful for the results' interpretation, so they will be recalled in Section 5.

Imola was the first European venue and the fourth in general in 2022 F1 calendar. The racetrack was constructed in the 1950 s and has hosted F1 races since the 1980 s. Following the tragic events of the 1994 racing weekend, during which Roland Ratzenberger and Ayrton Senna lost their lives during the qualifying session and race respectively, the track layout was significantly altered to enhance safety. Following a 15-year hiatus, the F1 returned to Imola in 2020. The Imola racetrack is homologated by the FIA as a first-category circuit and has achieved the third "three star" level of the FIA Environmental Accreditation Program. The Autodromo (see Fig. 1) has permanent stands with a capacity of 35,000 spectators and some large lawn areas. In addition, during major events, the installation of additional stands is planned. Fig. 1 shows the study area and the location of the grandstands around the racetrack. The circuit is located in close proximity to the city centre of Imola. Consequently, the F1 venue may be considered an urban event, and several traffic control measures have been implemented to mitigate the impact of the event on local traffic, including temporal traffic closures and the pedestrianization of the city centre (Comune di Imola, 2022b). Furthermore, temporary parking areas have been realized in the proximity of the racetrack (Comune di Imola, 2022a).

Table 1
Schedule of Gran Premio del Made in Italy e dell'Emilia–Romagna.

Friday (22/4/2022)	
Time	**Event**
9:55–10:40	F3 practice session
11:05–11:50	F2 practice session
13:30–14:30	F1 first practice session
15:00–15:30	F3 qualifying session
15:55–16:25	F2 qualifying session
17:00–18:00	F1 qualifying session*
18:30–19:15	Porsche SuperCup practice session*
*: Due to bad weather, this session was delayed	
Saturday (23/4/2022)	
Time	**Event**
10:35–11:20	F3 sprint race
12:30–13:30	F1 second practice session
14:00–14:30	Porsche SuperCup qualifying session
16:30–17:00	F1 sprint race
17:55–18:45	F2 sprint race
Sunday (24/4/2022)	
Time	**Event**
8:50–9:40	F3 feature race
10:20–11:25	F2 feature race
12:05–12:40	Porsche SuperCup race
13:00–13:30	F1 drivers' parade
14:44–14:48	National Italian anthem; Frecce Tricolori flight over
15:00–17:00	F1 race, followed by podium ceremony

4. Materials and methods

The analysis was based on data provided by Vodafone Italia, one of the three main mobile phone operators in Italian market. With regards to the data typology, they could be ascribed among the network-driven events mentioned in Section 1. According to AGCOM, the Italian authority for the communication, the monitored sample consisted in 20 million SIMs (Subscriber Identity Module) approximately, which corresponds to one third of Italian population, and foreign users when are roaming-connected to Vodafone Italia network. The data processing is in full compliance with the national legislation on data protection and the provisions of the GDPR regulation. The workflow steps of analysis are presented in Fig. 2.

In general, telecommunications analytics are based on network logs, which are records of the position of mobile phones in relation to the antennas to which they are connected. The objective is therefore to assign the user to a known location, and different techniques are available (Calabrese et al., 2015). The antennas are georeferenced and provide one or more coverage areas. With regard to the technical aspects of this analysis, further details are provided in Table 2.

It is important to note that the coverages typically overlap to avoid areas that are not adequately covered and are optimized using simulation software that can estimate ideal coverage areas through the use of heuristics. In order to pursue the scope of this analysis and to best render the attendants' behaviours, the antennas reported in Table 2 were classified in accordance with their position and coverage areas. The classification was as follows: (a) some antennas covered the effective circuit area, (b) some antennas located within Imola have been used to distinguish between resident and non-resident spectators, (c) the antennas located outside Imola have been employed to detect the provenance of visitors. Prior to proceeding, a pre-processing phase was completed in order to manage the hysteresis cycles, which have the potential to affect the final estimation of location. Moreover, to prevent erroneous redundancy (e.g., double counting) or misleading estimates, non-mobile phone SIMs, such as those used for the Internet of Things (IoT), have been excluded from the dataset. With regard to Fig. 1 and Table 2, it is noteworthy that the circuit was covered by the permanent telecom network, and no additional supporting antennas were installed. This limitation of the study design precludes the possibility of splitting the study area into smaller partitions, such as the main entrances and grandstands.

It is therefore necessary to consider the reported value as an aggregated figure. Although this approach may result in the loss of information regarding the detailed mobility patterns of spectators, it allows for the delineation of homogeneous groups that are sufficiently large to avoid being masked due to privacy and time thresholds. The processed data are of high resolution in both space and time and pertain to interactions conducted via the Packet Switch (PS) protocol. The protocol allows for the collection of data at an average frequency of between two and three minutes per user per day, which equates to approximately 40 million records per minute. PS logs can be utilised to ascertain certain information, including the unique user ID, which is inherently anonymised, the timestamp, and the unique antenna ID. The interactions can be classified into three principal categories: single and georeferenced records, which are captured in real time and originate from the traffic between cell phones and the network; data usage by cell phones; and other types of events that are not included in the aforementioned categories. Prior to undergoing processing, the data underwent irreversible anonymisation and aggregation. Following the anonymisation and aggregation of the datasets, they were expanded to the entire population through the use of an inferential model and appropriate calibration algorithms. The algorithms took into account a number of factors,

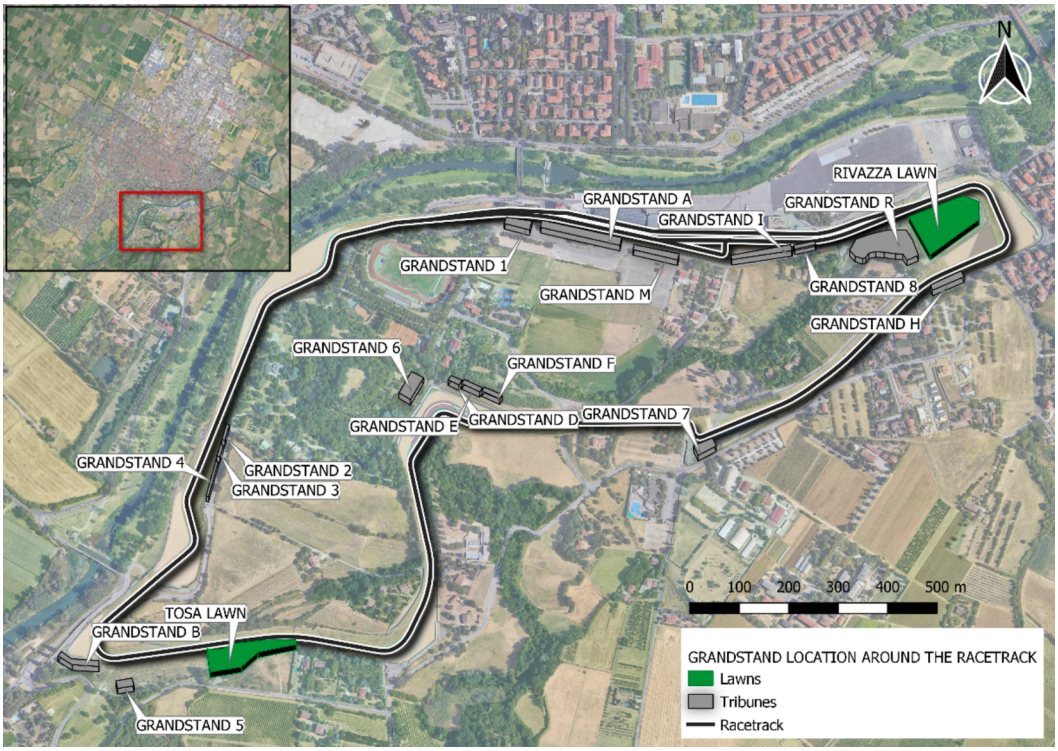

Fig. 1. Grandstand location map. Own elaboration.

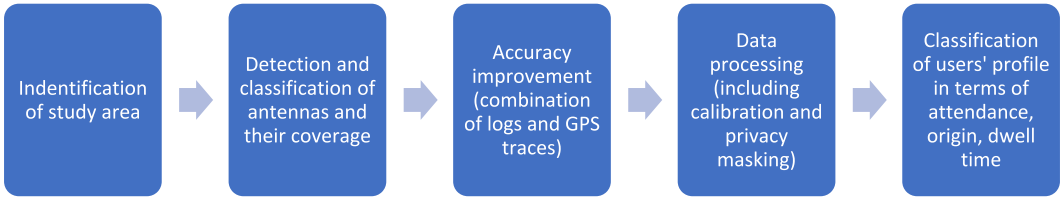

Fig. 2. Analysis workflow.

Table 2
Main details about mobile phone network coverage in the study area (Imola + Autodromo).

Technology	Number of masts	Number of antennas	Average radius (m)
2G	21	53	720 *m*
4G	36	207	650 *m*

including the market share of the operator (i.e., local by place and age class; national for foreign SIMs that are roaming on the network; by business or consumer SIM type) and socio-demographic characteristics of users (i.e., official statistics such as the National Statistics Institute–ISTAT census data). The data allowed for the identification of analogous and homogeneous mobility behaviours, thus facilitating the construction of user profiles. A significant variable is the minimum temporal threshold above which a mobile phone user detected within a given area can be classified as an event attendee or not. This parameter has the potential to influence all subsequent analyses, as it determines the distinction between individuals who should be considered event attendees and those who should be considered passers-by. The following sections provide a concise overview of the configurations utilized during the analysis.

4.1. Attendance

The principal objective of the analysis was to monitor all individuals present at the event, including spectators, personnel and members of the press, within the designated study area. The attendance analysis permitted an investigation of users' behaviours with respect to the dynamics of their participation, including arrival and departure times and the time spent at the event. Furthermore, the analysis permitted an investigation of the geographical origins of users, which were categorised as either national or international. The time threshold was set at 60 min of consecutive time spent within the analysed area, namely the circuit. Individuals who did not meet the minimum 60-minute threshold within the defined area were excluded from the dataset.

4.2. Origin

The analysis provided an overview of the attendees' origin and revealed the international and national scale of the event. With regard to Italians, a retrospective analysis of the 12 months preceding the event was conducted to ascertain the location where each user had spent the majority of their nights. This resulted in a precise, yet anonymised, indication of the places where the spectators had previously resided. It should be noted that this does not necessarily correspond to the registered residence. Consequently, it represents a valuable and ground-truth information source for studying, for instance, the distance an attendee is

willing to travel to participate in the event. Foreign nationals are allocated to a specific country based on the Mobile Country Code (MCC) of their SIM cards, which constitutes part of the metadata available when they connect to the Vodafone Italia network. Moreover, this analysis permitted the distinction between residents of Imola and other users who were categorised as tourists, defined as individuals who spent at least one night within the study area. To include spectators who did not spend any night within the study area and, therefore, could not be labelled as residents nor tourists, a new category, that of daily visitors, was introduced. This classification was the only feasible option available, given the lack of additional data that could be used to distinguish between fans, media reporters, and teams' workers. Additionally, as a preliminary constraint, analyses were conducted using a single day as the time reference, which prevented the investigation of individuals who attended more than one day.

4.3. Arrival and departure

The analysis of arrivals and departures is designed to provide insights into the general attractiveness of the event and the dynamics of inflow and outflow. The result of this analysis is a distribution of arrival and departure times, classified into 30-minute slots and divided according to the profiles of users, including residents and tourists. This classification is in accordance with the definitions set forth in Section 4.2. Users were allocated a specific time slot for arrival and departure based on the single first or last moment in which they were registered in the area, respectively.

4.4. Dwell time

The dwell time analysis was developed as a method for further examining the appeal of the event and its success. The result was a distribution of several dwell times. The initial category of dwell time, which spans a duration of between one and four hours, was deemed appropriate for the identification of attendees at F1 events. The second category, which lasts between four and six hours, was designed to identify individuals who arrived at the event with significant advance notice and thus attended prior events. The third category, which lasts between six and eight hours, was designed to identify attendees who participated in most the event's activities. The fourth category, which lasts for more than eight hours, was designed to identify attendees who participated in the entire event as spectators or insiders. This category includes members of the event's organising crew, journalists, and other relevant personnel.

5. Results

The diverse range of metrics and aggregation techniques employed in the analysis of the dataset permitted the generation of a variety of analyses. The following paragraphs will present the main results, detailing the trends and offering a comprehensive overview of attendance, arrival, departure and dwell times at the circuit, with respect to the schedule reported in Table 1.

5.1. Trends in attendance

Overall presences are reported in Fig. 3 and accounted on Friday 36,000 people, 61,000 on Saturday and 89,000 on Sunday, totalling 186,000 spectators, while Fig. 4 reports the cumulative attendant amount. It is worth noting that, while the previous sums all the attendees, both spectators and workers, and could be considered as the sum of all those who have been detected for at least 60 min within the study area, the latter should be interpreted as the 'effective' presence per each time slot (i.e., sum of arrivals and departures per each 30-minutes interval, as described in Section 4.3). According to the analysis by spectator categories in Fig. 3, "Inhabitant" labelled spectators amounted

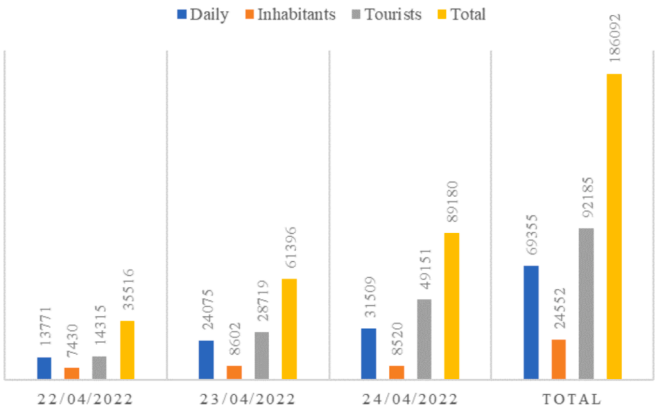

Fig. 3. Presences by day.

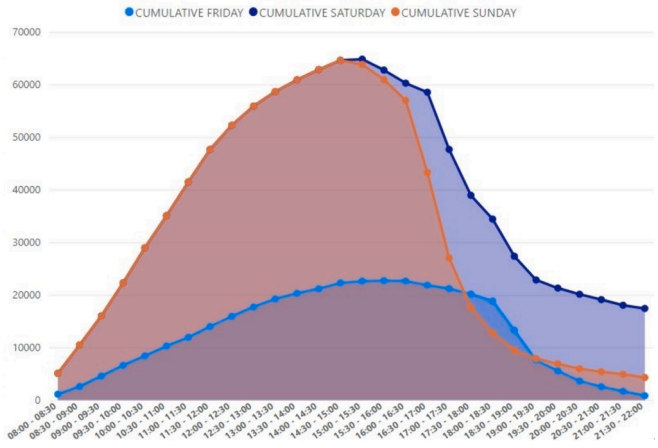

Fig. 4. Cumulative attendees (sum of arrivals and departures per time slot).

13 % of the overall crowd and they attended as follow: 7,430 people on Friday (21 %), 8,602 on Saturday (14 %) and 8,520 on Sunday (9.5 %). Unlike inhabitants, presences labelled as 'tourists' and 'daily visitors', grew with a higher increasing rate. With regards to Fig. 4, a noteworthy overlap in trends could be detected on Saturday and Sunday until early afternoon, while the Sunday drop at 16:30–17:00 time slot should be interpreted as the end of race.

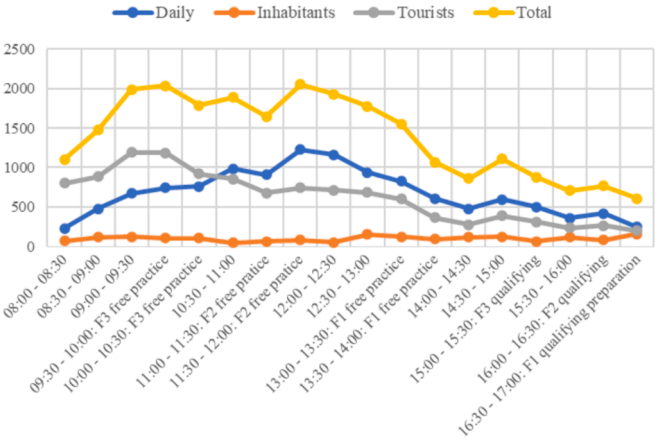

Fig. 5. Arrival time at circuit; spectators' profile–Friday.

5.2. Overview of arrival time at circuit

Fig. 5 shows the arrival time detailed by day, spectators' profile and time slots on Friday. On the first day of agonistic activities, the most–-frequented time slots were 11:30–12:00 (2,055 users, 90 min before the start of F1 free practice session), followed by the slot 9:30–10:00 (2,036 users, during the F3 free practice session) and 9:00–9:30 (1,987 users, time slot without any scheduled agonistic activity). On Saturday the most frequented slots were 10:00–10:30 (4,421; just before the F3 Sprint Race start), 10:30–11:00 (4,364; during the F3 Sprint Race) and 9:30–10:00 (4,109; time slot with no agonistic activity) (Fig. 6). On Sunday the three slots with the most entries were 10:00–10:30 (6,687; just before the start of F2 Feature Race), 11:00–11:30 (6,418; end of F2 Feature Race) and 9:30–10:00 (6,278; during the F3 Sprint Race) (Fig. 7).

5.3. Overview of departure time at circuit

On Friday time slots with the highest value of registered users were 19:00–19:30 (5,731; during the Porsche SuperCup free practice session) and 18:30–19:00 (5,553; after the F1 qualifying session, which ended late due to some suspensions) (Fig. 8).

On Saturday the most–frequented time slots were 17:00–17:30 (10,877; at the end of the 'Sprint' race, lasted just 30 min and won by Red Bull driver Max Verstappen followed by Charles Leclerc on Ferrari and Sergio Perez on Red Bull), 17:30–18:00 (8,723; in the immediate vicinity of the F2 Sprint Race start) and 18:30–19:00 (7,076; at the end of F2 Sprint Race) (Fig. 9). With details on attendant profile, 22 % of 'daily' spectators (5,530) left the track at the end of the F1 'Sprint' race, while another 2,914 left during the 18:30–19:00 time slot. Similar attitude, albeit with different proportions, was detected for 'tourists': 4,541 (15 % of the total) left at the end of the F1 'Sprint' race and 3,864 during the 18:30–19:00 time slot. On the other hand, there has been a significant drop in departures in the 18:00–18:30 time slot, corresponding to the F2 Sprint race. With detail on the geographical origin, adding the 'inhabitants' and Italian profiles, the most–frequented time slot was 17:00–17:30 (8,275), followed by 17:30–18:00 (6,697) and 18:30–19:00 (4,897). Similar trend for "foreigners": 2,602 departures in the 17:00–17:30 range, 2,179 in the 18: 30–19:00 range and 2,026 in the 17:30–18:00 range. Sunday outflows was reported in Fig. 10. Timeslots with the highest number of departures from the track were 17:00–17:30 (16,219, after the end of the F1 race and the podium ceremony), followed by 16:30–17:00 (13,732, immediately after the chequered flag) and 17:30–18:00 (9,432). Focusing on Fig. 11, which highlights the time slots during the race, and considering the sum of 'inhabitants' and Italians profiles, which could be supposed to support Ferrari in majority, the time slot 17:00–17:30 counted 13,243 departures, while 16:30–17:00 slot 11,145. In addition, an increase in departures was detected during the time span 16:00–16:30, when the Ferrari's driver

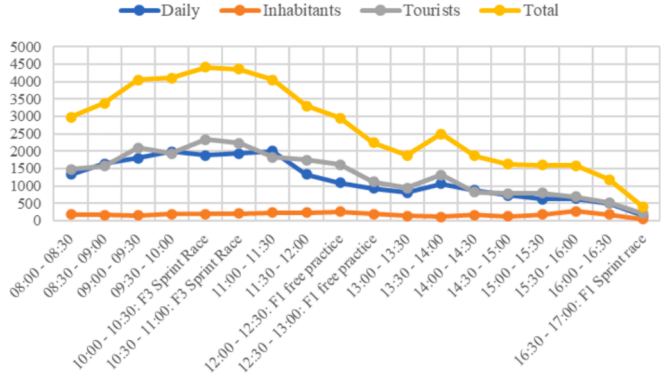

Fig. 6. Arrival time at circuit; spectators' profile–Saturday.

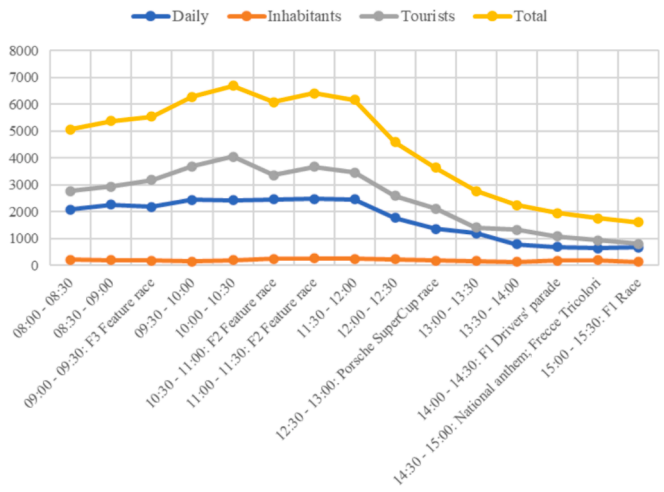

Fig. 7. Arrival time at circuit; spectators' profile–Sunday.

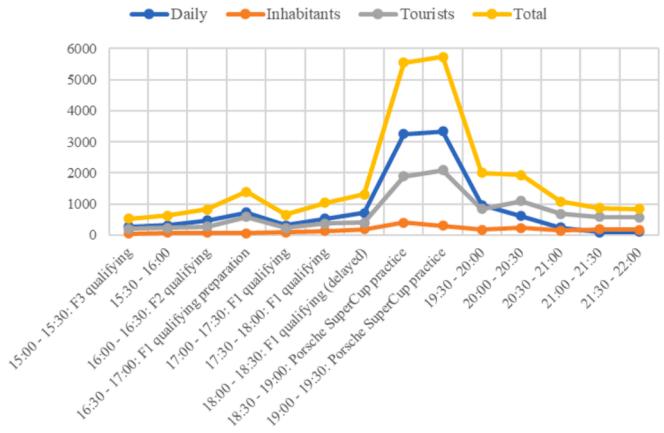

Fig. 8. Departures from the circuit; spectators' profile–Friday.

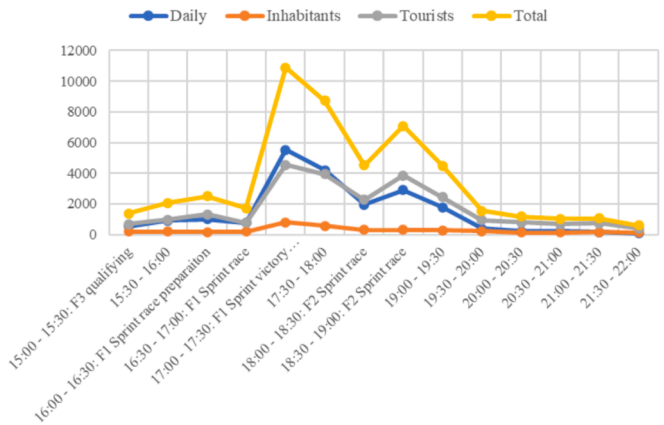

Fig. 9. Departures from at circuit; spectators' profile–Saturday.

Charles Leclerc crashed.

5.4. Dwell time at circuit

Dwell time analysis can offer some prospects about the attendance preferences, the willingness to staying within the circuit and the 'loyalty' of spectators to the agonistic events, which can be used to marketing purposes (e.g., merchandising, collateral events such as fan zone

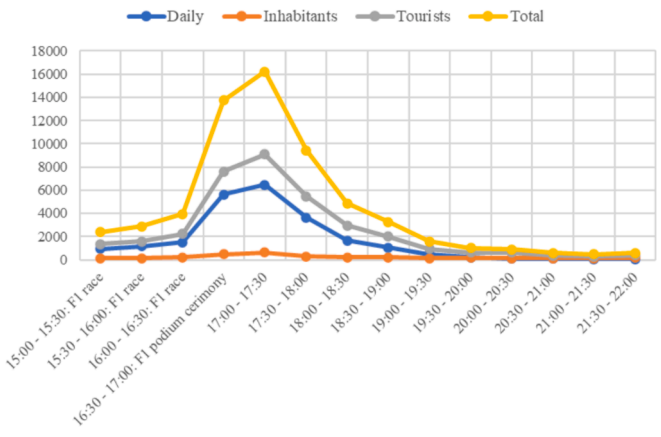

Fig. 10. Departures from at circuit; spectators' profile–Sunday.

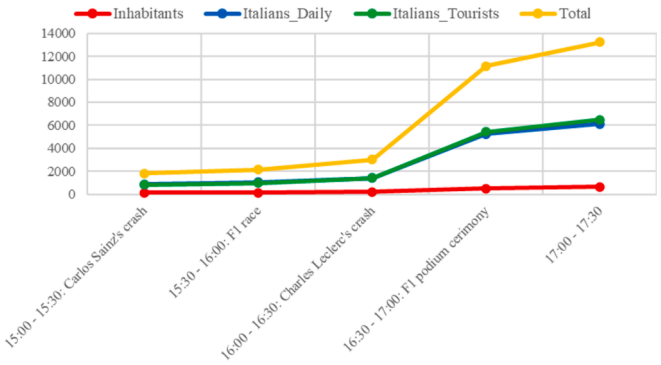

Fig. 11. Departures from circuit; focus on as-supposed Ferrari's fans–Sunday.

sizing and placement, etc.). Even if these topics are not related to the main focus of the paper, an overview of dwell time will be briefly reported as follow. By computing dwell time with arrival and departure time, different thresholds were applied (Table 3): 1 to 4 h; 4 to 6 h; 6 to 8 h; More than 8 h. Time classes are described in Section 4.4.

In detail, on Friday, 15,423 spectators (equal to 43 % of the presences registered on the day) attended the racetrack for at least 8 h. Despite the working day, 4,320 'inhabitants' (58 % of the global number for this profile) remained in the circuit for at least 8 h. As regards the geographical origin, foreigners ('daily' + 'tourists') remained in the racetrack in proportion more than Italians: 45 % for at least 8 h compared to 35 % of Italians. It should be noted that on Friday, due to the prohibitive weather conditions, the F2 free practice session was cancelled, therefore the most assiduous spectators suffered a long interruption from 10:40 (end of F3 free practice session) at 13:30 (start of first F1 free practice session). With regards to Saturday, spectators who spent at least 8 h in the circuit were 22,780, equal to 37 % of the total recorded on the day; one out of two (11,097) was 'tourist', while just under one out of four (6,974) was 'daily'. As regards the geographical origin, 34 % of Italians against 32 % of foreigners spent at least 8 h in the racetrack. On Sunday, 28,881 spectators attended the activity on the track for a period of time between one hour and four hours (32 % of the total) against 25,214 assiduous (28 %). compared to the other days, spectators of this time interval constituted 26 % of the total on Friday and 27 % on Saturday. However, the values are uneven if we consider the 'daily' and 'tourists' profiles for Italians and foreigners. Overall, 'tourists' foreigners, as an absolute value ten times the 'daily' foreigners, attended a period of time including between one hour and four hours were 7,276 (39 % of the total for the day), while 'daily' Italians who remained at the racetrack for a period of time between one hour and four hours were 10,133 (34 % of the total).

Table 3
Dwell time comparison with regards to day and spectators' profile.

Friday (22/4/2022)				
Dwell time	**Daily**	**Inhabitants**	**Tourists**	**Total**
1–4 h	4,276 (45 %)	1,930 (20 %)	3,302 (35 %)	9,508 (100 %)
4–6 h	2,573 (56 %)	538 (12 %)	1,493 (32 %)	4,604 (100 %)
6–8 h	3,326 (55 %)	643 (11 %)	2,055 (34 %)	6,024 (100 %)
>8 h	3,624 (23 %)	4,320 (29 %)	7,479 (48 %)	15,423 (100 %)
Saturday (23/4/2022)				
Dwell time	**Daily**	**Inhabitants**	**Tourists**	**Total**
1–4 h	7,219 (43 %)	2,347 (14 %)	7,205 (43 %)	16,771 (100 %)
4–6 h	4,315 (45 %)	696 (7 %)	4,772 (48 %)	9,783 (100 %)
6–8 h	5,573 (46 %)	850 (7 %)	5,654 (47 %)	12,077 (100 %)
>8 h	6,974 (31 %)	4,709 (21 %)	11,097 (49 %)	22,780 (100 %)
Sunday (24/4/2022)				
Dwell time	**Daily**	**Inhabitants**	**Tourists**	**Total**
1–4 h	11,203 (39 %)	2,347 (8 %)	15,331 (53 %)	28,881 (100 %)
4–6 h	6,921 (40 %)	758 (4 %)	9,557 (56 %)	17,236 (100 %)
6–8 h	6,792 (38 %)	787 (4 %)	10,301 (58 %)	17,880 (100 %)
>8 h	6,595 (26 %)	4,627 (18 %)	13,992 (56 %)	25,214 (100 %)

5.5. *Origin of attendance*

Italian attendance over the weekend was 141,105 spectators (24,844 on Friday, 45,332 on Saturday and 70,929 on Sunday). Analysing the origin by place, spectators from Emilia–Romagna constituted the largest group (59,747 in the three days, 42 % of the total), followed by Lombardy (16,060), Veneto (10,406), Lazio (9,556) and Tuscany (8,544). With regards to provinces, Bologna counted 37,374 spectators over the three days (of which 24,552 'inhabitants') equal to 62 % of the people coming from Emilia–Romagna region and 26 % of Italians. The following provinces were Rome (7,081), Modena (6,101), Milan (5,971) and Ravenna (5,937). The other provinces of Emilia–Romagna totalled 22,373 spectators as follows: 2,841 Reggio Emilia, 2,414 Forlì–Cesena, 1,959 Ferrara, 1,819 Parma, 894 Rimini, 408 Piacenza respectively. With details of foreigners, 45,073 presences were recorded in the three days of competitions (10,701 on Friday, 16,083 on Saturday and 18,289 on Sunday). Table 4 shows the most-represented countries with regards to agonistic day, while choropleth map in Fig. 12 highlights the total number of foreign spectators.

6. Discussion

A noteworthy relation between in- and outflows and the schedule of the weekend reported in Table 1 was identified. Furthermore, a comparison of the number of attendees revealed a discernible but predictable upward trend. Additionally, dwell time analysis revealed several key insights into the behaviour of attendees and the appeal of the event. In general, the number of attendees on Friday and Saturday was lower than on Sunday, but those who did attend were more assiduous (i.e., they spent a greater amount of time at the event), while on Sunday the audience was larger but less attentive. With regard to Sunday, the dwell

Table 4
Most-represented countries with regards to day.

Country	Friday (22/4/2022)	Saturday (23/4/2022)	Sunday (24/4/2022)	Total
United Kingdom	2,971	3,572	3,917	10,460
The Netherlands	1,392	2,054	2,199	5,645
Germany	1,167	1,565	1,685	4,417
France	581	864	1,008	2,453
Spain	508	859	912	2,279
USA	317	792	1,066	2,175

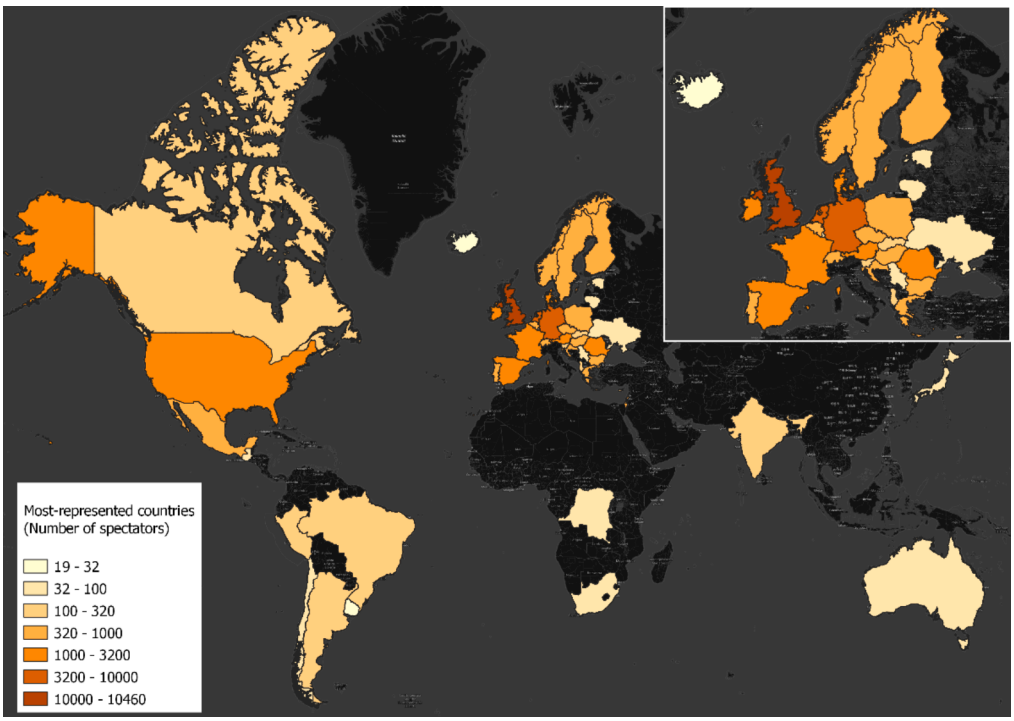

Fig. 12. Most-represented countries (zoom: European countries). Own elaboration.

time was found to be proportionately lower than on the other days. Indeed, the proportion of spectators who attended solely the F1 race was considerably higher than on other days. This finding is consistent with previous research efforts (Hall et al., 2010), as the joint analysis of arrival, departure and dwell time revealed the presence of a group of loyal fans. As they are expected to maximise the time spent at the venue location, they have specific transportation needs (Martins et al., 2022), such as early and late collective transportation services, secure and accessible parking lots, multi-day fares for public transportation.

A detailed examination of the arrival times at the circuit reveals a symmetry throughout the weekend, despite the varying schedule of the competitive sessions. The majority of spectators arrived at the circuit within the first half of the day, even on Sunday when the weekend's main event (i.e., the F1 race) was scheduled for 3p.m. This finding is consistent with the trends observed in other major sporting events (Leilei et al., 2012). The time slot that recorded the highest number of entrants was 10:00–10:30, with 13,000 accesses to the tribunes over the three days. It should also be noted that the gates were opened every day at 8:00. The analysis of spectators' profiles revealed that those who identified as 'inhabitants' exhibited a more attenuated presence at the racetrack compared to 'daily' and 'tourist' attendees. This was evidenced by the fact that the time slot with the highest number of entrances was 16:30–17:00. The data revealed that the highest number of attendees arrived at the racetrack at 15:00 on Friday (immediately before the start of the qualifying session), 15:30–16:00 on Saturday (before the 'Sprint' race, time slot without any scheduled agonistic activity) and 11:00–11:30 on Sunday (three and a half hours before the green light of the F1 race and during the F2 Feature Race). With regard to departures, the analysis revealed that the Sunday schedule did not include any competitive or entertainment events at the conclusion of the F1 race and found evidence on the disappointing weekend for the Italian-based team Ferrari and for both the drivers. Charles Leclerc finished in sixth place after making a significant error in the latter stages of the race (54th lap out of 63), while Carlos Sainz retired his vehicle during the initial lap due to a crash. The unsatisfactory result for Ferrari was confirmed by the outflows, which reached their highest values at the immediate end of the race, as reported in Fig. 10. A significant, albeit empirical, relation between Ferrari's disappointing racing results and the departures is evident even when analysing the geographical origin of attendees and their mobility patterns, as shown in Fig. 11, which highlights the frustration of supporters. With regards to the origin of attendants, the majority of foreigners came from countries with a longstanding tradition of motorsport. In fact, 10,460 (equivalent to 23 % of total foreign attendance) came from the United Kingdom. Subsequently, 5,645 individuals came from the Netherlands (12.5 %), the country of the world champion Max Verstappen, who is likely to benefit from a considerable following of Dutch fans. Germany (9.8 %) and Spain (5 %) followed with 4,417 and 2,543 spectators, respectively. Finally, the United States of America (4.8 %) accounted for 2,175 attendees. In general, however, European spectators remain in a clear majority compared to non-Europeans, even in the context of the travel restrictions following the COVID-19 pandemic.

The argument presented here can be interpreted as a robust result of the analysis and is a topic worthy of further discussion. As previously acknowledged, the necessity for updated information on people behaviours during major events is important for an effective management of flows, and hence to reduce the externalities on the transportation systems and the hosting locations (McCullough et al., 2020), especially in urban areas (Giuliano & Lu, 2021) or when the attendance is significantly higher than the local population (Not, 2021). The most evident externalities are the non-recurrent congestion along the roads (Bao et al., 2023; Giuliano & Lu, 2021; Rabadi et al., 2015), the saturation of the parking lots (Bao et al., 2023), the increased travel times for attendees (Bao et al., 2023) and the increased risk of crashes (Cigognetti et al., 2024). This is exemplified by the presented case study, given that the racetrack is situated in close proximity to the city centre of Imola, a small town (69,000 registered inhabitants). Given the spatial and temporal resolution of mobile phone data and the time span of this analysis (i.e., 30 min), the results provide a response to the issues raised by Leilei et al. (2012) and Kwoczek et al. (2014). In fact, an analysis based on cell phone data is able to provide real data when modelling or predicting the demand behaviours of individuals when approaching or departing from the event location (Bao et al., 2023; Kwoczek et al., 2014, 2015).

7. Conclusions and future works

Formula One (F1) is regarded as the pinnacle of motor racing due to the advanced engineering and the exceptional abilities of the drivers. Since the inaugural races in the 1950 s, Formula One has ascended to become one of the most-watched sports worldwide, rivalling the Olympics and the FIFA World Cup in terms of global popularity. The season calendar typically comprises approximately 20 races held in various locations around the globe. It is notable that European countries continue to play a significant role in the hosting of these events. In light of the increasing interest in hosting F1 races, the competition is intense, necessitating that prospective hosts guarantee the highest standards. This encompasses the implementation of ongoing infrastructural improvements, the involvement of a diverse range of stakeholders, and the development of an effective marketing strategy. From the perspective of crowd management in urban settings, a F1 venue could be considered an intermediate event between ongoing events throughout the season, such as football matches, and the 'once-in-a-lifetime' events such as the Olympics. Indeed, they are analogous to the former due to their repetitiveness, although they occur only once per consecutive year. Moreover, monitoring the movements of individuals is crucial for understanding their behaviour in reaching the venue, their spatial distribution, and, from a more general perspective, the economic impact of the event itself. A fundamental element of assessing the influence of significant occurrences in urban environments is the accurate estimation of the actual number of attendees and individuals directly involved. It should be noted that this figure may differ from the official number provided by ticket sales, as it does not account for workers (e.g. crews, journalists, etc.) or complimentary and promotional tickets.

This study offers a comprehensive examination of the patterns and mobility behaviours exhibited during the 2022 F1 Emilia-Romagna and Made in Italy Grand Prix, held in Imola, Italy. The data was obtained from a set of mobile phone records belonging to Vodafone Italia, one of three principal mobile phone operators in the Italian market. The data processing, which was based on inferential statistics and in compliance with privacy laws, revealed the actual number of attendees, their characteristics (i.e., whether they were residents of Imola, tourists or visitors), their movements (when they entered the circuit's gates; when they left; how long they remained) and their origin. In general terms, this research can be categorised as descriptive analytics, although it is not as far removed from diagnostic analytics (H. Huang et al., 2021) as one might expect. Moreover, thanks to the repetitiveness of this event (i.e., contracts between local Authorities and F1 management are usually signed for more than one year) and the *a–priori* knowledge of the schedule of planned activities (e.g., agonistic sessions), it is possible to predict the mobility choices and hence plan adequate mobility–related policies in order to minimize the impacts of the following events.

The results were consistent with the weekend schedule, the expected behaviour of attendees, and the tickets sold. Additionally, consistency with the existing literature on attendees' trends during major sport events was found. This work is intended to be a relevant case study in demonstrating how big data, and in particular mobile phone traces, can be a decisive source of information in implementing effective in- and outflow policies, even for other events. While the F1 race can be considered the most significant event in an annual programme, the Imola Autodromo hosts other major events, such as other motorsport races and concerts. Widening the gaze, the described approach and methodology could be useful in monitoring flows during similar racing venues, such as car or motorcycle races held in both street circuits and urban contexts. Some examples of the previous typology include the F1 venues of Monte Carlo, Monaco, Singapore, and Albert Park in Melbourne, Australia; some of the IndyCar Series tracks; the Formula E circuits. Indeed, permanent racetracks in dense urban environments, such as Monza, Italy; Interlagos, Brazil; and Zandvoort, The Netherlands, can benefit from the application of this methodological approach.

In the lights of the findings, the authors considered a number of additional topics. Given the pervasive presence of connected devices, it would be appropriate to conduct analyses on the modal share in getting to Imola and the circuit. This can be aimed in order to gain a detailed overview of attendance flows and plan outflow strategies to avoid congestion issues. In particular, the use of collective means of transport, such as public transportation lines or shuttle buses, which is a common practice during a major event (Ceder & Perera, 2014; Chirieleison & Scrucca, 2017), can be further assessed and managed. The majority of the future developments can be achieved by integrating cell phone data with additional sources of data, such as traffic loops or real-time feeds from public transportation agencies (e.g., Automatic Passenger Counter, Automatic Vehicle Location), and by overcoming the technical limitations outlined in Sections 4.2 and 5. This can be achieved by modifying the aggregation thresholds or improving the telecommunications network in the vicinity of the circuit, even temporarily, for instance by installing additional antennas. This latter consideration is a key component in order to achieve a better spatial–temporal resolution and enhance the potential of telecommunications data in monitoring large crowds during major events. Moreover, in light of the taxonomy proposed by H. Huang et al. (2021), enhancements can be made to achieve 'predictive' or 'prescriptive analytics', thereby enabling the triggering of new analyses and the enhancement of insights and value.

CRediT authorship contribution statement

Alessandro Nalin: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Andrea Simone:** Visualization. **Claudio Lantieri:** Writing – review & editing. **Denis Cappellari:** Writing – review & editing, Conceptualization. **Glauco Mantegari:** Writing – original draft, Methodology, Conceptualization. **Valeria Vignali:** Writing – review & editing, Supervision, Investigation, Conceptualization.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to express their gratitude to Andrea Zaramella and Dario di Sorte of Vodafone Italia for data provision and support during analysis and to the President Gian Carlo Minardi, the Director Pietro Benvenuti, and the technical staff of the 'Autodromo Enzo e Dino Ferrari' International Circuit for the help.

References

Afandizadeh Zargari, S., Memarnejad, A., Mirzahossein, H., 2021. A structural comparison between the origin-destination matrices based on local windows with socioeconomic, land-use, and population characteristics. J. Adv. Transp. 2021, 1–17. https://doi.org/10.1155/2021/9968698.

Altin, L., Ahas, R., Silm, S., Saluveer, E., 2022. Megastar concerts in tourism: a study using mobile phone data. Scand. J. Hosp. Tour. 22 (2), 161–180.

Aversa, P., Furnari, S., Haefliger, S., 2015. Business model configurations and performance: a qualitative comparative analysis in Formula One racing, 2005–2013. Ind. Corp. Chang. 24 (3), 655–676.

Balduini, M., Brambilla, M., Valle, E.D., Marazzi, C., Arabghalizi, T., Rahdari, B., Vescovi, M., 2019. Models and practices in urban data science at scale. Big Data Res. 17, 66–84.

Bao, Y., Yang, H., Gao, Z., Xu, H., 2023. How do pre-event activities alleviate congestion and increase attendees' travel utility and the venue's profit during a special event? Transp. Res. B Methodol. 173, 332–353. https://doi.org/10.1016/j.trb.2023.05.013.

Battistini, R., Nalin, A., Simone, A., Lantieri, C., Vignali, V., 2022. How do university student cyclists ride? the case of university of Bologna. Appl. Sci. 12 (22), 11569. https://doi.org/10.3390/app122211569.

Battistini, R., Lantieri, C., Simone, A., Dondi, G., Vignali, V., 2023. A decision support system for the safety evaluation of urban pedestrian crossings. Transp. Res. Procedia 69, 655–662.

Berlingerio, M., Calabrese, F., Di Lorenzo, G., Nair, R., Pinelli, F., Sbodio, M.L., 2013. AllAboard: A System for Exploring Urban Mobility and Optimizing Public Transport Using Cellphone Data. In: Salinesi, C., Norrie, M.C., Pastor, Ó. (Eds.), Advanced Information Systems Engineering, Vol. 7908. Springer, Berlin Heidelberg, pp. 663–666. https://doi.org/10.1007/978-3-642-40994-3_50.

Bonnetain, L., Furno, A., Krug, J., & Faouzi, N.-E. E. (2019). Can We Map-Match Individual Cellular Network Signaling Trajectories in Urban Environments? Data-Driven Study. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(7), 74–88. https://doi.org/10.1177/0361198119847472.

Bwambale, A., Choudhury, C., Hess, S., 2019. Modelling long-distance route choice using mobile phone call detail record data: a case study of Senegal. Transportmetrica a: Trans. Sci. 15 (2), 1543–1568. https://doi.org/10.1080/23249935.2019.1611970.

Caceres, N., Romero, L. M., & Benitez, F. G. (2020). Exploring strengths and weaknesses of mobility inference from mobile phone data vs. Travel surveys. *Transportmetrica (Abingdon, Oxfordshire, UK)*, *16*(3), 574–601.

Caceres, N., Wideberg, J.P., Benitez, F.G., 2007. Deriving origin–destination data from a mobile phone network. IEE Proceedings. Intel. Trans. Sys. 1 (1), 15.

Calabrese, F., Pereira, F.C., Lorenzo, G.D., Liu, L., Ratti, C., 2010. The Geography of Taste: Analyzing Cell-Phone Mobility and Social Events. In: In *Lecture Notes in Computer Science*, p. (p. 22).. https://doi.org/10.1007/978-3-642-12654-3_2.

Calabrese, F., Ferrari, L., Blondel, V., 2015. Urban sensing using mobile phone network data: a survey of research. ACM Comput. Surv. 47 (2), 1–20.

Carneiro Pereira, L., & Prevedouros, P. D. (2022). *Analysis of Naturalistic Driving Events in Honolulu (Pre-COVID)*. 138–149.

Ceder (Avi), A., Perera, S., 2014. Detecting and improving public-transit connectivity with case studies of two world sport events. Transp. Policy 33, 96–109. https://doi.org/10.1016/j.tranpol.2014.03.001.

Chamberlain, D.A., Edwards, D., Lai, J., Thwala, W.D., 2019. Mega event management of formula one grand prix: an analysis of literature. Facilities 37 (13/14), 1166–1184.

Chen, H., Song, X., Xu, C., & Zhang, X. (2021). Using Mobile Phone Data to Examine Point-of-Interest Urban Mobility. In *Journal of Urban Technology* (Vol. 27, Issue 4, p. 43). https://doi.org/10.1080/10630732.2021.1882175.

Chen, M., Mao, S., Liu, Y., 2014. Big data: a survey. Mobile Networks and Applications 19 (2), 171–209. https://doi.org/10.1007/s11036-013-0489-0.

Chin, K., Huang, H., Horn, C., Kasanicky, I., & Weibel, R. (2019). Inferring fine-grained transport modes from mobile phone cellular signaling data. In *Computers, Environment and Urban Systems* (Vol. 77). https://doi.org/10.1016/j.compenvurbsys.2019.101348.

Chirieleison, C., Scrucca, L., 2017. Event sustainability and transportation policy: a model-based cluster analysis for a cross-comparison of hallmark events. Tour. Manag. Perspect. 24, 72–85. https://doi.org/10.1016/j.tmp.2017.07.020.

Cigognetti, T., Carra, M., Ghirardi, A., Assefa, N.G., Ferretto, L., Ventura, R., Maternini, G., Barabino, B., 2024. Mega-events: assessing road safety through an operating framework. an application for the milano-cortina 2026 winter olympic games. Infrastructures 9 (3), 51. https://doi.org/10.3390/infrastructures9030051.

Cik, M., Cornelia, H., Horn, C., Schulze, G., Traub, M., Schweighofer, E., Hötzendorfer, W., & Fellendorf, M. (2017). *Using cell phone and social media data to enhance safety at mega events*.

Cik, M., Fellendorf, M., 2019. Cell Phone based Origin-Destination Matrices for Transport Modelling. Transp. Res. Procedia 41, 551–553. https://doi.org/10.1016/j.trpro.2019.09.097.

Comune di Imola. (2022a). *Mappa dei parcheggi e programma degli eventi collaterali in occasione del Gran Premio di F1 2022*. https://visitareimola.it/eng/wp-content/uploads/2022/04/MappaInfo-A3-DEFINITIVA-AGGIORNATA.pdf.

Comune di Imola. (2022b). *Ordinanza n. 741 del 15/04/2022. Modifiche alla viabilità stradale di Imola in occasione del Gran Premio di Formula Uno 2022*. https://www.comune.imola.bo.it/novita/notizie/2022/04/le-modifiche-alla-viabilita-in-occasione-del-gran-premio-di-f1.

Demissie, M.G., Phithakkitnukoon, S., Sukhvibul, T., Antunes, F., Gomes, R., Bento, C., 2016. Inferring passenger travel demand to improve urban mobility in developing countries using cell phone data: a case study of senegal. IEEE Trans. Intell. Transp. Syst. 17 (9), 2466–2478. https://doi.org/10.1109/TITS.2016.2521830.

Doyle, J., Hung, P., Farrell, R., & Mcloone, S. (2014). Population Mobility Dynamics Estimated from Mobile Telephony Data. In *Journal of Urban Technology* (Vol. 21, Issue 2, p. 109). https://doi.org/10.1080/10630732.2014.888904.

Fairley, S., Tyler, B.D., Kellett, P., D'Elia, K., 2011. The formula one Australian grand prix: exploring the triple bottom line. Sport Management Review 14 (2), 141–152.

Fekih, M., Bellemans, T., Smoreda, Z., Bonnel, P., Furno, A., Galland, S., 2021. A data-driven approach for origin–destination matrix construction from cellular network signalling data: a case study of Lyon region (France). Transportation 48 (4), 1671–1702. https://doi.org/10.1007/s11116-020-10108-w.

Fiadino, P., Valerio, D., Ricciato, F., Hummel, K.A., 2012. Steps towards the Extraction of Vehicular Mobility Patterns from 3G Signaling Data. In: Pescapè, A., Salgarelli, L., Dimitropoulos, X. (Eds.), Traffic Monitoring and Analysis, Vol. 7189. Springer, Berlin Heidelberg, pp. 66–80. https://doi.org/10.1007/978-3-642-28534-9_7.

Gandomi, A., Haider, M., 2015. Beyond the hype: big data concepts, methods, and analytics. Int. J. Inf. Manag. 35 (2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007.

Gerard-Reimer, C.C., 2020. Race Cartels: how Constructor Collaboration Is Curbing Innovation in Formula 1. *Vand. J. Ent. & Tech.*l. 23, 855.

Gezici, F., Er, S., 2014. What has been left after hosting the Formula 1 Grand Prix in Istanbul? Cities 41 (a), 44–53.

Ghasemi, N., Acerra, E., Vignali, V., Lantieri, C., Simone, A., Imine, H., 2020. Road Safety Review update by using innovative technologies to investigate driver behaviour. Transp. Res. Procedia 45, 368–375.

Giuliano, G., & Lu, Y. (2021). Analyzing Traffic Impacts of Planned Major Events. *Transportation Research Record: Journal of the Transportation Research Board*, *2675*(8), 432–442. https://doi.org/10.1177/0361198121998710.

González, M. C., Xu, Y., & González, M. C. (2017). *Collective benefits in traffic during mega events via the use of information technologies*. https://doi.org/10.6084/m9.

Hadachi, A., Pourmoradnasseri, M., Khoshkhah, K., 2020. Unveiling large-scale commuting patterns based on mobile phone cellular network data. In. J. Transp. Geogr. 89 https://doi.org/10.1016/j.jtrangeo.2020.102871.

Hall, J., O'Mahony, B., Vieceli, J., 2010. An empirical model of attendance factors at major sporting events. Int. J. Hosp. Manag. 29 (2), 328–334. https://doi.org/10.1016/j.ijhm.2009.10.011.

Hasselwander, M., Tamagusko, T., Bigotte, J.F., Ferreira, A., Mejia, A., Ferranti, E.J.S., 2021. Building back better: the COVID-19 pandemic and transport policy implications for a developing megacity. Sustain. Cities Soc. 69, 102864 https://doi.org/10.1016/j.scs.2021.102864.

Huang, H., Cheng, Y., Weibel, R., 2019. Transport mode detection based on mobile phone network data: a systematic review. In. Trans. Res. Part c: Emerg. Technol. 101, p. 297). https://doi.org/10.1016/j.trc.2019.02.008.

Huang, H., Yao, X. A., Krisp, J. M., & Jiang, B. (2021). *Analytics of location-based big data for smart cities: Opportunities, challenges, and future directions*. *90*.

Huang, Z., Ling, X., Wang, P., Zhang, F., Mao, Y., Lin, T., Wang, F.-Y., 2018. Modeling real-time human mobility based on mobile phone and transportation data fusion. Trans. Res. Part c, Emerg. Technol. 96, 251–269.

Kim, M.K., Kim, S.-K., Park, J.-A., Carroll, M., Yu, J.-G., Na, K., 2017. Measuring the economic impacts of major sports events: the case of Formula One Grand Prix (F1). Asia Pacific J. Tourism Res. 22 (1), 64–73.

Kontaxi, A., Ziakopoulos, A., Yannis, G., 2021. Trip characteristics impact on the frequency of harsh events recorded via smartphone sensors. IATSS Research 45 (4), 574–583. https://doi.org/10.1016/j.iatssr.2021.07.004.

Kwoczek, S., Di Martino, S., Nejdl, W., 2014. Predicting and visualizing traffic congestion in the presence of planned special events. J. Vis. Lang. Comput. 25 (6), 973–980. https://doi.org/10.1016/j.jvlc.2014.10.028.

Kwoczek, S., Di Martino, S., & Nejdl, W. (2015). Stuck Around the Stadium? An Approach to Identify Road Segments Affected by Planned Special Events. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 1255–1260. https://doi.org/10.1109/ITSC.2015.206.

Lefebvre, S., Roult, R., 2011. Formula One's new urban economies. Cities 28 (4), 330–339. https://doi.org/10.1016/j.cities.2011.03.005.

Lefebvre, S., Roult, R., 2013. Territorial and touristic branding: urban history and the festive and economic perspectives of Montreal's Formula One Grand Prix. Loisir et Société / Society and Leisure 36 (1), 43–59.

Leilei, D., Zheng-liang, S., Jin-gang, G., & Hong-tong, Q. (2012). Study on traffic organization and management strategies for large special events. *2012 International Conference on System Science and Engineering (ICSSE)*, 432–436. https://doi.org/10.1109/ICSSE.2012.6257222.

Liguori, G., Amadori, M., Iacorossi, F., Marcucci, E., Simone, A., Lantieri, C., 2017. From people to people-the self-benefit of crowdsourced cycling data as part of the european cycling challenge. Procedia Int. Cycling Confer. 2017, 1–21.

Liu, D., Gratton, C., 2010. The impact of mega sporting events on live spectators' images of a host city: a case study of the Shanghai F1 Grand Prix. Tour. Econ. 16 (3), 629–645. https://doi.org/10.5367/000000010792278347.

Lowes, M., 2018. Toward a conceptual understanding of formula one motorsport and local cosmopolitanism discourse in urban placemarketing strategies. Commun. Sport 6 (2), 203–218. https://doi.org/10.1177/2167479517697955.

Ma, L., Chen, Q., Han, K., Gao, Y., & Li, D. (2018). *A tale of two stations: Analyzing metro ridership with big data*.

Mamei, M., Colonna, M., 2016. Estimating attendance from cellular network data. Int. J. Geogr. Inf. Sci. 30 (7), 1281–1301.

Marques-Neto, H.T., Xavier, F.H., Xavier, W.Z., Malab, C.H.S., Ziviani, A., Silveira, L.M., Almeida, J.M., 2018. Understanding human mobility and workload dynamics due to different large-scale events using mobile phone data. J. Netw. Syst. Manag. 26 (4), 1079–1100.

Martins, R., Pereira, E., Rosado, A., Marôco, J., McCullough, B., Mascarenhas, M., 2022. Understanding spectator sustainable transportation intentions in international sport tourism events. J. Sustain. Tour. 30 (8), 1972–1991. https://doi.org/10.1080/09669582.2021.1991936.

Mastromarco, C., Runkel, M., 2009. Rule changes and competitive balance in Formula One motor racing. Appl. Econ. 41 (23), 3003–3014.

McCartney, G.J., 2005. Hosting a recurring mega-event: visitor raison d'être. J. Sport Tourism 10 (2), 113–128.

McCullough, B.P., Orr, M., Watanabe, N.M., 2020. Measuring Externalities: the imperative next step to sustainability assessment in sport. J. Sport Manag. 34 (5), 393–402. https://doi.org/10.1123/jsm.2019-0254.

Mohammadi, S., Kamrani, M., Khattak, A. J., & Chakraborty, S. (2019). *Social Influence on Driver Decisions Using Modeling and Gossip Algorithms*.

Mourão, P., 2017. The economics of motorsports: The case of Formula One. Springer.

Mourao, P.R., 2018. Smoking gentlemen—how formula one has controlled CO2 emissions. Sustainability 10 (6), 1841.

Müller, M. (2015). What makes an event a mega-event? Definitions and sizes. In *Leisure Studies* (Vol. 34, Issue 6, p. 627). https://doi.org/10.1080/02614367.2014.993333.

Nalin, A., Vignali, V., Lantieri, C., Cappellari, D., Zamengo, B., Simone, A., 2024. Assessing veracity of big data: an in-depth evaluation process from the comparison of Mobile phone traces and groundtruth data in traffic monitoring. J. Transp. Geogr. 118, 103930 https://doi.org/10.1016/j.jtrangeo.2024.103930.

Neilson, A., Indratmo, Daniel, B., & Tjandra, S. (2019). Systematic Review of the Literature on Big Data in the Transportation Domain: Concepts and Applications. *Big Data Research, 17*, 35–44. https://doi.org/10.1016/j.bdr.2019.03.001.

Not, E., 2021. Mining mobile application usage data to understand travel planning for attending a large event. Inform. Technol. Tourism 23 (3), 291–325. https://doi.org/10.1007/s40558-021-00204-7.

Parsafard, M., Chi, G., Qu, X., Li, X., Wang, H., 2019. Error measures for trajectory estimations with geo-tagged mobility sample data. IEEE Trans. Intell. Transp. Syst. 20 (7), 2566–2583. https://doi.org/10.1109/TITS.2018.2868182.

Pereira, R.H.M., 2018. Transport legacy of mega-events and the redistribution of accessibility to urban destinations. Cities 81, 45–60. https://doi.org/10.1016/j.cities.2018.03.013.

Pintér, G., & Felde, I. (2021). Analyzing the Behavior of Soccer Fans from a Mobile Phone Network Perspective: Euro 2016, a Case Study. *arXiv Preprint arXiv:2108.09291*.

Ponieman, N.B., Sarraute, C., Minnoni, M., Travizano, M., Zivic, P.R., Salles, A., 2015. Mobility and sociocultural events in mobile phone data records. AI Commun. 29 (1), 77–86.

Pozdnukov, A., & Waddell, P. (2016). *Demand Forecasting and Activity-based Mobility Modelling from Cell Phone Data*.

Preuss, H., 2006. Lasting effects of major sporting events. Institute of Sport Science 1–11.

Qian, C., Li, W., Duan, Z., Yang, D., Ran, B., 2021. Using mobile phone data to determine spatial correlations between tourism facilities. J. Transp. Geogr. 92.

Rabadi, G., Khallouli, W., Salem, M.A., Ghoniem, A., 2015. Planning and management of major sporting events: a survey. Int. J. Plan. Schedul. 2 (2), 154. https://doi.org/10.1504/IJPS.2015.072122.

Ramasamy, B., Wu, H., Yeung, M., 2020. Hosting annual international sporting events and tourism: Formula 1, golf or tennis? Tour. Econ. 13548166211029053.

Ricciato, F., Widhalm, P., Pantisano, F., Craglia, M., 2017. Beyond the "single-operator, CDR-only" paradigm: an interoperable framework for mobile phone network data analyses and population density estimation. Pervasive Mob. Comput. 35, 65–82. https://doi.org/10.1016/j.pmcj.2016.04.009.

Roult, R., Auger, D., Lafond, M.-P., 2020. Formula 1, city and tourism: a research theme analyzed on the basis of a systematic literature review. Int. J. Tourism Cities 6 (4), 813–830. https://doi.org/10.1108/IJTC-02-2020-0025.

Sadeghvaziri, E., Rojas, M. B., & Jin, X. (2016). Exploring the Potential of Mobile Phone Data in Travel Pattern Analysis. *Transportation Research Record: Journal of the Transportation Research Board, 2594*(1), 27–34. https://doi.org/10.3141/2594-04.

Scannapieco, M., Virgillito, A., & Zardetto, D. (2013). *Placing big data in official statistics: A big challenge*. Paper for the New Techniques and Technologies for Statistics conference.

Schneiders, C., Rocha, C., 2022. Technology innovations and consumption of Formula 1 as a TV sport product. Sport Mark. Q. 31 (3).

Sevtsuk, A., Ratti, C., 2010. Does urban mobility have a daily routine? learning from the aggregate data of mobile networks. The Journal of Urban Technology 17 (1), 41–60.

Shoman, M.M., Imine, H., Acerra, E.M., Lantieri, C., 2023. Evaluation of cycling safety and comfort in bad weather and surface conditions using an instrumented bicycle. IEEE Access 11, 15096–15108.

Smith, A., 2012. Events and urban regeneration: The strategic use of events to revitalise cities. Routledge.

Steenbruggen, J., Tranos, E., Nijkamp, P., 2015. Data from mobile phone operators: a tool for smarter cities? Telecommun. Policy 39 (3–4), 335–346.

Storm, R. K., Nielsen, C. G., & Jakobsen, T. G. (2019). *The Impact of Formula One on Regional Economies in Europe*.

Thevenin, T., Vuidel, G., 2021. Exploring daily mobility in space and time: the geographer project. J. Transp. Geogr. 93, 103082 https://doi.org/10.1016/j.jtrangeo.2021.103082.

Wang, F., & Chen, C. (2018b). *Passively-Generated Mobile Phone Data: Data Issues and Processing Required*.

Wang, F., Chen, C., 2018a. On data processing required to derive mobility patterns from passively-generated mobile phone data. Trans. Res. Part c: Emerg. Technol. 87, 58–74. https://doi.org/10.1016/j.trc.2017.12.003.

Wang, F., Chen, W., Zhao, Y., Gu, T., Gao, S., Bao, H., 2017. Adaptively exploring population mobility patterns in flow visualization. IEEE Trans. Intell. Transp. Syst. 18 (8), 2250–2259. https://doi.org/10.1109/TITS.2017.2711644.

Warren, J., Lipkowitz, J., Sokolov, V., 2019. Clusters of driving behavior from observational smartphone data. IEEE Intell. Transp. Syst. Mag. 11 (3), 171–180. https://doi.org/10.1109/MITS.2019.2919516.

Widhalm, P., Yang, Y., Ulm, M., Athavale, S., González, M.C., 2015. Discovering urban activity patterns in cell phone data. Transportation 42 (4), 597–623. https://doi.org/10.1007/s11116-015-9598-x.

Wischer, T., Cik, M., & Fellendorf, M. (2023). Graph Supported Mode Detection within Mobile Phone Data Trajectories. *Transportation Research Record: Journal of the Transportation Research Board*, *2677*(3), 18–32. https://doi.org/10.1177/03611981221150399.

Xavier, F., Silveira, L., Almeida, J., Ziviani, A., Malab, C., & Marques-Neto, H. (2012). *Analyzing the workload dynamics of a mobile phone network in large scale events*. 37–42. https://doi.org/10.1145/2413236.2413245.

Xiao, J., & Lou, Y. (2018). *A smartphone-based parking guidance system with predictive parking availability information*. 2018 Transportation research annual meeting.

Yabe, T., Jones, N.K.W., Rao, P.S.C., Gonzalez, M.C., Ukkusuri, S.V., 2022. Mobile phone location data for disasters: A review from natural hazards and epidemics. Computers, Environment and Urban Systems, p. 94.

Yang, C., Zhang, Y., Zhan, X., Ukkusuri, S.V., Chen, Y., 2020. Fusing Mobile phone and travel survey data to model urban activity dynamics. J. Adv. Transp. 2020, 1–17. https://doi.org/10.1155/2020/5321385.

Zahedi, S., Shafahi, Y., 2018. Estimating activity patterns using spatio-temporal data of cell phone networks. Int. J. Urban Sci. 22 (2), 162–179. https://doi.org/10.1080/12265934.2017.1331139.

Zhang, Y., Zhong, W., Wang, D., & Lin, F.-T. (2021). Understanding the spatiotemporal patterns of nighttime urban vibrancy in central Shanghai inferred from mobile phone data. In *Regional Sustainability* (Vol. 2, Issue 4, p. 297). https://doi.org/10.1016/j.regsus.2021.11.006.