



HAL
open science

Reviewing Trip Purpose Imputation in GPS-based Travel Surveys

Minh Hieu Nguyen, Jimmy Armoogum, Jean-Loup Madre, Cédric Garcia

► **To cite this version:**

Minh Hieu Nguyen, Jimmy Armoogum, Jean-Loup Madre, Cédric Garcia. Reviewing Trip Purpose Imputation in GPS-based Travel Surveys. *Journal of Traffic and Transportation Engineering*, 2020, 7 (4), pp.395-412. 10.1016/j.jtte.2020.05.004 . hal-02916874

HAL Id: hal-02916874

<https://hal.science/hal-02916874>

Submitted on 1 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Review paper

2

3 **Reviewing trip purpose imputation in**

4 **GPS-based travel surveys**

5

6 **Minh Hieu Nguyen^{a, b, *}, Jimmy Armoogum^a, Jean-Loup Madre^a, Cédric Garcia^a**

7

8 ^a *Université Gustave Eiffel / AME / DEST, F-77447 Marne la Vallée Cedex 2, France*

9 ^b *University of Transport and Communications / Faculty of Transport Economics, Hanoi, Vietnam*

10

11 **Highlights**

- 12 • Defining purpose imputation from GPS data in the context of the revolution of GPS-based travel
- 13 surveys and two research domains (Transportation Science, Human Geography).
- 14 • Reviewing 25 papers found through a systematic survey and careful selection on Scopus, Web of
- 15 Science, ScienceDirect and TRID.
- 16 • Surveying three steps of purpose detection, which are (1) trip end detection, (2) input feature
- 17 selection and (3) main algorithms and validation, to clarify pros and cons of existing methods
- 18 coupled with proposing future research directions.

19

20 **Abstract**

21 The Global Positioning System (GPS) has motivated rapid advances in mobility data collection. A

22 massive amount of spatio-temporal information has made it possible to know where a person was and

23 when, but not how and why (s)he travelled, creating the need for inference models. Compared with

24 mode detection, purpose imputation has been insufficiently studied. However, the relative lack of

25 attention to purpose identification does not mean that this field has not emerged. For this paper, which is

26 the first review dedicated to inferring trip purposes from GPS data, 1,162 non-duplicate papers from four
27 databases (Scopus, Web of Science, ScienceDirect and TRID) were screened, and a corpus of 25
28 publications was selected for examination. Based on these papers, the purpose imputation problem is
29 defined in the contexts of the evolution of GPS-based travel surveys and two research domains,
30 Transportation Science (TS) and Human Geography (HG). Subsequently, three steps of the purpose
31 detection process, namely (1) trip end detection, (2) input feature selection and (3) main algorithms and
32 validation, are surveyed. During these procedures, the differences between studies in TS and those in
33 HG are highlighted. Finally, unresolved issues related to data and feature selection, algorithms and
34 assessment are discussed substantially to provide potential research directions. This review may be an
35 informative reference for those newly accessing the GPS-based purpose imputation field and/or
36 intending to develop solutions to this problem.

37

38 **Keywords:**

39 GPS; Travel Survey; Purpose Imputation; Transportation Science; Human Geography; Review; Activity
40 Inference

41

*Corresponding author.

E-mail addresses: hieunm@utc.edu.vn (Minh Hieu Nguyen),
jimmy.armoogum@univ-eiffel.fr (Jimmy Armoogum), jean-loup.madre@univ-eiffel.fr
(Jean-Loup Madre), cedric.garcia@univ-eiffel.fr (Cédric Garcia).
ORCID: <https://orcid.org/0000-0003-4863-5446> (Minh Hieu Nguyen)

42 1 Introduction

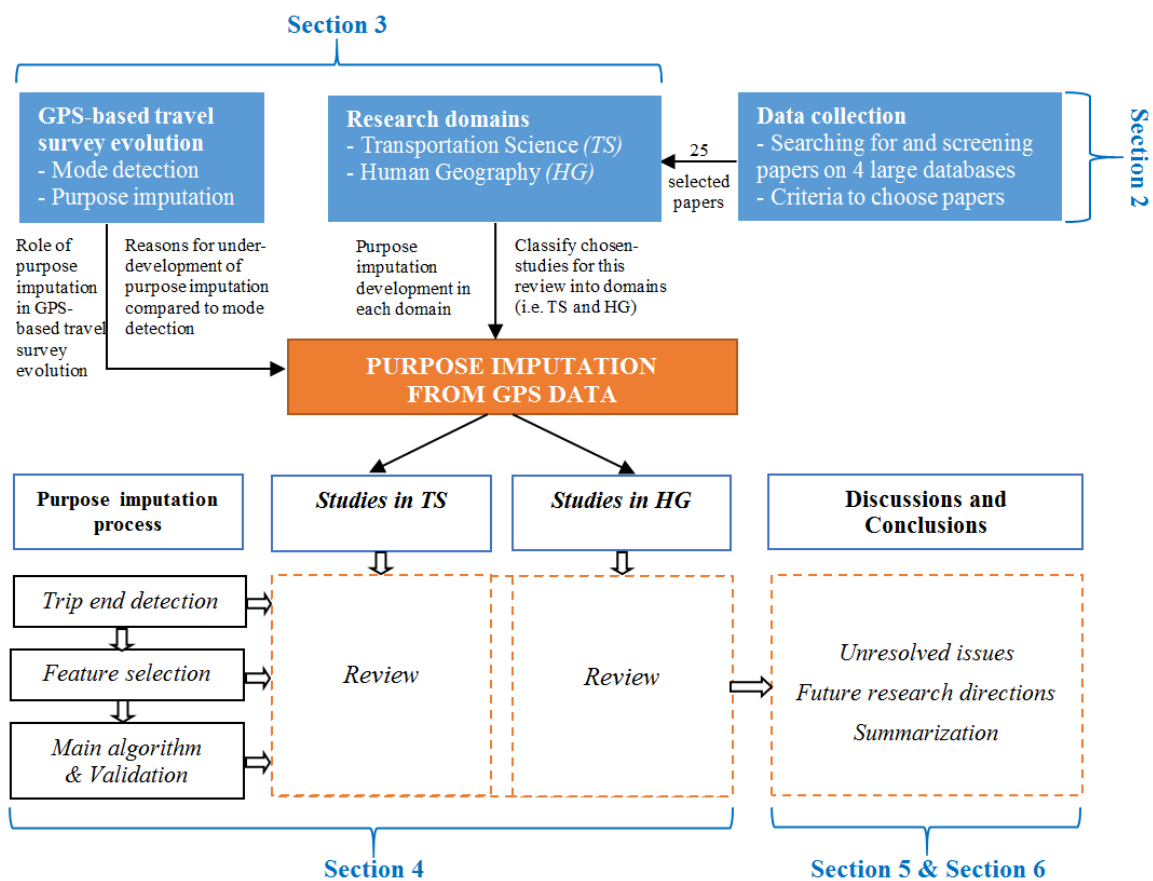
43 The Global Positioning System (GPS) has induced a technological revolution in mobility surveys.
44 Initially, surveying travel behaviors depended on face-to-face interviews with pen and paper, followed by
45 telephone and computer-aided methods (Armoogum et al., 2014). However, self-administered
46 investigations are subject to either under-reporting or over-reporting of trip rate and time (Bricka et al.,
47 2012; Stopher and Shen, 2011; Wolf et al., 2003), not to mention increasing cost, high non-response
48 rate and long gaps between periodic household travel surveys (Armoogum et al., 2009; Ortúzar et al.,
49 2011; Patrick et al., 2009), underlining a need for improving survey quality with the support of new
50 technologies. GPS has been a satisfactory answer to this demand thanks to its major benefits over
51 traditional techniques, including a mitigation of burden on participants, an extension of survey periods,
52 and the provision of more accurate data with spatial and temporal details (Bachu et al., 2001; Bohte and
53 Maat, 2009; Zhao et al., 2015). In the infancy of travel surveys, on-board devices and then wearable
54 loggers were used. More recently, the ubiquity of smartphones with inbuilt GPS sensors and the
55 proliferation of dedicated applications for collecting and validating mobility data have put handheld
56 devices in the past (Stopher et al., 2018) and created opportunities for a wider reach of GPS-assisted
57 surveys (Marra et al., 2019; Semanjski et al., 2017; Thomas et al., 2018; Xiao et al., 2016; Zhao et al.,
58 2015).

59 GPS data, on the one hand, can draw a high-resolution picture of where and when a participant
60 stayed or passed, but on the other hand, such data require sophisticated approaches to extract trip
61 characteristics, which are not included directly (Du and Aultman-Hall, 2007). The two most important trip
62 attributes are transportation mode and purpose. In comparison with the former, the latter has been
63 insufficiently studied (Ermagun et al., 2017). No existing review has been dedicated solely to purpose
64 imputation from GPS data, whose content makes up a small part of discussions about broad topics such
65 as the comparison between small and big data (Chen et al., 2016), mining GPS data (Lin and Hsu,
66 2014), and evaluation of a series of new positioning technologies (Lee et al., 2016).

67 Although focusing on mode detection solutions, Gong et al. (2014) and Shen and Stopher (2014)
68 devote more effort to synthesizing how to derive purposes from GPS data than do other authors. The

69 former study has shed light on reviewing purpose imputation in the context of GPS data processing,
 70 carefully presenting both positive and negative sides of the methods examined. The latter authors go
 71 further by classifying algorithms into three categories according to their mathematical complexity. Only
 72 concentrating on main algorithms to infer purposes would be a limitation as input feature selection and
 73 method validation are focal respects in a classification problem.

74 To fill the mentioned-above gaps, conducting a critical synthesis to suggest potential future research
 75 directions is necessary. The remainder of this paper is structured into five parts (see Figure 1). Section
 76 2 describes the procedures of searching for and selecting papers for this review. In Section 3, the GPS
 77 data-based purpose imputation problem is defined in the context of the travel survey revolution and two
 78 research domains (Transportation Science and Human Geography). Section 4 reviews three primary
 79 steps to infer purpose, that is, trip end detection, feature selection, main algorithm and validation. A
 80 substantial discussion about unsolved issues is presented in Section 5 before Section 6 reveals
 81 conclusions.



82
 83
 Fig. 1 Description of review process and the structure of the rest of this paper

84 2 Data collection

85 In order to make this review systematic, a search for papers published by mid-January 2020 on four
86 databases (Web of Science, Scopus, TRID and Science Direct) was conducted by a function using
87 Boolean operations, that is, (“*activity inference*” OR “*activity type*” OR “*activity types*” OR “*trip purpose*”
88 OR “*trip purposes*”) AND (“*GPS*” OR “*social media*”). To the best of our knowledge, *purpose(s)* and
89 *activity/activities* are interchangeable terminologies in the literature because a purpose refers to an
90 activity performed at a trip destination. To limit the number of candidates found but still ensure sufficient
91 pinpointing of publications, *trip*, *type(s)* and *inference* were added. Because some authors only indicate
92 *social media* but not *GPS* in titles, abstract and keywords (e.g., Cui et al., 2018), *social media* were
93 coupled with *GPS*. Only journal articles, conference papers and book chapters/lecture notes published
94 officially and in English were examined.

95 The numbers of papers found was: 1,098 (Scopus), 127 (TRID), 103 (Web of Science) and 38
96 (ScienceDirect). After removing duplicates, 1,162 were kept. Then, a screening of their titles, abstracts
97 and keywords was undertaken. For papers including *social media* but not *GPS* in titles, abstracts and
98 keywords, research data descriptions in full texts were scrutinized. Finally, a corpus of 25 publications
99 (see Table 1) was selected based on the four criteria that follow: (1) available full text, (2) showing
100 clearly method with variables/features used, (3) detecting at least three purposes from GPS data, (4)
101 achieving ≥ 1 citation, ≥ 3 citations and ≥ 6 citations for those published in 2019¹, 2016–2018 and
102 2001–2015, respectively. The fourth criterion enables the removal of publications with low impacts on
103 the literature. The number of citations was based on statistics from Google Scholar. As can be seen in
104 Figure 2, starting from 2001, the field of purpose imputation from GPS data introduced eight selected
105 papers for the first 10 years. Since 2012, there has been at least one paper annually.

¹ Papers are usually published online in *article in press*; therefore, they may have citations before published formally.

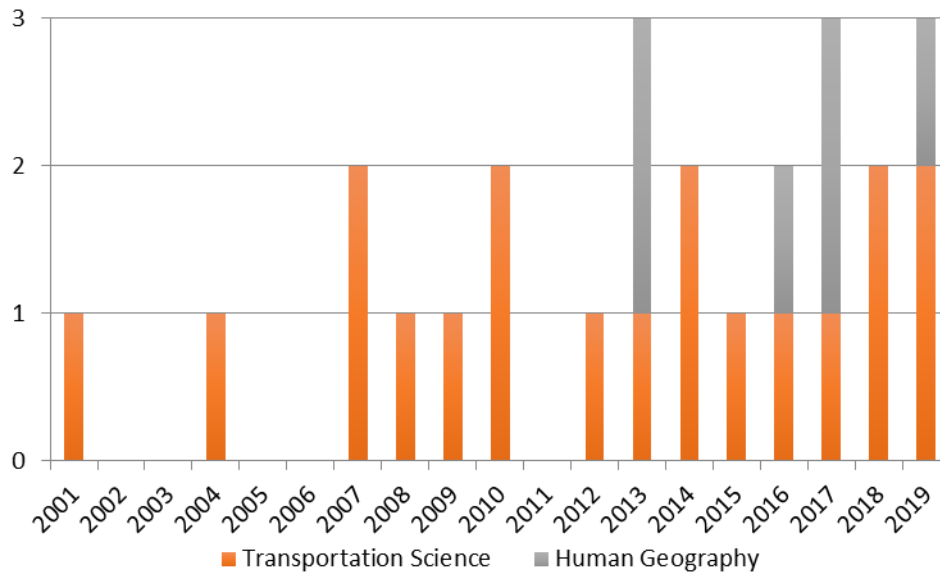


Fig. 2 The number of selected papers according to year of publication and domain

3 Defining trip purpose imputation from GPS data

3.1 In the context of GPS-based travel survey revolution

The concern over missing trips in conventional data collection techniques triggered the use of GPS in the early 2000s (Chen et al., 2014; Shen and Stopher, 2014). GPS enables passively, continuously recording movement and activity at high resolution (Auld et al., 2009; Wolf et al., 2014), and thus it aids effectively in measuring (Forrest and Pearson, 2005; Wolf et al., 2003) and correcting (Bricka and Bhat, 2006) the influences of the missing trip problem in self-reported travel surveys. People tend to neglect trips they consider unimportant (Richardson et al., 1996); therefore, purposes are always inferred prior to audit and estimate impacts of underreported trips.

At the beginning of the 21st century, GPS loggers were installed on vehicles to take advantage of on-board electricity, resulting in the non-existence of mode detection problems and scrupulous attention to inferring trip purpose. Nevertheless, with the recent availability of person-based multimodal GPS data collected by wearable devices or smartphones, transportation mode identification has become the heart of GPS-based research.

The secondary role of purpose detection results from its complex nature. In contrast to visible

123 transportation modes, inferring purposes involves seeking the most probable activity at a specific
124 location type. In fact, different people can engage in dissimilar activities at the same place in cases of
125 multiple functionality (e.g., a shopping mall with restaurants and a cinema inside). Additionally, while
126 movement parameters (e.g., instantaneous speed, acceleration), which are usually provided by devices
127 or easily calculated from coordinates and timestamps, are sufficient for labelling modes (Feng and
128 Timmermans, 2013; Nguyen et al., 2019a; Schuessler and Axhausen, 2009), purpose prediction
129 depends largely on external sources like land use data (Oliveira et al., 2014; Xiao et al., 2016). GPS
130 traces and/or GIS data may not be accurate enough to recognize the places of activities (Meng et al.,
131 2017). Shortage of data also prevents researchers from conducting more studies. A trip comprises
132 some single-mode segments (i.e. stages/trip legs) but has only one purpose.

133 3.2 *In research domains*

134 GPS data are useful for research in two domains, namely Transportation Science (TS) and Human
135 Geography (HG).

136 3.2.1 *Transportation Science*

137 Transportation Science (TS) studies persons' or objects' movement from place to place by an array of
138 methods like physics, operation research, probability, control theory (Hall, 2003). Trip purposes coupled
139 with their respective locations are useful since they are predictors of mode choice and travel forecast
140 models (Burbidge and Goulias, 2009; Ho and Mulley, 2013).

141 Because of long-lasting and unresolved drawbacks of data quality collected in conventional travel
142 surveys (Chen et al., 2016), TS researchers pay assiduous attention to the potential of GPS for
143 supplementing and even replacing the traditional methods of data collection (Auld et al., 2009). That is
144 to say, the focus of TS is on developing and validating methods of deriving trip purpose and other
145 characteristics from GPS data.

146 The majority of purpose imputation studies are found within TS (see Table 1). In this domain,
147 developing methods are undertaken after travel surveys are completed (i.e. post-collection analysis).
148 Twelve out of 19 studies gather data using personal devices or smartphones ((Bohte and Maat, 2009;

149 Chen et al., 2010; Cui et al., 2018; Deng and Ji, 2010; Feng and Timmermans, 2015; Gong et al., 2018;
150 Montini et al., 2014; Oliveira et al., 2014; Shen and Stopher, 2013; Stopher et al., 2008; Xiao et al.,
151 2016; Yazdizadeh et al., 2019). TS studies is to regularly collect ground truth, which refers to confirmed
152 travel diaries corresponding to GPS streams, through prompted recall surveys following up
153 GPS-enabled ones (Auld et al., 2009). A variety of strategies to collect ground truth have been carried
154 out, including the use of paper diaries (Chen et al., 2010; McGowen and McNally, 2007; Oliveira et al.,
155 2014; Reumers et al., 2013; Wolf et al., 2001), websites (Bohte and Maat, 2009; Cui et al., 2018; Deng
156 and Ji, 2010; Feng and Timmermans, 2015; Krause and Zhang, 2019; Lu et al., 2012; Montini et al.,
157 2014; Shen and Stopher, 2013; Stopher et al., 2008), telephone (Xiao et al., 2016), mail (Cui et al.,
158 2018) and smartphone (Yazdizadeh et al., 2019). Another way is to manually label activities by means
159 of visualizing trip ends in association with POI (i.e. Point Of Interest) data (Gong et al., 2018; Liao et al.,
160 2007).

161 3.2.2 *Human Geography*

162 Human Geography (HG), a major discipline of the subject field of geography (Gibson, 2009), studies the
163 interrelationships between people, place, and environment and how they vary spatially and temporally
164 (Castree et al., 2013). It places attention on elements of human activity and organization, such as
165 culture, urbanization, population and transport (Castree et al., 2013; Fang et al., 2017).

166 Similar to TS, GPS-based purpose imputation studies in HG undertake post-collection analyses in
167 order to acquire general knowledge on mobility and whereabouts of activities. They do so by
168 segmenting trajectories into parts of stopping and moving before semantically enriching them through
169 inference algorithms, usually called the semantic enrichment process (Furletti et al., 2013; Gautama et
170 al., 2017; Prelipcean et al., 2017).

171 Compared to studies in TS, GPS-based purpose imputation studies in HG are fewer and published
172 within the last decade (see Figure 2 and Table 1). Most HG studies use data from specific dweller
173 groups (e.g., cyclists [Usyukov, 2017] or taxi passengers [Chen et al., 2019; Gong et al., 2016; Wang et
174 al., 2017]) gathered by on-board devices (Chen et al., 2019; Furletti et al., 2013; Gong et al., 2016;
175 Wang et al., 2017). Ground truth is optional because the common sense or the travel patterns from

176 previous surveys are sufficient for validation (Chen et al., 2019; Gong et al., 2016; Usyukov, 2017;
 177 Wang et al., 2017). In Furletti et al. (2013) and Reumers et al (2013), ground truth is collected, enabling
 178 application of learning methods to semantically annotate activities to GPS traces. The poor prediction of
 179 non-home and non-work purposes in Reumers et al. (2013) and of “services, shopping” in Furletti et al.
 180 (2013) may not be a concern since the main travel patterns are obtained.

181 **Table 1** Overview of purpose imputation studies in Transportation Science and Human Geography

Indicator	Transportation Science	Human Geography
Focus	Methods of deriving purposes from GPS data	Semantic enrichment for GPS trajectories
Validation	Directly by ground truth	Indirectly by previously known knowledge ⁽¹⁾ or by ground truth
Mainly assessed by	Quantitative measurement (e.g., accuracy, precision, recall)	Consistent with previously known knowledge
Main devices used to collect GPS data	Personal devices and smartphones	Devices attached to vehicles
Data source	Travel surveys	Travel surveys and taxi (probe) data
Studies for this review	19: (Bohte and Maat, 2009; Chen et al., 2010; Cui et al., 2018; Deng and Ji, 2010; Feng and Timmermans, 2015; Gong et al., 2018; Krause and Zhang, 2019; Liao et al., 2007; Lu et al., 2012; McGowen and McNally, 2007; Meng et al., 2017; Montini et al., 2014; Oliveira et al., 2014; Shen and Stopher, 2013; Stopher et al., 2008; Wolf et al., 2004, 2001; Xiao et al., 2016; Yazdizadeh et al., 2019)	6: (Chen et al., 2019; Furletti et al., 2013; Gong et al., 2016; Reumers et al., 2013; Usyukov, 2017; Wang et al., 2017)

182 **Note:** ⁽¹⁾ refers to common sense and information extracted from large-scale travel household surveys.

183 4 Reviewing GPS data-based purpose imputation process

184 Purpose imputation, which is a classification problem, is addressed by a three-step process, including
185 (1) trip end detection, (2) feature selection, and (3) main algorithm and validation.

186 4.1 *Trip end detection*

187 Rule-based routine is the key to handling the trip detection challenge. The minimum threshold of
188 duration of the nearly stationary status of the observed object is heavily utilized. The 120s threshold
189 yields the best trip identification performance in Wolf et al. (2001), agreed by Deng and Ji (2010) and
190 Stopher et al. (2008), but smaller than 180s in Bohte and Maat (2009), Krause and Zhang (2019), and
191 Yazdizadeh et al. (2019) and 300s in Wolf et al. (2004). The highest is 10 minutes in Usyukov (2017).
192 Shen and Stopher (2013) and Gong et al. (2018) do not report the cut-off value. The low and high
193 thresholds coincide with over-segmentation and under-segmentation, respectively.

194 Cui et al. (2018) manually examine both GPS data and reported travel diaries to merge or divide
195 trajectories into trips and activities, enabling elimination of spurious activities. Liao et al. (2007) detect
196 trip ends by using street maps to develop a complex probabilistic method. Clustering points is
197 implemented in Chen et al. (2010) and Montini et al. (2014).

198 In case of in-vehicle devices used, a trip end is detected through signals showing the change in
199 engine status between on and off (Lu et al., 2012). However, this method is not reliable enough, and
200 thus the three-minute rule of stability is applied (Krause and Zhang, 2019; Wolf et al., 2004). For taxi
201 trips, a trip end is easily detected as a drop-off point (Chen et al., 2019; Gong et al., 2016; Wang et al.,
202 2017).

203 Identifying trip ends does not exist in McGowen and McNally (2007) because the authors geo-code
204 origins and destinations of trips in the California household travel survey. Geo-coded data act as a
205 surrogate of GPS data.

206 4.2 *Feature selection*

207 All features are classified into four categories, including (1) geographic data, (2) activity-related and

208 trip-related, (3) participant-related and (4) others.

209 4.2.1 Geographic data

210 Geographic data are employed most frequently to infer activity types. Except for Montini et al. (2014)
211 and Reumers et al. (2013), all other studies utilize GIS-related features. In studies developing random
212 forest models, which compute and order variables by their importance (Gong et al., 2018; Montini et al.,
213 2014; Yazdizadeh et al., 2019), variables based on the distances to POIs are the most significant. Gong
214 et al. (2018) and Oliveira et al. (2014), by dropping out and employing again features, emphasize that
215 adding spatial features considerably improves activity predictions, especially for non-mandatory
216 purposes like eating out, banking and shopping (Oliveira et al., 2014). This argument is supported by
217 Reumers et al. (2013), who reach only 20%, 26.9% and 7.1% success rates for social, shopping and
218 leisure activities respectively due to lack of spatial features. In Lu et al. (2012), a land use-based model
219 successfully labels 72.3% of total cases while those based on social demographics only and
220 previous/next trip attributes solely produce accuracies of 51.9% and 60.6%, respectively.

221 The threshold of distance between a trip end and a place varies among studies, possibly resulting
222 from the combination of four factors, namely (1) GIS data types, (2) device types, (3) place types and (4)
223 quality of GPS data. An explanation of these factors follows:

224 • First, GIS data, which are available for regions or downloaded directly from Google Places (Cui
225 et al., 2018; Furletti et al., 2013; Meng et al., 2017) or OpenStreetMap (Furletti et al., 2013; Krause and
226 Zhang, 2019), exist in two forms, that is, a point representing a (small) place (i.e. POI) or a polygon
227 representing a large place. Once polygon-shaped information is used, if a trip end is perfectly within a
228 land use area (Deng and Ji, 2010; Xiao et al., 2016) or within a short distance of it (e.g., 25m of a
229 commercial area in Oliveira et al. [2014]), the type of this area is assigned to the trip end. Once POIs are
230 used, the higher distance thresholds are determined, shorter than 150m to a church in Oliveira et al.
231 (2014), for example.

232 • Second, unlike personal devices, onboard devices fail to access within places. Consequently,
233 the distance thresholds, which are indicated as the maximum walkable distance between an
234 origin/destination and a parking lot (for car data) or a drop-off point (for taxi data), tend to be set at a

235 (very) high value of 1000m (Furletti et al., 2013), 500m (Lu et al., 2012) or 300m (Krause and Zhang,
236 2019). Since taxi is a nearly door-to-door service, the threshold would be smaller (e.g., 200m in Gong et
237 al. [2012]).

238 • Third, distance to known places and often visited could be higher than that to unknown ones.
239 Bohte and Maat (2009) determine 50m for most POIs and 100m for homes and workplaces whose
240 addresses are provided by participants.

241 • Fourth, for data with limited quality, the distance thresholds to POIs are high—for instance, at
242 250m in New York due to strong urban canyon effects (Chen et al., 2010) or 150m owing to the use of
243 low-cost devices (Usyukov, 2017).

244 Original types of POIs and polygon-shaped locations vary across cities/regions and are much more
245 numerous than activity types, which indicates a need for grouping them (Deng and Ji, 2010; Gong et al.,
246 2016; Xiao et al., 2016; Yazdizadeh et al., 2019). Twenty polygon-level and 18 POI-level land use types
247 in Shanghai (China), for instance, are aggregated into 10 categories (Xiao et al., 2016). Street maps are
248 used by Liao et al. (2007) to support trip end detection.

249 Schedules of land use types, which are not employed in all TS studies, are deployed in HG ones (see
250 Table 2). A location will be disregarded if it closes when a person visits it (Furletti et al., 2013; Gong et
251 al., 2016). Working time is unavailable in GIS data, leading Furletti et al. (2013) and Gong et al. (2016) to
252 assign it to each type. To give an illustration, opening hours of shopping malls, government agencies
253 and bars will be 9:00am–10:00pm, 8:00am–6:00pm and 2:00pm–03:00am, respectively (Gong et al.,
254 2016).

255 4.2.2 *Trip and activity related*

256 An activity is the main business undertaken in a significant location; hence, it is in harmony with a
257 purpose. A trip is a one-way course of travel to conduct an activity. Trip and activity have a close
258 relationship (Ho and Mulley, 2013). Temporal attributes of trip and activity (i.e., time of day or day of
259 week) are the same and listed as features of activity. Yet, mode, speed, distance, start time and end
260 time characterize a trip. Except for Furletti et al. (2013), other authors deploy features related to either
261 activity or trip or both to construct their models.

262 As for activity, the four characteristics, namely duration, time of day, day of week and start time are
263 used most frequently (see Appendix); however, some of them rather than all are integrated in each
264 study. If home type is neglected, duration and start time are sufficient for the models (Feng and
265 Timmermans, 2015; Oliveira et al., 2014). The reasons would be that start time may be equivalent to
266 time of day (e.g., morning, afternoon, night) and the confusion of duration between work and home does
267 not exist. Activity history-related features occur in two studies (Chen et al., 2010; McGowen and
268 McNally, 2007). Chen et al. (2010) conclude that the frequency of visiting locations, estimated from
269 one-day data, carries little meaning. Based on the frequency of visiting a specific location per day,
270 McGowen and McNally (2007) estimate the accumulated duration and give value to a dummy variable
271 (1 for re-visiting case). Temporal profiles of activities are sufficient for detecting basic purposes (such as
272 work and home) (Reumers et al., 2013). Shen and Stopher (2013) propose tour²-based corrections to
273 enhance accuracy by approximately 8%. Notably, activity information is not available for HG studies
274 using taxi data (Chen et al., 2019; Gong et al., 2016; Wang et al., 2017).

275 Trip attributes are employed in Deng and Ji (2010), Feng and Timmermans (2015), Gong et al.
276 (2018), Lu et al. (2012), Montini et al. (2014), Oliveira et al. (2014), Xiao et al. (2016) and Yazdizadeh et
277 al. (2019). Cui et al. (2018) use trip characteristics but not activity ones to impute purposes. Importantly,
278 mode information is taken directly from ground truth in all existing studies while mode and purpose must
279 be imputed in GPS data processing (Shen and Stopher, 2014). In this sense, mode detection and
280 purpose imputation are now considered separately, although they should be combined in a continuous
281 process. Travel mode is not included in HG studies since most of them use vehicle-based data (Chen et
282 al., 2019; Furletti et al., 2013; Gong et al., 2016; Usyukov, 2017; Wang et al., 2017).

283 4.2.3 *Participant related*

284 Personal information is used in TS studies but not in HG ones, many of which employ anonymous data
285 (i.e., taxi data) (Chen et al., 2019; Gong et al., 2016; Wang et al., 2017) (see Table 2).

286 Addresses of home and work, which are seen in Bohte and Maat (2009), Gong et al. (2018), Montini
287 et al. (2014), Shen and Stopher (2013), Stopher et al. (2008), Wolf et al. (2004) and Yazdizadeh et al.

² Tour is defined as trips and activities occurring between a participant leaving home and returning home (Shen and Stopher, 2013)

288 (2019), help in boosting considerably the overall classification performance since the number of work
289 and home activities is by far greater than the number of other purposes. In case home address is not
290 provided, home location can be estimated satisfactorily by examining the first trip's origin and the last
291 trip's destination every day (Usyukov, 2017).

292 Among socio-demographics, age and occupation (Cui et al., 2018; Deng and Ji, 2010; Gong et al.,
293 2018; Montini et al., 2014; Oliveira et al., 2014; Xiao et al., 2016; Yazdizadeh et al., 2019) are used
294 much more frequently than the remainder (see Appendix).

295 In comparison with the two above feature categories, participant-related features are less informative
296 for directly distinguishing activities and play a supportive role in increasing the classifier power (Lu et al.,
297 2012; Shen and Stopher, 2013). Without participants' information, 96.8% of activities of 329 people in
298 the Netherlands are correctly identified by Feng and Timmermans (2015).

299 4.2.4 *Other features*

300 In TS, Gong et al. (2018) test the use of weather-based variables (e.g., temperature, precipitation and
301 snow accumulation) in combination with the three above-mentioned feature categories by data collected
302 in two distinct seasons in Hakodate, Japan. The authors report that weather features decrease the
303 model accuracy without further explanation of reasons. This is a negative and surprising result given the
304 significant relationship between weather and mobility in the literature (Böcker et al., 2013; Cools et al.,
305 2010; Liu et al., 2017).

306 Social media (e.g., Twitter and Foursquare) information is useful for enhancing purpose inference in
307 both TS (Cui et al., 2018; Meng et al., 2017; Yazdizadeh et al., 2019) and HG (Chen et al., 2019; Wang
308 et al., 2017). Since social media data are noise, an information retrieval process is first conducted to find
309 a place (i.e. POI) matching each post (e.g., tweet or check-in) (Cui et al., 2018; Meng et al., 2017). In this
310 way, each post can be arranged into a category (e.g., food, education, leisure [Meng et al., 2017]).
311 Then, the number of posts that belong to a specific category and are close to a trip end (e.g., with
312 distance less than 250m [Yazdizadeh et al., 2019]) is used as a predictor variable. While studies in TS
313 use social media data in tandem with geographical data achieved from other sources (e.g., Google
314 Places and OpenStreetMap), city-wide POIs and their categories are achieved directly from Foursquare

315 check-in data in HG (Chen et al., 2019; Wang et al., 2017). It should be noted that using social media
 316 data creates a lack of robustness owing to a variety of social media services and coverage. To give an
 317 illustration, in China, Weibo, a domestic social network, is far more common than Twitter or Facebook.
 318

Table 2 Summary of using input features in two domains

Feature category	Specific features	TS	HG
Geographic data	Polygon-based; POI; Street map	√	√
	Working time of land use types	-	√
Activity-related	Duration, Time of day, Day of week, Start time	√	√
	Activity history	√	√
Trip-related	Travel mode, Mode of next trip, Mode of previous trip	√	-
	Speed, Distance, Duration, Start time, End time	√	√
Participant-related	Home address, Work address, Occupation, Age, School address, Frequently visited places, Working hours, Gender, Driving license, Employment status, Income, Education degree, Race, Family structure, Household information, Marital status, Driving frequency	√	-
Others	Social networking: Foursquare, Twitter	√	√
	Weather: Temperature, Precipitation, Snow accumulation	√	-

319 4.3 Main algorithm and validation

320 This study follows the classification of Gong et al. (2014) to review three method categories in turn,
 321 namely (1) rule-based, (2) probability-based and (3) machine learning, before discussing how to
 322 validate prediction results.

323 4.3.1 Rule-based method

324 Deterministic rules were the most common method for the initial era of purpose imputation because they
 325 are simple and easily interpretable. On the basis of polygon-based land use data and temporal
 326 characteristics of activities, Wolf et al. (2001) infer 10 purpose types to trips by 13 participants in Atlanta
 327 (USA) with a 79.5% accuracy. Using addresses for home, work and two frequently visited grocery

328 stores, Stopher et al. (2008) report a set of rules to achieve an accuracy of over 60%. They improve their
329 rules by deploying tour information to correctly classify 66.5% of 4,133 trips in the Greater Cincinnati
330 region (USA) (Shen and Stopher, 2013). The lowest accuracy is 43%, reported by Bohte and Maat
331 (2009), who use data of 1,104 respondents in the Netherlands to classify seven purposes by
332 participants' home and work addresses together with a geographic database. They assign a location
333 type to a trip end if the location has the shortest distance to the trip end. However, due to the error of
334 GPS points, the closest would not be the right choice. A notable point from these four studies is that the
335 accuracy is inversely proportional to the sample size, explained by an increase in the heterogeneity of
336 travel patterns and activities in diverse samples. What's more, the quality of rules hinges on experts'
337 knowledge of transport conditions in the research area. Rules are the least transferable since daily lives
338 and travel patterns in different environments vary, requiring great efforts to re-calibrate and re-set up
339 thresholds. Deterministic approaches have been used to construct the first steps of deriving purposes
340 from GPS data. It is now integrated as a way to choose potential locations of a trip end or to estimate
341 location of home for both TS and HG studies (Furletti et al., 2013; Usyukov, 2017; Xiao et al., 2016) (see
342 Table 3).

343 4.3.2 *Probability-based method*

344 Probabilistic approaches are more flexible than rules because they simultaneously take into
345 consideration spatial and temporal confusion between purposes.

346 In TS, the first effort is estimating probability of each POI based on distance to cluster of trip end (Wolf
347 et al. 2004) rather than choosing the closest like Bohte and Maat (2009). Nevertheless,
348 socio-demographic variables are included in a deterministic way. Chen et al. (2010) introduce a
349 multinomial logit model to analyze 49-person data in high-density areas of New York and classify four
350 purposes into home-based group and non-home-based group with accuracies of 67% and 78%,
351 respectively. Using data from 1,352 participants and more variables with the addition of trip-related
352 variables, Oliveira et al. (2014) create a two-level nested logit model to distinguish 12 purposes with an
353 accuracy of 60% at the expense of a time-intensive computation. Therefore, statistical methods would
354 be better than rule-based ones in dense environments with high confusion between location types, yet

355 they become much more complex and fail to adequately address a list with various purposes.

356 In HG, probability-based methods are the most common. Similar to Chen et al. (2010), Usyukov
357 (2017) develops a multinomial model using start time and duration of activity to discriminate work
358 activities from others. Furletti et al. (2013) propose temporal and spatial rules to choose a set of
359 potential locations corresponding to each stop of a car. Only locations that are within a
360 1,000m-radius-buffer and open when a stop occurs are kept. Subsequently, the probability of each
361 location is estimated based on the gravity model. The chosen location type has the largest probability.
362 Gong et al. (2016) re-use the temporal and spatial compatibility criteria, which are integrated into
363 Bayes's rules with distance decay effect, and attractiveness of POI (i.e. service capability, size and
364 fame) to create a probability function of visiting each POI. Wang et al. (2017) utilize probabilistic
365 distributions (multinomial and Dirichlet) to first categorize POIs into 10 topics (e.g., shopping, university,
366 office) based on the similarity of functionality, and then assign topics to both origin and destination of
367 each trip. The addition of time helps in uncovering the semantic meanings of trip purposes.

368 Thresholds of (working) time, on the one hand, are impossibly set for every place, resulting in
369 misclassification for specific cases. On the other hand, annotating activities with a degree of
370 approximation is accepted for the HG authors, who aim at semantically enriching trajectories and
371 discovering general activity patterns rather than focusing on precision at an individual level. Thanks to
372 their flexibility, probabilistic methods are more powerful than rule-based ones. Rules are specific cases
373 of probabilities. Compared with rules, statistical algorithms are less dependent on the subjective
374 knowledge of researchers and are thus more transferable. They are capable of deducing purposes from
375 far larger data sets (e.g., one-week data of 6,600 taxis [Gong et al., 2016]) (see Table 3).

376 4.3.3 *Machine learning method*

377 Supervised machine learning (SML) algorithms are trendy in TS and considered a better choice than
378 both deterministic and probabilistic methods. Their evolution can be divided into two main periods.

379 The period between 2001 and 2010 has witnessed decision tree models (C4.5 [McGowen and
380 McNally, 2007] and C5.0 [Deng and Ji, 2010]). A more complex process, hierarchical conditional
381 random field, is proposed mainly to detect stops (Liao et al., 2007). Most of these studies are

382 small-scale tests. Deng and Ji (2010) use 226 trips by 36 participants while Liao et al. (2007) use data
383 from three persons to train their model before validating it with the trips of a fourth person. With small
384 sizes, samples would be homogeneous, resulting in high accuracy levels of 85.2%–90.6% (Liao et al.,
385 2007) or 87.6% (Deng and Ji, 2010). With a larger sample of 17,000 households, McGowen and
386 McNally (2007) report the precision at 73%–74% for five major activities and only 63%–64% for
387 disaggregated purposes.

388 Since 2011, a boom in building up and enhancing a host of SML algorithms has been seen. Lu and Liu
389 (2012) and Oliveira et al. (2014) continue employing decision tree and fail to generate an accuracy of
390 over 80%. In contrast, Montini et al. (2014) pioneer a random forest model correctly classifying 84.4% of
391 total trips by 156 persons in Zurich, Switzerland. Notably, random forest has replaced decision tree to
392 become the most-utilized method (Feng and Timmermans, 2015; Gong et al., 2018; Montini et al., 2014;
393 Yazdizadeh et al., 2019). It hits 96.8% accuracy with 329-person data in the Netherlands (Feng and
394 Timmermans, 2015). However, it tends to encounter bias towards the categorical variables with many
395 more levels (Deng et al., 2011) and the activity types occurring more frequently in training sets (Montini
396 et al., 2014).

397 Some authors display impressive performances with improved neural networks. Xiao et al. (2016) use
398 particle swarm optimization instead of back propagation in artificial neural network structure to obtain a
399 96.5% accuracy. By estimating parameters between nodes of different layers using probability
400 distribution based on Bayesian inference rather than fixed values, Cui et al. (2018) successfully label
401 90.5% of activities.

402 Evaluating SML algorithms across studies would be questionable owing to the differences in purpose
403 lists, features used, sizes and homogeneous levels of samples, not to mention the effects of the trip
404 under-reporting problem in prompted recall surveys. However, in some studies reporting performances
405 of multiple methods (Cui et al., 2018; Feng and Timmermans, 2015; Meng et al., 2017; Xiao et al.,
406 2016), comparisons are feasible. Unsurprisingly, newly proposed models yield the best performances.
407 Random forest has the second highest accuracy (Cui et al., 2018; Meng et al., 2017), while support
408 vector machine only generates reasonable accuracy levels, which are much lower than the highest (Cui
409 et al., 2018; Meng et al., 2017; Xiao et al., 2016).

410 Contrasting opinions about the power of Bayesian networks are reported. Observing the accuracy of
411 46.2%, Feng and Timmermans (2015) conclude that it is not appropriate for inferring purpose,
412 contradicting the high accuracy levels of 87.8% in Meng et al. (2017) and 86.6% in Xiao et al. (2016).

413 Recently, with the view that a purpose, which is decided at origin, will determine the trip end type,
414 Krause and Zhang (2019) integrate a decision tree-based purpose inference into a model aimed at
415 predicting vehicle destination. The purpose of a trip is inferred by features related to the trip origin before
416 the derived purpose is used to estimate the trip destination.

417 Based on the comparison of Oliveira et al. (2014), probability-based models are outweighed by SML
418 algorithms. SML is the most transferable because it learns from data to make predictions instead of
419 being explicitly programmed by rules. In addition, it is suited to big data like GPS and GIS data as well as
420 a host of input variables (Gong et al., 2018; Xiao et al., 2016; Yazdizadeh et al., 2019). For example, 29
421 features are deployed in Cui et al. (2018). As for SML's shortcomings, they need ground truth to train
422 models and numerous data to avoid the overfitting problem. In that SML is a complex structure with
423 many hyper-parameters (e.g., learning rate, number of trees), it is challenging to interpret magnitude of
424 variables. A typical matter for SML is how to form training and test sets effectively (see Table 3). The
425 common way is by dividing equal proportions of observations belonging to each purpose (e.g., 75% for
426 training and 25% for testing [Montini et al., 2014]). However, Gong et al. (2018) highlight that data in
427 training and test sets coming from different, distinct seasons weaken prediction results.

428 Authors in HG develop both SML and unsupervised machine learning (UML) models. Reumers et al.
429 (2013) introduce a decision tree model, employing temporal features, to impute six purposes. Without
430 ground truth, Chen et al. (2019) propose a novel combination of autoencoder and K-means.
431 Autoencoder is an unsupervised neural network whose encoder maps a 22-dimension (i.e. original
432 features) input vector to a more representative vector with 10 dimensions. Then, each trip that is
433 represented by a 10-dimension vector is classified into one among five clusters (i.e. purposes) by
434 K-means algorithm. One hundred thousand trips are used to train the autoencoder network. Because
435 there is only one study, it is difficult to assess the transferable level and the importance of splitting data
436 for UML. UML's shortcomings are complex structure with a need for big data coupled with low
437 interpretation. Furthermore, its prediction is validated reasonably at aggregated level only (see Table 3).

Table 3 *Synthesis of purpose inference methods*

	Rule-based	Probability-based	Supervised machine learning	Unsupervised machine learning
Transferable level	Low	Medium	High	-
Suitable data size	Small	Medium and big	Big	Big
Role of data selection/division	Minor	Minor	Major	-
Number of variables	Small	Small	Large	Large
Ground truth	Optional	Optional	Mandatory	No
Mainly used in	Estimating home address; Choosing location candidate	Human Geography	Transportation Science	Human Geography
Power/performance	Low	Reasonable	High	Reasonable
Interpretation	High	Medium	Low	Low

439 4.3.4 Purpose list and validation

440 Prediction assessment is based on counting the matching times between the prediction and the ground
441 truth for each category before summarizing them in a confusion matrix (Sokolova and Lapalme, 2009).
442 The main measurement indexes are precision and recall for purposes and accuracy for the model.
443 While the majority of studies report the confusion matrix, some (Chen et al., 2010; Lu and Liu, 2012)
444 present the precision levels at aggregate levels only. The accuracy is not reported in Gong et al. (2018)
445 or Krause and Zhang (2019).

446 Generally, the overall accuracy is inflated by identification of home trips, which are inferred at a high
447 success rate and account for the largest percentage. When home activities are taken into consideration,
448 the accuracy increases by 15% to 80% (Oliveira et al., 2014).

449 How to define purposes is scarcely presented. Most authors do not pay attention to mode transfer
450 activity, although Montini et al. (2014) and Oliveira et al. (2014) do. Transferring from one mode to
451 another is the purpose of a separate movement stage but not of a trip connecting between two
452 significant locations. Because transition between stages during a trip is frequent, the inclusion of mode

453 transfer exaggerates the model performance.

454 In a case of unavailable ground truth, Usyukov (2017) uses data from the 2006 Ontario Transportation
455 Tomorrow Survey, which is believed to sufficiently represent the travel patterns of inhabitants. Because
456 the shares of purposes extracted from GPS data are nearly in line with those obtained from the
457 household survey, the model performs satisfactorily. The consistency between spatio-temporal patterns
458 extracted from GPS data and the common distribution of activities according to time and space is
459 deployed to validate models (Chen et al., 2019; Gong et al., 2016; Wang et al., 2017). To provide
460 illustrations, the spatial distribution of home purpose is positively correlated with that of residential
461 areas, and the direct distribution of all trips aligns with the geometry of central business districts in the
462 research area (Gong et al., 2016). Since trips for the same purpose tend to have similar temporal
463 distribution, the model of Wang et al. (2017)—with its smaller estimate of average time difference of trips
464 annotated identically, as compared to its counterpart generated by a previous method—is validated.

465 **5 Discussion**

466 *5.1 Feature selection*

467 The above review emphasizes the great importance of geographic databases in purpose imputation, but
468 relying on such databases would be the biggest barrier to carrying out purpose detection in areas where
469 facilities and locations are not geo-coded well. GIS data, which themselves are numerous, are
470 categorized differently in different regions, requiring great efforts to merge them with GPS data,
471 especially for the simultaneous use of GIS data of several regions. In fact, to limit the complexity of
472 problems and amount of computation, almost all existing studies involve experiments conducted in a
473 specific area and neglect purpose imputation for long-distance trips. Krause and Zhang (2019) stress
474 that including long-distance trips may exert adverse effects on purpose prediction results, but they do
475 not show a detailed figure. Accordingly, it is worth digging and testing more features related to
476 participants, activities and trips. Without geographic variables, a model tends to become more
477 transferable and to be utilized not only to include several areas in a comparative analysis but also to
478 identify purposes of long-distance trips.

479 As for participant-related features, more socio-demographics put a heavier burden on participants
480 and thus result in a higher incomplete survey rate (Armoogum et al., 2014). Additionally, some locational
481 information, such as school addresses, would be sensitive to parents (Feng and Timmermans, 2015).
482 Therefore, a model using the minimum number of these features would be welcomed. Future studies
483 should carry out more tests and report the sensitivity of their models while adding or dropping
484 person-specific variables.

485 Transportation mode and purpose are desired by researchers. In GPS data processing, mode
486 detection is the step prior to purpose imputation (Shen and Stopher, 2014). However, up to now, they
487 have been handled separately. As discussed above, trip mode is informative for inferring purposes, but
488 all existing studies use mode information extracted from ground truth. Therefore, deducing purposes
489 from GPS data with the support of the results of trip identification is promising, likely introducing the
490 potential for incorporating both in an automated process.

491 Using social media information enables an improvement of detection but with lack of robustness.
492 When driving, attending classes or working, people are less likely to tweet compared to when going
493 shopping or eating out. Moreover, social media data would be biased toward young people and against
494 the elderly (Chen et al., 2019). There are a number of social networking services (e.g., Facebook,
495 Twitter, Foursquare, Weibo), and they offer different ways of accessing and exploiting their data.
496 Accordingly, more efforts should be invested in testing different social media networking services to
497 clarify how they complement participant, activity and geography-related variables to help in obtaining
498 satisfactory purpose imputation results.

499 Reports of the negative or marginal roles of some features spur more tests. Historical trip frequency
500 should be extracted from data collected during an adequate period, rather than from one day as in Chen
501 et al. (2010). The negative effects of variables related to seasonal phenomena (e.g., snow, temperature)
502 on purpose imputation results require more attention to draw a comprehensive conclusion and uncover
503 reasons for their influences.

504 *5.2 Data selection and enhancing algorithm*

505 Developing and enhancing machine learning-based methods require masses of data collected over

506 several years (Feng and Timmermans, 2015; Gong et al., 2018), leading to a concern over data
507 selection. First, as a result of the seasonal variability of travel patterns in different weather conditions
508 (Liu et al., 2017), training and test sets should not be formed by data from distinct seasons (Gong et al.,
509 2018). Second, different people display different numbers of each type of activity. Division based on the
510 rate of participants may induce a lack of specific trip types in the training process. In addition, different
511 people tend to carry out the same activity in dissimilar ways; therefore, training and test sets should
512 have a similar distribution of socio-demographics. The most common way is now randomly using the
513 same percentage of each purpose type. This method ignores the logic between consecutive trips in a
514 chain. If many more trips come from several specific participant groups (e.g., the young group), the
515 prediction may be weakened in that the model over-fits these groups. Accordingly, thinking about how to
516 populate training and test sets to ensure similar distribution with respect to seasonal conditions, people
517 and activity would be benefit future researchers.

518 Generally, basic machine learning algorithms (e.g., decision tree, random forest, artificial neural
519 network, support vector machine) have been constructed. To bring about methodological improvement,
520 combining them in hybrid models to limit the long-lasting problem of over-fitting should be considered.
521 Bayesian network may need to be tested in different circumstances to clarify the contrasting reports in
522 Feng and Timmermans (2015), Meng et al. (2017) and Xiao et al. (2016).

523 5.3 *Validation and assessment*

524 The variability of purpose categories prevents cross-evaluation among studies, raising the question of
525 what constitutes an optimal list of activity types. An ideal list may be determined at the minimum level
526 based on the balance between the specific research targets and the requirements of resolution quality
527 for the travel demand forecast (Oliveira et al., 2014). Furthermore, it is worth considering the nature of
528 data instead of using predefined purpose categories. In this regard, UML methods may be applied to
529 classify activities into groups sharing mutual characteristics with unknown labels. The results of these
530 classes would be serious candidates for deciding purpose lists. In any case, the mode transfer purpose
531 should not be included if trip is taken into consideration. To limit the inflation of the overall accuracy, a
532 higher penalty may be imposed on misclassified home cases, or home activities may be disregarded.

533 The error propagation from trip end detection to purpose imputation result would be a useful area for
534 research. Dividing a trajectory is subject to under- and/or over-segmentation problems. While there is no
535 way to detect neglected trip ends in order to fix under-segmentation, over-segmentation is generally
536 preferred (Montini et al., 2014; Shen and Stopher, 2014). The challenge is how to detect a spurious trip
537 end, which is either a mode transfer point or simply a stop at a traffic light. A potential solution is to
538 detect purposes of all potential trip ends with a certain level for each. Subsequently, using the most
539 certain ones as well as logic in trip chain or tour could help in detecting and eliminating wrongly labelled
540 trip ends. The rate of unidentified spurious ends reflects the error propagation level of the trip end
541 detection step.

542 Duration, start time and end time of an activity are estimated from timestamps of GPS points.
543 Estimation from GPS is subject to technical and random errors (Chen et al., 2010; Stopher et al., 2008)
544 while using prompted recall travel diary tends to be prone to under-reporting. Therefore, it is better to
545 compare start time, end time and duration of both to determine the accuracy rather than using the
546 number of matches between detection results and reported diary only.

547 Inferring purpose from GPS surveys is seen predominantly in Belgium, the USA, the Netherlands,
548 Japan, Canada, Switzerland and Australia (Bohte and Maat, 2009; Gong et al., 2018; Stopher et al.,
549 2008; Wolf et al., 2004, 2001; Yazdizadeh et al., 2019), while among developing countries, China is the
550 only venue of such studies (Deng and Ji, 2010; Xiao et al., 2016). Because working, living and transport
551 conditions in emerging countries are distinct from those in developed countries (Nguyen et al., 2019b;
552 Nguyen and Pojani, 2018; Pojani, 2020), it is worth carrying out more experiments in these nations.

553 **6 Conclusions**

554 Purpose imputation from GPS data is a complex problem that a wealth of studies have attempted to
555 address. Here, we provide a critical synthesis of these studies based on the process of addressing a
556 purpose inference challenge and according to two domains, TS and HG. In reality, researchers always
557 confront one of two situations: analyzing a post-collected dataset or designing a survey to create an
558 enhanced inference model. In both cases, positioning the domain of research is vital. For example, a
559 post-gathered dataset without ground truth should be treated as a study in HG to semantically annotate

560 GPS stream. On contrary, if aiming at a methodological contribution, researchers may need a dataset
561 with corresponding travel diary collected in a prompted recall survey. Although applying the same
562 process with trip end detection, feature selection and prediction result evaluation, studies from TS and
563 HG draw up various schemes with both similarities and differences (see Section 4) due to dissimilarity of
564 research focus (see Table 1). This review, therefore, could help with efforts to obtain a clear
565 understanding of purpose imputation in both domains. Furthermore, the unresolved issues mentioned in
566 Section 5 could serve as suggestions for future research. Dealing with these issues is the key to (1)
567 integrating purpose and mode imputation into an automated and continuous process, enabling GPS
568 travel surveys to replace, in part or completely, conventional techniques and (2) gaining a better
569 understanding of how human beings travel from GPS data.

570 **Conflict of interest**

571 The authors do not have any conflict of interest with other entities or researchers.

572 **Acknowledgments**

573 The authors highly appreciate (1) constructive comments of two anonymous reviewers and the editor as
574 well as (2) the support of Ha Anh Nguyen, a PhD student at Department of Civil Engineering under
575 Monash University, as regards searching for publications on SCOPUS database.

576 This paper is a part of Minh Hieu Nguyen's PhD thesis (12/2016-01/2020) under supervision of Dr.
577 Jimmy Armoogum at Université Gustave Eiffel/AME/DEST (France) and co-funded by the Ministry of
578 Education and Training of Vietnam (The educational program 911).

579 Appendix: Summarizing 25 selected papers on purpose imputation

Author	Data, study area, collection device	Trip end detection	Feature selection	Method type	Method description	Accuracy	Ground truth
I. Transportation Science							
1. Wolf et al. (2001)	156 trips; 13 participants; 3-day survey; in Atlanta (US) in 2000 In-vehicle GPS devices	RB	- GI: POI data and polygon-based land use - Activity: start time, duration	RB	Matching table for 10 categories (home, shop, go to work, pick-up/drop-off, change mode, recreation, personal business, eat, unknown)	79.5%	Paper diary
2. Wolf et al. (2004)	39 participants (28 fulltime workers and 11 retirees); 30 days in Borlange (Sweden) in 2002 In-vehicle GPS devices	RB	- GI: POI data and polygon-based land use - Activity: duration, time of day, day of week - Participant: home address, profession, working hours	PB	Estimation of probability based on distance to POI. 10 categories (home, leisure, long-term shopping, daily shopping, work, school, work related, private business, pick up/drop off)	NA	NA
3. Liao et al. (2007)	4 participants; 7 days per person; 40,000 points with 10,000 segments per person	PB	- GI: street map, POI data - Activity: stay duration, time of day, day of week	SML	Hierarchical Conditional Random Fields to recognize activities and places	85.2%-90.6%	Manually label
4. McGowen and McNally (2007)	170,000 activities of 17,000 households per day; 2-day survey, in California (US) in 2000-2001; Geo-coding location reported	NA	- GI: polygon-based land use - Activity: time of day, duration, start time, activity history (repeat duration, repeat location) - Participant: age, gender, driving license, employment status.	SML	Discriminant analysis Classification (regression) tree model 5 activities (school/work, discretionary, coupling, maintenance, egress through wait for vehicle)	73% 74%	Paper diary
5. Stopher et al. (2008)	21 participants in the Sydney Household Travel and 45 ones in pilot Travel Behavior Change Program in Canberra (Australia) Personal devices	RB	- GI: land use - Activity: duration - Participant: home, school, work, frequently visited stores	RB	Heuristic rules. 10 purposes (home-based {work, education, shopping, eating, personal/medical, recreational, pick-up/drop-off, other}, non-home-based {work-other, other-other})	Over 60%	Web-based diary
6. Bohte and Maat (2009)	1,104 participants; one week per person; 3 municipalities in the Netherlands in 2007 Personal devices	RB	- GI: POI and polygon-based data - Participant: home, work address	RB	Heuristic rule based on the closest distance to POI. 7 purposes (work, study, shop, social visit, recreation, home, other)	43%	Web-based diary

7. Deng and Ji (2010)	36 participants; 226 trips; 2.5-month data collection in Shanghai (China)	RB	- GI: polygon-based data - Participant: occupation, income, family structure, age - Activity: weekdays, weekend, time of day - Trip: speed, modes, distance, duration	SML	Decision tree (C5.0) with adaptive boosting 6 purposes (work, school, pick up/drop off, shopping/recreation, business visit, others)	87.6%	Web-based diary
8. Chen et al. (2010)	25 participants for one day and 24 participants for 5 weekdays; in New York (US) Personal devices	Clustering trip ends	- GI: polygon-based land use - Activity: time of day, duration, historical activity frequency	RB & PB	Multinomial logit for high-density areas Deterministic matching for low-density areas 4 purposes (work/school, personal business, shopping, recreation) for two groups (home-based, non-home-based)	67% – 78%	Paper diary
9. Lu et al. (2012)	3,188 trips; Twin Cities Metro Area, Minnesota (US) in 2008 In-vehicle devices	Engine on/off signal	- GI: polygon-based data - Participant: income, race, education level - Trip: start and end time of current, previous and next trip, duration	SML	Decision tree 10 purposes (home, work, shopping, daycare, dining, driving others, services, school, social/recreation, others)	73.4%	Web-based diary
10. Shen and Stopher (2013)	4,133 trips; Greater Cincinnati region (USA) in 2009 Personal devices	NA	- GI: land use - Activity: duration, start time, end time, tour - Participant: home, school, work, frequently visited stores	RB	Heuristic rules. 5 purposes (home, work, education, shopping, other)	66.5%	Web-based diary
11. Oliveira et al. (2014)	1,352 participants (subsample of Atlanta Household Survey); 22,734 activities (10,512 non-home ones); in 2011 in Atlanta (US) Personal devices	NA	- GI: land use - Activity: duration - Participant: occupation, age, household's information, etc. - Trip: trip mode, mode of next trip, etc.	PB & SML	Two-level Nested Logit Model Decision tree 12 purposes (<i>mode transfer</i> , work, pick up, drop off, maintenance, work, work-related, attending class, shopping, eating, religious ones, entertainment, social visit)	60% 65%	Paper diary

12. Montini et al. (2014)	156 participants; 6,938 activities; in 2012 in Zurich (Switzerland) Personal devices every 1s	Clustering trip ends	- Activity: day of week, duration, start time. - Participant: education level, age, income, marital status, home and work address - Trip: mode before and after activity	SML	Random forest. 8 purposes (<i>mode transfer</i> , home, work, shopping, recreation, business, pick up/drop off, others)	84.4%	Web-based diary
13. Feng and Timmermans (2015)	329 participants; 10,545 activities except "return home" purpose; 2012-2013 in Rijnmond Region (the Netherlands) Personal devices every 3s	NA	- GI: POI data - Activity: start time, duration - Trip: mode (walk, bike, bus, car, taxi, tram, metro and train)	SML	Random forest Decision tree Bayesian belief network 11 purposes (paid work, daily shopping, help parents/children, non-daily shopping, recreation, social, voluntary work, service, leisure, pick up, study)	96.8% 69.8% 46.2%	Web-based diary
14. Xiao et al. (2016)	321 participants; 2,409 person days; 7,039 activities; 10/2013-06/2015 in Shanghai (China) Smartphones every 1s	NA	- GI: POI and polygon-based data - Participant: age, gender, income, occupation, child in household - Activity: duration, time of week, start time - Trip: mode (walk, bike, e-bike, bus, car)	SML	Artificial neural network with particle swarm optimization Artificial neural network with back propagation Support vector machine Bayesian network Multinomial logit 6 purposes (home, work/education, eat out, shopping, social visit, pick up/drop off)	96.5% 93.8% 87.1% 86.6% 80.2%	Telephone-based collection
15. Meng et al. (2017)	8,631 GPS trajectories in California Travel Household Survey	NA	- GI: POI data from Google places - Activity: time and duration, previous activity - Trip: modes - Social network: Twitter	SML	Bayesian network Random forest K-nearest neighbour Artificial neural network Support vector machine	87.8% 60.1% 38.5% 45.4% 33.7%	Available but not reported
16. Gong et al. (2018)	20 participants; 9,981 trips (5,512 in summer & 4,469 in winter); 12/2012-04/2013 in Hakodate (Japan) Smartphones every 30s	RB	- GI: POI data - Participant: home & work addresses, gender, age, household's information (size, vehicle number), driving frequency - Activity: time of day, day of time, time of week - Trip: duration, length - Weather: temperature, snow accumulation and precipitation	SML	Random forest 6 purposes (home, commute, meal, shopping, recreation, others)	Only for separate purposes	Manual labelling

17. Cui et al. (2018)	10,474 participants (42,431 households); Bay Area, California (US) in 2012-2013 Personal and in-vehicle devices	Comparison between GPS data and ground truth	- GI: POI data from Google places - Participant: age, occupation/employment - Trip: departure time, duration, distance - Social network: Twitter	SML	Bayesian neural network Support vector machine Artificial neural network K-nearest neighbor Random forest 6 purposes (eat out, personal, recreation, shopping, transportation, education)	90.5% 49.9% 48.2% 45.1% 61.9%	Web-based and mail-based diary
18. Yazdizadeh et al. (2019)	6,845 participants with validated data; 102,904 trips; 10-11/2016 in Montreal (Canada) Smartphones	RB	- GI: land use data - Participant: age, occupation, gender, home's neighbourhood value, home/work/education addresses - Activity: day of week, time of day - Trip: travel time, modes, origin and destination in comparison with Montreal Island - Social network: Foursquare	SML	Random forest 6 purposes (education, health, leisure, shopping, return home, work)	71.26%	Prompted recall by smartphone
19. Krause and Zhang (2019)	260 participants, 36,000 trips, 70 days; 10/2011-02/2012 in 22 states of US On-board devices every minute	RB	- GI: POI data - Participants' characteristics - Trip: start-of-trip information	SML	Decision tree (J48) 7 purposes (home, school, shopping, social, work, driving, other)	NA	Online web-based diary
II. Human Geography							
20. Furletti et al. (2013)	28 participants; around 30,000 trips; in Flanders (Belgium) In-vehicle devices	NA	- GI: POI data, opening/working time	PB	Gravity model based on distance 6 purposes (services, food, daily shopping, shopping, education, leisure)	43%	Daily diaries
21. Reumers et al. (2013)	1,250 households; one-week survey; 290 activities (test set); in Flanders (Belgium) in 2006 – 2007 GPS-enabled Personal Digital Assistants (PDA) every 1s	NA	- Activity: start time, duration	SML	Decision tree. 6 purposes (home, work, bring-get, leisure, shopping, social)	75.9%	Paper diary
22. Gong et al. (2016)	6,600 taxis; one-week survey; 06/2019 in Shanghai (China) In-vehicle devices every 10s	Signal of whether passenger on-board	- GI: POI data, opening time - Activity: drop-off time	PB	Bayes' rules based on spatial and temporal constraints 9 purposes (in-home, work-related, transfer, dining, shopping, recreation, schooling, lodging, medical)	NA; use spatio-temporal pattern; trip length; trip direction	NA, use common sense
23. Usyukov (2017)	108 cyclists, 541 survey days (5 days per cyclist), 2011 in Waterloo (Canada) Low-cost personal devices every 5s	RB	- GI: land use data - Trip: start time, end time	RB & PB	Heuristic rule for home activities Discrete choice for non-home activities 3 purposes (home, work, other)	NA, proportion of activities	NA, use Transportation Tomorrow Survey (2006)

24. Wang et al. (2017)	188,363 trajectories with 13,455 pick-up and 17,926 drop-off points; in NYC (USA), 01-06/2015 In-vehicle devices	Signal of whether passenger on-board	- Trip: start time - Social network: Foursquare	PB	General probabilistic model 3 purposes (work-oriented, entertainment-oriented, nightlife-oriented)	NA, use spatio-temporal pattern	NA, use common sense
25. Chen et al. (2019)	110,000* of 13 million taxi trips; in Manhattan, NYC (USA) In-vehicle devices	Signal of whether passenger on-board	- Trip: time of day, day of week, duration - Social network: Foursquare	UML	Autoencoder and K-means 5 purposes (dining, recreation, work, homing, others)	NA, use spatio-temporal pattern	NA, use common sense
580	GI is "Geographic Information"; RB is "Rule-based"; PB is "Probability-based"; SML is "Supervised Machine Learning"; UML is "Unsupervised Machine Learning"; NA is "Not Available"; 100,000 for training Autoencoder and 10,000 for evaluating model.						
581							

582 **References**

- 583 Armoogum, J., Axhausen, K.W., Madre, J.-L., 2009. Chapter 36 Lessons from an Overview of National
584 Transport Surveys, from Working Group 3 of COST 355: “Changing Behavior Toward a More
585 Sustainable Transport System, in: Transport Survey Methods: Keeping up with a Changing
586 World. Emerald Group Publishing Limited, pp. 621–634.
- 587 Armoogum, J., Bonsall, P., Browne, M., Christensen, L., Cools, M., Cornelis, E., Diana, M., Guilloux, T.,
588 Harder, H., Hegner Reinau, K., Hubert, J.-P., Kagerbauer, M., Kuhnimhof, T., Madre, J.-L.,
589 Moiseeva, A., Polak, J., Schulz, A., Tébar, M., Vidalakis, L., 2014. Survey Harmonisation with New
590 Technologies Improvement (SHANTI). IFSTTAR.
- 591 Auld, J., Williams, C., Mohammadian, A. (Kouros), Nelson, P., 2009. An automated GPS-based
592 prompted recall survey with learning algorithms. *Transp. Lett. Int. J. Transp. Res.* 1, 59–79.
- 593 Bachu, P., Dudala, T., Kothuri, S., 2001. Prompted Recall in Global Positioning System Survey:
594 Proof-of-Concept Study. *Transp. Res. Rec. J. Transp. Res. Board* 1768, 106–113.
595 <https://doi.org/10.3141/1768-13>
- 596 Böcker, L., Dijst, M., Prillwitz, J., 2013. Impact of Everyday Weather on Individual Daily Travel
597 Behaviours in Perspective: A Literature Review. *Transp. Rev.* 33, 71–91.
598 <https://doi.org/10.1080/01441647.2012.747114>
- 599 Bohte, W., Maat, K., 2009. Deriving and validating trip purposes and travel modes for multi-day
600 GPS-based travel surveys: A large-scale application in the Netherlands. *Transp. Res. Part C*
601 *Emerg. Technol.* 17, 285–297. <https://doi.org/10.1016/j.trc.2008.11.004>
- 602 Bricka, S., Bhat, C., 2006. Comparative Analysis of Global Positioning System-Based and Travel
603 Survey-Based Data. *Transp. Res. Rec. J. Transp. Res. Board* 1972, 9–20.
604 <https://doi.org/10.3141/1972-04>

605 Bricka, S.G., Sen, S., Paleti, R., Bhat, C.R., 2012. An analysis of the factors influencing differences in
606 survey-reported and GPS-recorded trips. *Transp. Res. Part C Emerg. Technol.* 21, 67–88.
607 <https://doi.org/10.1016/j.trc.2011.09.005>

608 Burbidge, S., Goulias, K., 2009. Active travel behavior. *Transp. Lett.* 1, 147–167.
609 <https://doi.org/10.3328/TL.2009.01.02.147-167>

610 Castree, N., Kitchin, R., Rogers, A., 2013. *A Dictionary of Human Geography*, 1st ed. Oxford University
611 Press. <https://doi.org/10.1093/acref/9780199599868.001.0001>

612 Chen, C., Bian, L., Ma, J., 2014. From traces to trajectories: How well can we guess activity locations
613 from mobile phone traces? *Transp. Res. Part C Emerg. Technol.* 46, 326–337.
614 <https://doi.org/10.1016/j.trc.2014.07.001>

615 Chen, C., Gong, H., Lawson, C., Bialostozky, E., 2010. Evaluating the feasibility of a passive travel
616 survey collection in a complex urban environment: Lessons learned from the New York City case
617 study. *Transp. Res. Part Policy Pract.* 44, 830–840. <https://doi.org/10.1016/j.tra.2010.08.004>

618 Chen, C., Liao, C., Xie, X., Wang, Y., Zhao, J., 2019. Trip2Vec: a deep embedding approach for
619 clustering and profiling taxi trip purposes. *Pers. Ubiquitous Comput.* 23, 53–66.
620 <https://doi.org/10.1007/s00779-018-1175-9>

621 Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M., 2016. The promises of big data and small data for travel
622 behavior (aka human mobility) analysis. *Transp. Res. Part C Emerg. Technol.* 68, 285–299.
623 <https://doi.org/10.1016/j.trc.2016.04.005>

624 Cools, M., Moons, E., Creemers, L., Wets, G., 2010. Changes in Travel Behavior in Response to
625 Weather Conditions: Do Type of Weather and Trip Purpose Matter? *Transp. Res. Rec. J. Transp.*
626 *Res. Board* 2157, 22–28. <https://doi.org/10.3141/2157-03>

627 Cui, Y., Meng, C., He, Q., Gao, J., 2018. Forecasting current and next trip purpose with social media
628 data and Google Places. *Transp. Res. Part C Emerg. Technol.* 97, 159–174.

629 <https://doi.org/10.1016/j.trc.2018.10.017>

630 Deng, H., Runger, G., Tuv, E., 2011. Bias of Importance Measures for Multi-valued Attributes and
631 Solutions, in: Honkela, T., Duch, W., Girolami, M., Kaski, S. (Eds.), *Artificial Neural Networks and*
632 *Machine Learning – ICANN 2011*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 293–300.
633 https://doi.org/10.1007/978-3-642-21738-8_38

634 Deng, Z., Ji, M., 2010. Deriving Rules for Trip Purpose Identification from GPS Travel Survey Data and
635 Land Use Data: A Machine Learning Approach, in: *Traffic and Transportation Studies 2010*.
636 Presented at the Seventh International Conference on Traffic and Transportation Studies (ICTTS)
637 2010, American Society of Civil Engineers, Kunming, China, pp. 768–777.
638 [https://doi.org/10.1061/41123\(383\)73](https://doi.org/10.1061/41123(383)73)

639 Du, J., Aultman-Hall, L., 2007. Increasing the accuracy of trip rate information from passive multi-day
640 GPS travel datasets: Automatic trip end identification issues. *Transp. Res. Part Policy Pract.* 41,
641 220–232. <https://doi.org/10.1016/j.tra.2006.05.001>

642 Ermagun, A., Fan, Y., Wolfson, J., Adomavicius, G., Das, K., 2017. Real-time trip purpose prediction
643 using online location-based search and discovery services. *Transp. Res. Part C Emerg. Technol.*
644 77, 96–112. <https://doi.org/10.1016/j.trc.2017.01.020>

645 Fang, C., Liu, H., Luo, K., Yu, X., 2017. Process and proposal for comprehensive regionalization of
646 Chinese human geography. *J. Geogr. Sci.* 27, 1155–1168.
647 <https://doi.org/10.1007/s11442-017-1428-y>

648 Feng, T., Timmermans, H.J.P., 2015. Detecting activity type from gps traces using spatial and
649 temporal information. *Eur. J. Transp. Infrastruct. Res.* 15, 662–674.

650 Feng, T., Timmermans, H.J.P., 2013. Transportation mode recognition using GPS and accelerometer
651 data. *Transp. Res. Part C Emerg. Technol.* 37, 118–130.
652 <https://doi.org/10.1016/j.trc.2013.09.014>

653 Forrest, T., Pearson, D., 2005. Comparison of Trip Determination Methods in Household Travel
654 Surveys Enhanced by a Global Positioning System. *Transp. Res. Rec. J. Transp. Res. Board* 1917,
655 63–71. <https://doi.org/10.3141/1917-08>

656 Furletti, B., Cintia, P., Renso, C., Spinsanti, L., 2013. Inferring Human Activities from GPS Tracks, in:
657 Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing, UrbComp
658 '13. ACM, New York, NY, USA, pp. 5:1–5:8. <https://doi.org/10.1145/2505821.2505830>

659 Gautama, S., Atzmueller, M., Kostakos, V., Gillis, D., Hosio, S., 2017. Observing Human Activity
660 Through Sensing, in: Loreto, V., Haklay, M., Hotho, A., Servedio, V.D.P., Stumme, G., Theunis, J.,
661 Tria, F. (Eds.), *Participatory Sensing, Opinions and Collective Awareness*. Springer International
662 Publishing, Cham, pp. 47–68. https://doi.org/10.1007/978-3-319-25658-0_3

663 Gibson, C., 2009. Human Geography, in: *International Encyclopedia of Human Geography*. Elsevier,
664 pp. 218–231. <https://doi.org/10.1016/B978-008044910-4.00275-3>

665 Gong, H., Chen, C., Bialostozky, E., Lawson, C.T., 2012. A GPS/GIS method for travel mode detection
666 in New York City. *Comput. Environ. Urban Syst., Special Issue: Geoinformatics 2010* 36,
667 131–139. <https://doi.org/10.1016/j.compenurbsys.2011.05.003>

668 Gong, L., Kanamori, R., Yamamoto, T., 2018. Data selection in machine learning for identifying trip
669 purposes and travel modes from longitudinal GPS data collection lasting for seasons. *Travel*
670 *Behav. Soc.* <https://doi.org/10.1016/j.tbs.2017.03.004>

671 Gong, L., Liu, X., Wu, L., Liu, Y., 2016. Inferring trip purposes and uncovering travel patterns from taxi
672 trajectory data. *Cartogr. Geogr. Inf. Sci.* 43, 103–114.
673 <https://doi.org/10.1080/15230406.2015.1014424>

674 Gong, L., Morikawa, T., Yamamoto, T., Sato, H., 2014. Deriving Personal Trip Data from GPS Data: A
675 Literature Review on the Existing Methodologies. *Procedia - Soc. Behav. Sci., The 9th*
676 *International Conference on Traffic and Transportation Studies (ICTTS 2014)* 138, 557–565.

677 <https://doi.org/10.1016/j.sbspro.2014.07.239>

678 Hall, R. (Ed.), 2003. Handbook of Transportation Science, 2nd ed. 2003 edition. ed. Springer, Boston.

679 Ho, C.Q., Mulley, C., 2013. Multiple purposes at single destination: A key to a better understanding of
680 the relationship between tour complexity and mode choice. *Transp. Res. Part Policy Pract.* 49,
681 206–219. <https://doi.org/10.1016/j.tra.2013.01.040>

682 Krause, C.M., Zhang, L., 2019. Short-term travel behavior prediction with GPS, land use, and point of
683 interest data. *Transp. Res. Part B Methodol.* 123, 349–361.
684 <https://doi.org/10.1016/j.trb.2018.06.012>

685 Lee, R.J., Sener, I.N., III, J.A.M., 2016. An evaluation of emerging data collection technologies for
686 travel demand modeling: from research to practice. *Transp. Lett.* 8, 181–193.
687 <https://doi.org/10.1080/19427867.2015.1106787>

688 Liao, L., Fox, D., Kautz, H., 2007. Extracting Places and Activities from GPS Traces Using Hierarchical
689 Conditional Random Fields. *Int. J. Robot. Res.* 26, 119–134.
690 <https://doi.org/10.1177/0278364907073775>

691 Lin, M., Hsu, W.-J., 2014. Mining GPS data for mobility patterns: A survey. *Pervasive Mob. Comput.*
692 12, 1–16. <https://doi.org/10.1016/j.pmcj.2013.06.005>

693 Liu, C., Susilo, Y.O., Karlström, A., 2017. Weather variability and travel behaviour – what we know
694 and what we do not know. *Transp. Rev.* 37, 715–741.
695 <https://doi.org/10.1080/01441647.2017.1293188>

696 Lu, Y., Liu, Y., 2012. Pervasive location acquisition technologies: Opportunities and challenges for
697 geospatial studies. *Comput. Environ. Urban Syst., Special Issue: Geoinformatics 2010* 36,
698 105–108. <https://doi.org/10.1016/j.compenvurbsys.2012.02.002>

699 Lu, Y., Zhu, S., Zhang, L., 2012. A Machine Learning Approach to Trip Purpose Imputation in
700 GPS-Based Travel Surveys. Presented at the 4th Conference on Innovations in Travel Modeling,

701 Florida, United States.

702 Marra, A.D., Becker, H., Axhausen, K.W., Corman, F., 2019. Developing a passive GPS tracking system
703 to study long-term travel behavior. *Transp. Res. Part C Emerg. Technol.* 104, 348–368.
704 <https://doi.org/10.1016/j.trc.2019.05.006>

705 McGowen, P.T., McNally, M.G., 2007. Evaluating the Potential To Predict Activity Types from GPS and
706 GIS Data. Presented at the Transportation Research Board 86th Annual Meeting Transportation
707 Research Board.

708 Meng, C., Cui, Y., He, Q., Su, L., Gao, J., 2017. Travel purpose inference with GPS trajectories, POIs,
709 and geo-tagged social media data, in: 2017 IEEE International Conference on Big Data (Big Data).
710 Presented at the 2017 IEEE International Conference on Big Data (Big Data), pp. 1319–1324.
711 <https://doi.org/10.1109/BigData.2017.8258062>

712 Montini, L., Rieser-Schüssler, N., Horni, A., Axhausen, K., 2014. Trip Purpose Identification from GPS
713 Tracks. *Transp. Res. Rec. J. Transp. Res. Board* 2405, 16–23. <https://doi.org/10.3141/2405-03>

714 Nguyen, M.H., Armoogum, J., Garcia, C., 2019a. Mode-Based Comparison of Data in Mobility Surveys
715 using GPS and Telephone. Presented at the 98th TRB Annual Meeting, Washington, D.C.

716 Nguyen, M.H., Ha, T.T., Tu, S.S., Nguyen, T.C., 2019b. Impediments to the bus rapid transit
717 implementation in developing countries – a typical evidence from Hanoi. *Int. J. Urban Sci.* 4,
718 464–483.

719 Nguyen, M.H., Pojani, D., 2018. Chapter Two - Why Do Some BRT Systems in the Global South Fail to
720 Perform or Expand?, in: Shiftan, Y., Kamargianni, M. (Eds.), *Preparing for the New Era of*
721 *Transport Policies: Learning from Experience, Advances in Transport Policy and Planning.*
722 Academic Press, pp. 35–61. <https://doi.org/10.1016/bs.atpp.2018.07.005>

723 Oliveira, M.G.S., Vovsha, P., Wolf, J., Mitchell, M., 2014. Evaluation of Two Methods for Identifying
724 Trip Purpose in GPS-Based Household Travel Surveys. *Transp. Res. Rec. J. Transp. Res. Board*

725 2405, 33–41. <https://doi.org/10.3141/2405-05>

726 Ortúzar, J.D.D., Armoogum, J., Madre, J., Potier, F., 2011. Continuous Mobility Surveys: The State of
727 Practice. *Transp. Rev.* 31, 293–312. <https://doi.org/10.1080/01441647.2010.510224>

728 Patrick, B., Martin, L.-G., Zmud, J., Madre, J., 2009. *Transport Survey Methods: Keeping Up with a*
729 *Changing World*. Emerald Group Publishing Limited.

730 Pojani, D., 2020. *Planning for Sustainable Urban Transport in Southeast Asia Policy Transfer,*
731 *Diffusion, and Mobility*. Springer, Cham.

732 Prelipcean, A.C., Gidófalvi, G., Susilo, Y.O., 2017. Transportation mode detection – an in-depth review
733 of applicability and reliability. *Transp. Rev.* 37, 442–464.
734 <https://doi.org/10.1080/01441647.2016.1246489>

735 Reumers, S., Liu, F., Janssens, D., Cools, M., Wets, G., 2013. Semantic Annotation of Global
736 Positioning System Traces: Activity Type Inference. *Transp. Res. Rec. J. Transp. Res. Board* 2383,
737 35–43. <https://doi.org/10.3141/2383-05>

738 Richardson, A.J., Ampt, E.S., Meyburg, A.H., 1996. Nonresponse Issues in Household Travel Surveys,
739 in: *Transportation Research Board Conference Proceedings. Presented at the Conference on*
740 *Household Travel Surveys: New Concepts and Research Needs Transportation Research Board;*
741 *Federal Highway Administration; Federal Transit Administration; and Bureau of Transportation*
742 *Statistics*.

743 Schuessler, N., Axhausen, K., 2009. Processing Raw Data from Global Positioning Systems Without
744 Additional Information. *Transp. Res. Rec. J. Transp. Res. Board* 2105, 28–36.
745 <https://doi.org/10.3141/2105-04>

746 Semanjski, I., Gautama, S., Ahas, R., Witlox, F., 2017. Spatial context mining approach for transport
747 mode recognition from mobile sensed big data. *Comput. Environ. Urban Syst.* 66, 38–52.
748 <https://doi.org/10.1016/j.compenvurbsys.2017.07.004>

749 Shen, L., Stopher, P.R., 2014. Review of GPS Travel Survey and GPS Data-Processing Methods. *Transp.*
750 *Rev.* 34, 316–334. <https://doi.org/10.1080/01441647.2014.903530>

751 Shen, L., Stopher, P.R., 2013. A process for trip purpose imputation from Global Positioning System
752 data. *Transp. Res. Part C Emerg. Technol.* 36, 261–267.
753 <https://doi.org/10.1016/j.trc.2013.09.004>

754 Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification
755 tasks. *Inf. Process. Manag.* 45, 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>

756 Stopher, P., FitzGerald, C., Zhang, J., 2008. Search for a global positioning system device to measure
757 person travel. *Transp. Res. Part C Emerg. Technol., Emerging Commercial Technologies* 16,
758 350–369. <https://doi.org/10.1016/j.trc.2007.10.002>

759 Stopher, P., Shen, L., 2011. In-Depth Comparison of Global Positioning System and Diary Records.
760 *Transp. Res. Rec. J. Transp. Res. Board* 2246, 32–37. <https://doi.org/10.3141/2246-05>

761 Stopher, P.R., Daigler, V., Griffith, S., 2018. Smartphone app versus GPS Logger: A comparative study.
762 *Transp. Res. Procedia, Transport Survey Methods in the era of big data: facing the challenges* 32,
763 135–145. <https://doi.org/10.1016/j.trpro.2018.10.026>

764 Thomas, T., Geurs, K.T., Koolwaaij, J., Bijlsma, M., 2018. Automatic Trip Detection with the Dutch
765 Mobile Mobility Panel: Towards Reliable Multiple-Week Trip Registration for Large Samples. *J.*
766 *Urban Technol.* 25, 143–161. <https://doi.org/10.1080/10630732.2018.1471874>

767 Usyukov, V., 2017. Methodology for identifying activities from GPS data streams. *Procedia Comput.*
768 *Sci., 8th International Conference on Ambient Systems, Networks and Technologies, ANT-2017*
769 *and the 7th International Conference on Sustainable Energy Information Technology, SEIT 2017,*
770 *16-19 May 2017, Madeira, Portugal* 109, 10–17. <https://doi.org/10.1016/j.procs.2017.05.289>

771 Wang, P., Liu, G., Fu, Y., Zhou, Y., Li, J., 2017. Spotting Trip Purposes from Taxi Trajectories: A General
772 Probabilistic Model. *ACM Trans. Intell. Syst. Technol. TIST* 9, 29:1–29:26.

773 <https://doi.org/10.1145/3078849>

774 Wolf, J., Bachman, W., Oliveira, M.S., Auld, J., Mohammadian, A. (Kouros), Vovsha, P., 2014. Applying
775 GPS Data to Understand Travel Behavior, Volume I: Background, Methods, and Tests.
776 Transportation Research Board, Washington, D.C. <https://doi.org/10.17226/22370>

777 Wolf, J., Guensler, R., Bachman, W., 2001. Elimination of the Travel Diary: Experiment to Derive Trip
778 Purpose from Global Positioning System Travel Data. *Transp. Res. Rec. J. Transp. Res. Board*
779 1768, 125–134. <https://doi.org/10.3141/1768-15>

780 Wolf, J., Oliveira, M., Thompson, M., 2003. Impact of Underreporting on Mileage and Travel Time
781 Estimates: Results from Global Positioning System-Enhanced Household Travel Survey. *Transp.*
782 *Res. Rec. J. Transp. Res. Board* 1854, 189–198. <https://doi.org/10.3141/1854-21>

783 Wolf, J., Schönfelder, S., Samaga, U., Oliveira, M., Axhausen, K., 2004. Eighty Weeks of Global
784 Positioning System Traces: Approaches to Enriching Trip Information. *Transp. Res. Rec. J.*
785 *Transp. Res. Board* 1870, 46–54. <https://doi.org/10.3141/1870-06>

786 Xiao, G., Juan, Z., Zhang, C., 2016. Detecting trip purposes from smartphone-based travel surveys
787 with artificial neural networks and particle swarm optimization. *Transp. Res. Part C Emerg.*
788 *Technol.* 71, 447–463. <https://doi.org/10.1016/j.trc.2016.08.008>

789 Yazdizadeh, A., Patterson, Z., Farooq, B., 2019. An automated approach from GPS traces to complete
790 trip information. *Int. J. Transp. Sci. Technol.* 8, 82–100.
791 <https://doi.org/10.1016/j.ijst.2018.08.003>

792 Zhao, F., Pereira, F.C., Ball, R., Kim, Y., Han, Y., Zegras, C., Ben-Akiva, M., 2015. Exploratory Analysis of
793 a Smartphone-Based Travel Survey in Singapore. *Transp. Res. Rec. J. Transp. Res. Board* 2494,
794 45–56. <https://doi.org/10.3141/2494-06>

795
796

797 **Author biographies and photos**



798

799 **Minh Hieu Nguyen** is a lecturer at Faculty of Transport Economics under University of Transport and
800 Communications (Vietnam). His main research focus is on deriving trip characteristics from GPS data, analyzing
801 travel behavior, and developing public transport in emerging countries.

802

803 **Jimmy Armoogum**, is a senior researcher at the Laboratory Economic and Social Dynamics of Transport (DEST)
804 of University Gustave Eiffel. He holds a PhD in Mathematical Statistic from the University of Brussels. His
805 research focuses on travel survey methodologies and its analysis, modeling including the long-term projection
806 of motorization and mobility. He has already completed several research projects for European and French
807 framework research programs.

808

809 **Jean-Loup Madre**, is the former Head of the Laboratory of Transport Economics and Sociology, PhD in
810 Mathematical Statistics and in Economics, Long experiences in Survey Methods (e.g. French NTS 1993-94). He
811 took part in European project such as MEST and TEST, Motorisation and Mobility in Metropolitan areas;
812 Longitudinal Approaches in Travel Behaviour (including long term projection), International comparisons
813 (SCENARIOS, KITE), SCENE.

814

815 **Cédric Garcia** is a research engineer in computational science and statistics, at the Laboratory of Economic and
816 Social Dynamics of Transport (DEST) of Université Gustave Eiffel. He holds a master degree in general
817 mathematics and a master degree in statistics. From 2013 to 2015, he worked as a mathematics teacher in high
818 school. From 2016 to 2017, he worked at Université Gustave Eiffel in an epidemiology research unit, where he

819 studied risk factors and did responsibility analyses and predictions on police report road crashes. Since 2018,
820 his work focuses on travel survey methodologies and its analysis, modeling immobility. Cédric is also a
821 participant in the Eurostat task force on the Harmonised travel survey guidelines.