



REVIEW ARTICLE



<https://doi.org/10.1057/s41599-020-0455-y>

OPEN

Estimating temporary populations: a systematic review of the empirical literature

Radoslaw Panczak ¹✉, Elin Charles-Edwards¹ & Jonathan Corcoran¹

ABSTRACT

The estimation of temporary populations is a well-established field, but despite growing interest they are yet to form part of the standard suite of official population statistics. This systematic review seeks to review the empirical literature on temporary population estimation and identify the contemporary “state of the art”. We identify a total of 96 studies that attempt to estimate or describe a method of estimation. Our findings reveal strong growth in the number of studies in recent decades that in part has been driven by the rise in both the type and availability of new sources of information, including mobile phone data. What emerges from this systematic review is the lack of any “gold standard” data source or methodology for temporary population estimation. The review points to a number of important challenges that remain for estimating temporary populations, both conceptually and practically. What remains is the need for clear definitions along with identification of appropriate data and methods that are able to robustly capture and measure the diverse array of spatial behaviours that drive temporary population dynamics. To our knowledge, this is the first review on this topic that brings together literature from various disciplines and collates methods used for estimation.

Background

Conventional population estimates capture a population at a single point in time. In doing so, such estimates ignore the short-term dynamism of populations that are caused by temporary population mobility, a territorial movement that does not result in a permanent change of usual residence (Bell and Ward, 2000). Temporary population movements are defined as moves of one or more night’s duration but can be broadened to also include diurnal movements, such as daily commuting (Smith, 1989).¹ Temporary population mobility can exert substantive impacts on both the size and composition of populations at the small-area level (Charles-Edwards and Panczak, 2018) impacting traffic, housing, retail sales, medical services and emergency preparedness to name a select few (Smith, 1989; Smith and House, 2007). As a consequence of these impacts, there exists a growing interest and need for temporary population estimates across a broad range of purposes including the planning and delivery of goods and services (Markham et al., 2013), fiscal equalisation (Graebert et al., 2014), retail analysis (Soundararaj et al., 2019), transport demand (Toole et al., 2015), emergency preparedness (Gao

¹Queensland Centre for Population Research, School of Earth and Environmental Sciences, The University of Queensland, Brisbane, QLD, Australia.
✉email: r.panczak@gmail.com

et al., 2014), and as better denominators for crime and epidemiological models (Charles-Edwards et al., 2020)—for a full range of purposes please refer to Supplementary Table 1, Part 1, ‘Source, type, purpose main data and methods’.

Scholarship concerned with developing a conceptual basis for temporary population estimates is longstanding with early work first appearing in the United States in the 1950s (Foley, 1954; Schmitt, 1956). This early work proposed building on approaches previously used for the estimation of resident populations to generate temporary population estimates. Schmitt (1956) was the first to recognise the value of symptomatic data as a source for such estimates. Several decades later, it was Smith (1987, 1989) that proposed a set of metrics (visitor-days and visitor-years) to capture measures of temporary populations as well as highlighting the difficulty of validating temporary population estimates. More recently, Bell (2000, 2004) argued that estimates of temporary populations are contingent on developing a rigorous understanding of temporary population mobility. Here he drew attention to four dimensions of temporary mobility (movement intensity, duration, seasonality, and spatial impact), which he argued drive change in the stock of temporary visitors at both the origin and at the destination.

Despite growing interest in temporary population statistics, they are not yet part of the standard suite of official population statistics. Historically, this has been due to a lack of the requisite data capturing temporary populations, but also nascent conceptualisation of temporary populations along with a lack of standard methodologies for their production. Indeed, Smith (1989), in his seminal paper on the estimation of temporary populations, noted that the development of a methodology applicable across a wide range of populations and situational contexts was ‘most likely impossible’ (p. 430). At the time of writing he could not have predicted the proliferation of geolocated data from mobile phones and other emergent technologies that have arisen in the 21st Century with potential for the sensing of temporary populations across large geographic areas and in widely differing contexts. Set against this backdrop, the current study assesses progress in the field of temporary population statistics. To this end, we conduct a systematic review of the international literature on the estimation of temporary populations to identify the contemporary state of the art. In doing so, we seek to understand the volume, geography, population coverage along with the data and methods used to generate temporary population estimates.

The remainder of the paper is structured as follows. In the section ‘Methods: Protocol and capture’, we detail the study methods including information sources and search terms, eligibility criteria and data extraction. In the section ‘Results and discussion: Estimating temporary populations’, we report the results of the review across a number of themes including: publication time frame, sources and reasons for estimates; data types; estimation methods; the geography of estimates; the temporal units of estimates; estimated population size and composition; and reproducibility. In the final section of the paper we summarise the current state of the art in the field.

Methods: Protocol and capture

The review was conducted using the principles of the PRISMA Statement for Reporting Systematic Reviews (Moher et al., 2009) (see Supplementary Table 1 online). In order to deal with complexity and lack of standardisation in the field and to account for the theory-building character of this work we additionally adopted the ‘mixed-method review’ methodology (Gough, 2015). This approach allowed us to combine methods of a systematic review (e.g. selection of keywords and sources and eligibility

criteria) with the openness and flexibility of traditional reviews. No review protocol was used in this study.

The articles that constituted the review were captured via a search of four databases; (1) Web of Science; (2) Scopus; (3) Google Scholar; and, (4) arXiv. A set of keywords that collectively encompass the breadth of synonyms associated with temporary populations were used. More specifically, we employed a suite of terms derived from the authors’ collective knowledge of the field, as well as key words that emerged during early stages of the review process. Terms included: ‘temporary population’, ‘working population’, ‘daytime population’, ‘mobile population’, ‘service population’, ‘floating population’, ‘elusive population’, ‘ambient population’, ‘seasonal population’, ‘non-resident population’, ‘real-time census’, ‘social sensing’ and ‘spatiotemporal population’. We additionally scrutinised the reference lists of included studies to ensure a full coverage of articles. Final searches were conducted on 24 November 2017. In order to enhance the currency and coverage of the articles captured principles of the living systematic review methodology were adopted (Elliott et al., 2014). To this end updates from three databases covering all original queries were configured along with periodic manual searches on the fourth database (arXiv) to ensure any new articles following the final search were included in the review. The final inclusion of studies into the review was conducted on 13 October, 2019. Supplementary Text 1 online contains details of the full search strategy.

To supplement the key search terms listed above, inclusion and exclusion criteria were developed. These were;

- Studies were included if they produced estimates or forecasts of temporary populations.
- Studies were required to detail their methodology for at least one geographical region and capture non-resident populations.
- Studies were excluded if they focused their analyses on certain population bases (for instance, social media or mobile phone users) without any extrapolation to the total population of a given case study region.
- Studies examining tourist and crowd behaviours were excluded as this is a distinct and voluminous literature deemed extraneous to the core motivation of this study.
- Book reviews and magazines were excluded.
- Only studies published in English were included.

Each of the above criteria were designed to capture studies that offered an in-depth empirical analysis of temporary population estimation. Beyond the above criteria, there were no restrictions on publication date, or publication status (for example, a published journal article versus an unpublished thesis). In cases where the same or a similar methodology was used across several studies, one representative article was selected for inclusion in the review and used to collect additional references for the supplementary materials.

We developed a data extraction template that was iteratively revised during the extraction process to ensure applicability across all included studies. One author (RP) extracted data from each study that was subsequently split into two major groups:

1. *Publication details*: Publication metadata included: author (s), title, source, year and publication type. We classified studies according to the primary source of data: mobile phone, remote sensing, social media, transport (e.g. cordon counts, traffic and smart card data), official statistics (e.g. national censuses, large scale or national demographic surveys and population registers), surveys (e.g. time use surveys, holiday accommodation occupancy surveys), Wi-Fi and data from utility providers (e.g. electricity consumption). Studies could belong to more than one category.

2. *Estimation*: Characteristics of the temporary population estimates reported in the study including whether estimates captured daytime or overnight populations, the country and region for which estimates were made, data sources, methods and software employed, data and population size, number and characteristics of the spatial and temporal units, estimate validation and the purpose for which the estimates were produced.

All analyses were conducted in R (R Core Team, 2019) version 3.6.1. Data management used several packages of the tidyverse family (Wickham, 2019a, b; Wickham and Bryan, 2019; Wickham et al., 2019; Wickham and Henry, 2019); descriptive statistics were done using sjmisc package (Luedecke, 2020); reporting was facilitated with knitr (Xie, 2019) accompanied by scales (Wickham, 2018) and bibtext (Francois, 2017); and graphs were produced using the ggplot2 package (Wickham, 2016).

Results and discussion: Estimating temporary populations
Publications time frame, sources and reasons for estimates.

Collectively, our search strategy initially identified around 22,900 documents (Fig. 1). After removing duplicates, titles and abstracts were screened for eligibility resulting in a reduction to 122 articles. The final stage involved a screening of full texts resulting in the removal of a further 68 articles. An additional 26 articles were identified from other sources, while 16 were identified by scanning updates of the search results. A total of 96 studies were selected for analysis (see Supplementary Text 2 online).

The 96 studies were published between 1925 and 2019 (median publication year, 2011) (Table 1; see Supplementary Table 2

online). Notably, there has been exponential growth in the number of studies since the first publication in 1925, with more than a third of the studies (35.4%) published since 2015. A total of 55% of the studies were published as articles followed by 22% as reports and 12% as conference proceedings or presentations. Taken together articles focused on various aspects of temporary population estimation, including emergency planning ($n = 22$), understanding service populations ($n = 8$), epidemiology ($n = 4$) and commuting ($n = 3$).

Data types. Official statistics, defined as large scale data collected by national statistical agencies (such as censuses, population surveys, administrative records or registers), formed the basis, or were an important component of the largest number of studies (Fig. 2). The first recorded report used data on place of work collected at the 1921 Census data from England and Wales to estimate daytime populations (Census of England and Wales 1921, 1925). Similar work was undertaken by the US Census Bureau in the 1950s (Bureau of the Census 1956), and such estimates continue to be produced in many countries around the world. A more recent innovation are workday population estimates, which combine estimates of working populations with night-time usual resident counts. These estimates were produced for England and Wales using data from the 2011 Census (Office for National Statistics 2013); see also related work in Supplementary Table 3 online]. In a similar vein, the US Census Bureau has constructed commuter-adjusted population estimates using data from the American Community Survey (McKenzie et al., 2013).

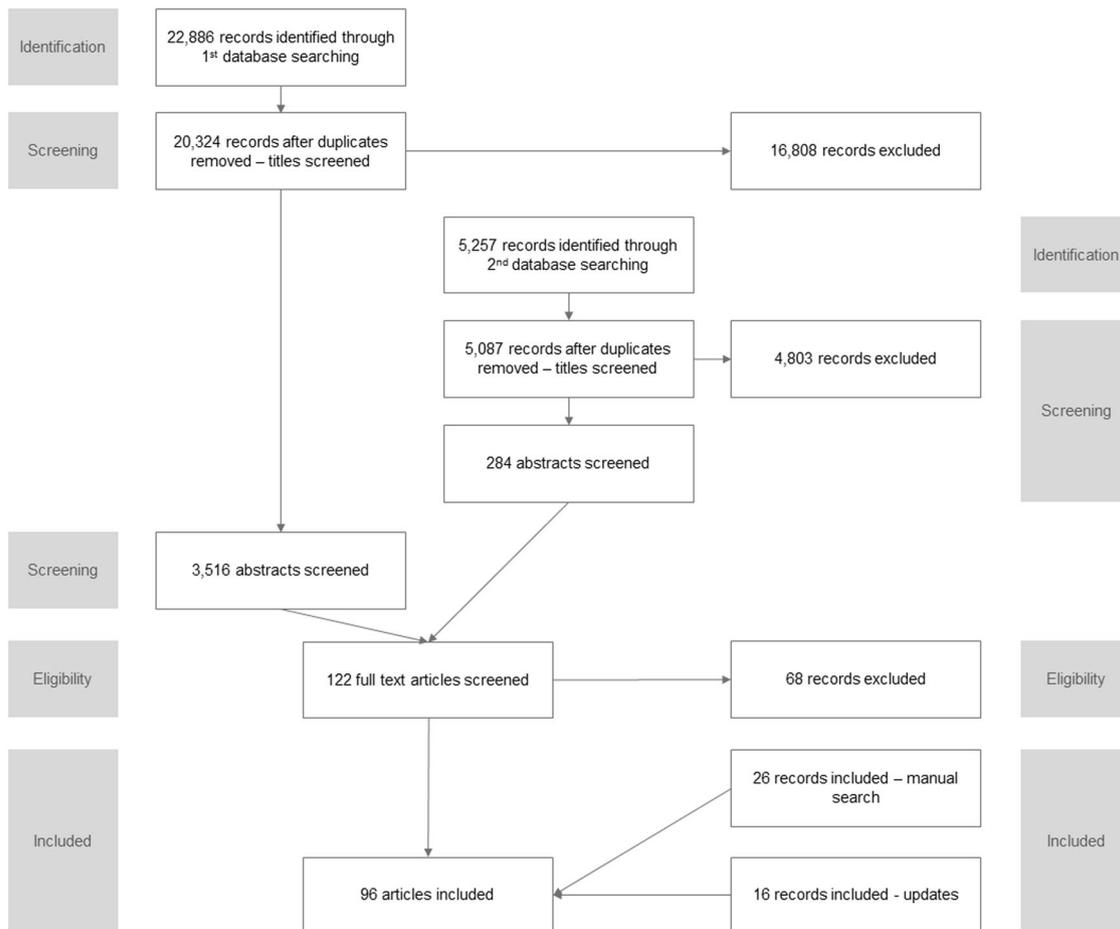


Fig. 1 Flowchart of study selection.

Table 1 Selected characteristics of included studies.

Characteristic	Value	n/median	%/min; max	
Year of publication		2011	1925; 2019	
Search method	Systematic	54	56.25	
	Manual	26	27.08	
	Update	16	16.67	
Type of publication	Article	53	55.21	
	Report	21	21.88	
	Conference	12	12.5	
	Thesis	7	7.29	
	Book chapter	3	3.12	
Purpose of publication ^a	Emergency planning	22	22.92	
	Other	9	9.38	
	Service population	8	8.34	
	Epidemiology	4	4.16	
	Commuting	3	3.12	
Main type of data ^a	Official statistics	53	55.20	
	Survey	40	41.68	
	Mobile phones	19	19.80	
	Transport	13	13.56	
	Other	11	11.44	
	Utilities	5	5.20	
	Remote sensing	2	2.08	
	Wi-Fi	2	2.08	
	Social media	1	1.04	
	Study region	Single city	52	54.17
	Country	17	17.71	
Administrative region	14	14.58		
Regions for estimates ^a	Multiple cities	10	10.42	
	Administrative	55	57.30	
	Grid	28	29.16	
	Custom	8	8.34	
	Points	7	7.30	
	Building	3	3.12	
Number of regions for estimates	Voronoi	2	2.08	
		34	1; 52,000	
Temporal unit of estimates	Minute/hour	35	36.46	
	Daytime/nighttime	21	21.88	
	Daytime	14	14.58	
	Month	10	10.42	
	Other	7	7.29	
	Day	2	2.08	
Size of estimated population	Weekday/weekend	2	2.08	
		355,100	1868; 53,349,074	
Size of dataset		722,000	422; 55,963,096	

^aIndicates that the study can belong to more than one category.

Census data have served as an input into a host of other population estimates. Census-based studies from the United States include the work from the Seattle City Planning Commission (1951), Fulton (1984), Gober and Mings (1984), Nelson and Nicholas (1992), McPherson and Brown (2003), Kobayashi et al. (2011), Swanson and Tayman (2011), Hodur and Bangsund (2015), Kim and Ahn (2017), Boeing (2018) and Esri (2018). These papers used a range of methods to adjust Census estimates at a variety of temporal and spatial scales, often accompanied by other data sources including place of work, transportation models or payroll data that helps to determine size and location of daytime or seasonal populations. Census data from the United Kingdom have formed a key input into the recent work of Smith et al. (2005) and Martin et al., (2015).

Census data has also been used for the estimation and analysis of temporary populations in a number of other countries. Taking advantage of the availability of the place of residence and place of enumeration data, Bell and Ward (1998); (see also subsequent work in Supplementary Table 3 online) looked at the one-day snapshot of the mobility of Australians across multiple census years and spatial scales. Similarly, Taylor (1998) used census data to estimate the mobility of indigenous Australians. Gao et al., (2014) and Qi et al., (2015) employed Chinese census data adjusted with information for tourists, school and patient populations extended with land use models to calculate daytime

Primary data types and year of publication

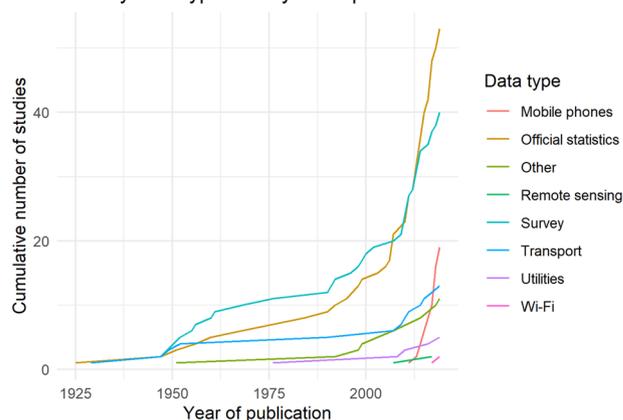


Fig. 2 Cumulative distribution of year of publication across data types. Studies can belong to more than one category; The ‘Social Media’ category has been excluded as it only has one member.

population estimates. The main advantage of census data is that it provides close to full enumeration of the population; however, it captures only a snapshot of temporary populations on one day or night, every 5 or 10 years. We know that many forms of temporary population mobility are highly seasonal, thus census data will not reflect the population present across other times of the year. To address this deficit a number of studies have drawn upon large administrative datasets such as health service utilisation records (Markham et al., 2013), building (Greger, 2015) or second home registries (Adamiak et al., 2017), however, these studies tend to target specific population groups or geographical regions, and may not directly capture the mobility of visitors. Notwithstanding these limitations, official statistics remain a crucial component of many temporary population estimates.

The second most common group of studies drew on survey data. Surveys vary widely with respect to the target population, spatial coverage and sample size and include urban travel surveys, destination surveys (Erickson, 1961; Happel and Hogan, 2002) and large-scale household surveys (Wurtele and Wellisch, 1968; Kavanaugh, 1990; Stutz et al., 1992; Roddis, 1996; Collins and Greaves, 2007; Lau, 2009; Horanont and Shibasaki, 2010; Charles-Edwards, 2011; Sekimoto et al., 2011; Himoto and Kimata, 2014; Kashiyama et al., 2017). Surveys have the singular advantage of being able to collect detailed information on the spatial and temporal dynamics of mobility, including the timing and duration of visits, as well as the characteristics of movers. They can be used to target transient populations such as seasonal snowbirds (Happel and Hogan, 2002) or indigenous Australians (Warchivker et al., 2000) not captured in official statistics. The major disadvantage of surveys are related to capturing rare events through random sampling. Even in large-scale origin-based surveys, sampling error can limit the utility of survey data for estimating temporary populations in small areas and over short time periods. Purposive surveys of accommodation providers or particular target populations fare better, but are usually limited with respect to both population and geographic coverage.

The third largest category of studies, and one that experienced the fastest growth in recent years, is work based on mobile phone data. Some of the early attempts to use this emergent source of data in tracking large-scale population distribution changes came from the SENSEable City Laboratory and their work in Milan (Ratti et al., 2006) and Rome (Reades et al., 2007). This was followed by work on seasonal residents in Estonia (Silm and Ahas, 2010). However these first studies did not attempt to

extrapolate the mobile phone user base to the whole population and the earliest such application was in response to the Haiti earthquake (Bengtsson et al., 2011). This approach was followed by large-scale estimates for Japan (Terada et al., 2013) and analyses of temporary populations in France and Portugal (Deville et al., 2014). The spatial coverage of mobile phone-based studies often reflects the availability and accessibility of such data. For instance, Italy and in particular Milan, have been the foci for a number of studies given that local telecom providers have offered access to the aggregated and anonymised call records (Khodabandelou et al., 2016). Mobile phone data are also used in parallel with other sources and methods for instance in dasymetric interpolation (Järv et al., 2017) or transport survey data (Lwin et al., 2016). While mobile phone data can provide detailed information on short-term change in population numbers at fine spatial resolutions they provide limited information on the characteristics of movers. Finally, as a commercially collected data set, they can be both difficult and costly to access.

Transportation data, collected via cordon counts, or more recently by smart cards, are an important source of information, particularly for diurnal population estimates. Early work from the United States can be traced back to Thornthwaite who estimated the daytime population of a Central Business District using data from a traffic survey (1929; cited in Wheeler and Brunn, 2002). Others followed including Breese (1947), Sharp (1955) and Weir (1960). The New York Regional Plan Association (1949), Seattle City Planning Commission (1951) and Institute for Research in Social Science at the University of North Carolina (1952) employed various transportation surveys and cordon counts to estimate the number of people in Central Business Districts. Similar work from the United Kingdom was undertaken by Menzler (1952). More recently and with the advent of public transport smart card data (Ma et al., 2017) along with technologies such as number plate recognition (Fehr & Peers, 2014) offer new data sources with utility for temporary population estimation. These studies have the advantage of longer collection periods, but tend to be limited to relatively narrowly defined geographic regions, usually cities, and tend to be employed for diurnal rather than seasonal population estimates (see Fehr & Peers, (2014) for an exception).

Less common, yet still important, are studies based on a range of symptomatic indicators including utilities data (Goldschmidt and Dahl, 1976; Monmouth County Planning Board, 2008; Rigall-I-Torrent, 2010; McKenzie and Canterford, 2016; Edmondson and Nantucket Data Platform team, 2019), remote sensing data (Taubenböck et al., 2007; Stathakis and Baltas, 2018) and Wi-Fi data (Kontokosta and Johnson, 2017; Crols and Malleson, 2019). At their most basic, these studies assume a constant relationship between a region's *de facto* population (including residents, diurnal and overnight visitors) and a single symptomatic indicator. Symptomatic variables theoretically enable the generation of continuous estimates of visitor populations in both a timely and cost effective manner, but are contingent on understanding the relationship between the symptomatic indicator and populations, which may vary across space and time. In addition, it can be difficult to identify appropriate symptomatic indicators when estimates are sought for a large number of regions. For example, a scoping study undertaken by the Australian Bureau of Statistics was unable to identify any symptomatic indicators that could be used to estimate visitor populations across five Local Government Areas in Australia (Lee, 1999). Contrary to our expectations, social media data were only used in one study (Lwin et al., 2016). Here geotagged Twitter data were used together with mobile phone records to predict scaled counts from a large person trip survey.

Slightly more than one-third of studies (38.5%) drew on multiple data sets. Census data was employed as a starting point,

for the calibration of models, or as a source of validation. For instance, the seminal work of LandScan—capturing the ambient population (an average population over a 24 hour period) at a 1 km resolution—(Bhaduri et al., 2007) used census data as one of the main inputs and Ma et al. (2017) used smart card transportation data to redistribute census (night-time) counts of population across a city region. Studies also draw on multiple data sources to estimate the components of temporary populations in particular regions (Edmondson and Nantucket Data Platform team, 2019), while Rigall-I-Torrent (2010) combined multiple symptomatic indicators with survey data to establish the relationship between household size, type and period of the year to produce population estimates. What emerges from this systematic review is the lack of any “gold standard” data source for temporary population estimation.

Estimation methods. Smith (1989) in his original study identified two approaches for the estimation of temporary populations: the ‘direct’ approach which draws on information collected directly from temporary residents via censuses and surveys and the ‘indirect’ which draws upon symptomatic variables that reflect changes in temporary populations. This distinction is increasingly fuzzy due to the emergence of new data sources such as mobile phone and Wi-Fi data which can be both treated as either direct or indirect sources depending on the methodology employed. In addition, a number of studies combine both direct and indirect sources to generate estimates for individual regions (Rigall-I-Torrent, 2010) or to apportion temporary visitors across geographic areas (Lwin et al., 2016). Moreover studies are increasingly drawing on sophisticated simulation and modelling approaches to produce temporary population estimates (see Crols and Malleson (2019) for an example).

The earliest method of deriving temporary population statistics was to use direct estimates, either from the Census (Census of England and Wales 1921, 1925), survey data (Thornthwaite, 1929) or by combining multiple formerly disparate sources of data (Breese, 1947). These studies often took advantage of availability of information about the workplaces of individuals (Office for National Statistics, 2013), their status of residence (Gober and Mings, 1984) or place of enumeration (Bell and Ward, 1998) to distinguish between two or more states of population distribution such as nighttime and daytime populations, or visitors and usual residents. An allied approach involves the scaling of population counts from one or more sources to the whole population based on expansion factors. This approach was used by both early survey-based studies (New York Regional Plan Association, 1949; Foley, 1952) and more contemporary analyses based on mobile phone or Wi-Fi data. In the latter case, various strategies of deriving factors were adopted, some based on the penetration rate of mobile data providers (Bengtsson et al., 2011), night-time Census populations (Deville et al., 2014) or ancillary survey data (Kontokosta and Johnson, 2017). Studies either used a single factor for an entire region (Thomas et al., 2017) or multiple factors, varying each geographically (Batra et al., 2018) or over population subgroups (Picornell et al., 2018).

Temporary populations have also been estimated using population accounts or component-based approaches. These methods used generic or area-specific equations to derive counts of temporary residents from usual residents by subtracting departures and adding arrivals to a spatial unit at a given time (Yong li, 1998). For instance Journey-to-Work data (from the Census) has been used to estimate the working population (McKenzie et al., 2013), second home users and usage could be used to capture seasonal populations in specific areas (Adamiak et al., 2017) or sales tax data could be used to derive equivalent

residents (Thakur, 2018). Additionally, separate estimates have been derived for specific subgroups of the population (Swanson and Tayman, 2011).

As discussed previously, symptomatic data such as electricity usage have been used to track seasonal variation in populations across specific regions (see Goldschmidt and Dahl, 1976). An allied strand of research draws on symptomatic data to redistribute temporary populations from larger to smaller spatial units. These include dasymetric techniques commonly used for mapping small area resident populations. The LandScan model (Bhaduri et al., 2007) for instance has the largest spatial coverage providing gridded population estimates of the nighttime and daytime population for the United States. Martin et al.’s (2015) model for Southampton on the other hand is an example of a more spatially and temporally detailed approach that can be adapted to various grid sizes and temporal resolutions. This model has also been implemented in other locations and settings (see Supplementary Table 3 online). The redistribution approach usually uses ancillary data such as land use (Batista e Silva et al., 2017) or building type (Greger, 2015) as well as data from Censuses, surveys, mobile phones (Järv et al., 2017), transportation (Ma et al., 2017) and social media (Lwin et al., 2016) as input.

Finally, a number of studies employ some form of modelling or simulation. These include agent-based simulation (Walker and Barros, 2012; Kashiyama et al., 2017; Crols and Malleson, 2019), cellular automaton models (Khakpour and Rød, 2016), and neural networks (Liu et al., 2018; Chen et al., 2018).

The geography of estimates. Only 18% of studies generated estimates for all geographic units within an entire country (Fig. 3). The United States was the best represented with a total of five studies, followed by Australia ($n = 3$) and England and Wales ($n = 2$). Only one developing country, Nepal, (Wilson et al., 2016) was represented and one study reported estimates from two countries (Deville et al., 2014). Studies covering a part of the country were again most frequent in the United States ($n = 32$). This was followed by China, Australia, Japan and the United Kingdom as well as a handful of developing nations (namely, Haiti, Bangladesh, Mozambique, Myanmar and Vietnam). A little more than half (54%) of the studies were focused on a single city.

Two types of spatial unit for the estimates dominated; 57% of studies used some form of administrative unit (for example, a statistical division, a county, a township or metropolitan region) and 29% used grids of varying sizes. Other types of spatial unit included points, buildings, or some custom units such as Voronoi

polygons. A number of studies ($n = 11$) reported estimates for more than one type of spatial unit.

There was a large variation in the number of spatial units used to produce estimates ranging from 1 to 52,000 with a median value of 34. A total of 10 studies reported on a single spatial unit including islands (Edmondson and Nantucket Data Platform team, 2019), a single administrative region (e.g. Fehr & Peers, 2014), a city (e.g. Goldschmidt and Dahl, 1976) and a university campus (Charles-Edwards and Bell, 2013). Studies reporting estimates for larger numbers of units provided estimates at the building level (Ara, 2014), high-resolution grids (Adamiak et al., 2017) or administrative areas (e.g. McKenzie et al., 2013), and were often alongside national or large sub-national area coverages. Half of the studies ($n = 47$) did not report the exact number of spatial units for which estimates were produced.

Temporal units of estimates. No study revealed efforts to produce and maintain an ongoing series of estimates. In contrast to the majority of studies presenting ad hoc analyses using opportunistic data sources only a few analyses and associated estimates received any form of update. Examples of these studies include, ESRI Daytime Population (Esri, 2018) and analyses of Australian Census from 1991 (Bell and Ward, 1998) to 2018 (Charles-Edwards and Panczak, 2018).

The majority of studies (77%) report on daytime (or both daytime and seasonal) populations and this reflected in the choice of the temporal units used in the estimates. A total of 36% of studies used minutes or hours as their temporal unit, with seconds as the most common unit that was split between daytime and night-time estimates (22%) or daytime only (15%) (Fig. 4).

Estimated population size and composition. There was a large variation in the sizes of estimated populations, ranging from 1,868 to 53,349,074 with a median value of 355,100. The size of dataset used for estimation ranged from 422 to 55,963,096 with median value of 722,000. Combining the size of the dataset and the size of estimated populations, largest populations were used or estimated when mobile phone data were employed, followed by smaller amount of transport-based studies (Fig. 5). Other large studies used monthly sales tax data from counties in North Carolina to derive “equivalent residents” surplus to the permanent population.

In terms of composition, the majority of studies focused on total populations present or working populations. For those studies concerned with population estimates of summer resort towns the total population was extended to include temporary or seasonal visitors (Erickson, 1961; Goldschmidt and Dahl, 1976).

Country of estimates & national coverage

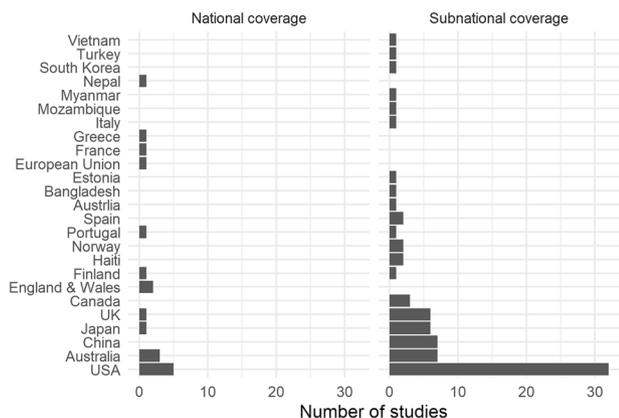


Fig. 3 Country of temporary population estimates and national coverage. Study of Deville et al. (2014) describes two countries.

Type of temporal units used

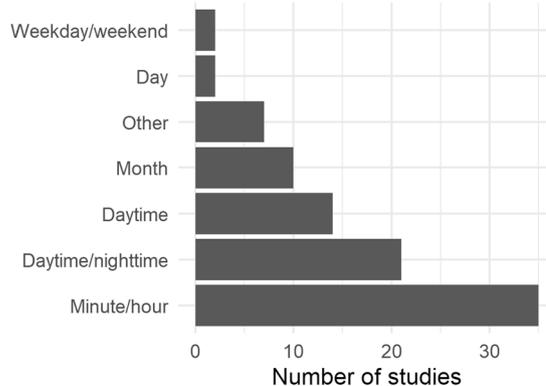


Fig. 4 Temporal units used for estimates.

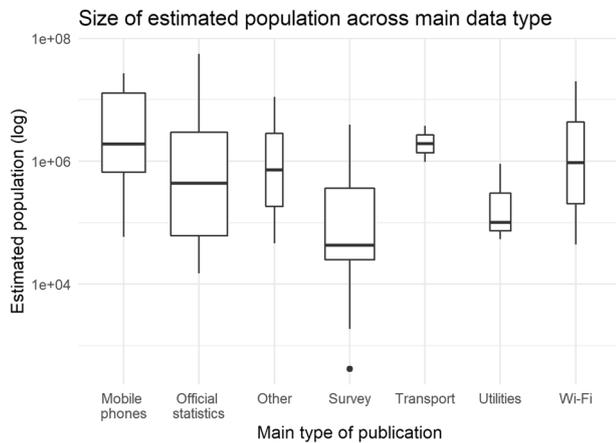


Fig. 5 Size of population by main data type. The first category used. Box width is proportional to the total number of studies. Studies that did not report on an estimated population size the size of dataset was used whenever possible.

The majority of studies (71%) did not report on any characteristics of estimated populations. This was often driven not by the theoretical or practical considerations but rather by data availability. For instance, many studies drawing on mobile phone data did not have access to any individual-level information (such as, age, sex, reason for visitation). Some studies sought to capture specific groups of temporary movers including retired individuals undertaking seasonal movements (Rose and Kingma, 1989; Happel and Hogan, 2002), school pupils (Campbell, 2010), second home users (Adamiak et al., 2017), indigenous populations (Warchivker et al., 2000), or on functional areas such as a university campus (Charles-Edwards and Bell, 2013).

Reproducibility. 61% of studies did not report estimates nor indicate where they could be obtained. Similarly, 95% of studies did not report on the type of software and/or code used to produce the estimates. In some cases, particularly when relatively simple methods were used (for example, equation based), this does not limit reproducibility or replicability of the studies or the possibility of their implementation across different situational contexts. However, for the remaining studies, particularly those based on proprietary data, the prospects of generating estimates across different areas allied with their validation are limited. More particularly, validation of the estimates was not common practice with only 36% reporting some attempt at validation with a lack of suitable data often quoted as a main reason for not undertaking the task.

Conclusions

This study sought to systematically review the empirical literature on temporary population estimation to identify the contemporary state of the art. To this end we revealed an interest and demand for temporary population estimates that is approaching 100 years in the making. However, it has been over the past two decades where there has been strong growth in the number of studies that is in part driven by the rise in both the type and availability of emerging sources of ICT-based data. This growth is likely to continue, set against the broader backdrop of 'big data' these new databases will need to be assessed and evaluated for their relevance for temporary population estimation along with their capacity to either unveil previously unknown temporary population dynamics or act as additional information source to bolster existing sources (both

traditional and emergent). To this end recent studies using credit card information (see for example, De Montjoye et al., 2015; Whitaker, 2019) would seem one of a number of promising new sources of data with the potential to reveal important dimensions of population mobility. That said, despite the rise in the number of studies using emergent sources of information, there remains no single source of data for the estimation of temporary populations. Rather traditional sources of information including the census and survey data remain an important backbone through which temporary population estimates are founded.

Notwithstanding the growing prevalence of temporary population studies, a disconnect currently exists with the lack of availability of such estimates through national statistical agencies. To this end the translation of temporary population studies into standardised products appears yet to occur in lieu of rising demand for such information that reflect the so far unresolved conceptual and methodological challenges. Often, in the absence of official estimates, cities and regions generate their own estimates (Lamb, 1999; Monmouth County Planning Board, 2008; Fehr & Peers, 2014; Hodur and Bangsund, 2015; Edmondson and Nantucket Data Platform team, 2019; Harris et al., 2019). However, it would appear that this situation is likely to change with the recent emergence of major coordinated programs of research focused on temporary population estimates to include ENhancing ACTivity and population mapping (ENACT) (Batista e Silva et al., 2017) and the Population 24/7 (Martin et al., 2015) projects (see also the subsequent work in Supplementary Table 3 online). It is likely that through these dedicated research programs the conceptual and methodological challenges will be systematically addressed and resolved.

It is important to highlight two interrelated limitations of the current study. The first pertains to the fragmented nature of the temporary population literature and its presence across multiple academic domains. For instance 28.1% of studies included in this review were found outside of the literature searches. Both the range and number of keywords (13 individual key words) used to capture studies of temporary populations reviewed here underscore this issue with the result of generating a large amount of false positive findings during the literature search. Second, there is a relatively large volume of grey literature represented. To this end, 27% of studies included in this study were published as reports. As such, this makes it difficult to systematically capture, classify as well as monitor the development in this field.

In sum, there is little doubt that the field of temporary population estimation will continue to grow however there remains a number of important challenges. Conceptually there is need for the research community and national statistical agencies to develop and agree on clear definitions that enable temporary populations to be classified in order to differentiate them from permanent residents. Practically, there remains much work to complete that comprehensively captures and measures the diverse array of spatial behaviours that involve different spatial circuits which are undertaken for a myriad of different purposes. Coordinated efforts that address both of these remaining challenges will ensure that we progress towards a more robust conceptual and empirical framework for the estimation of temporary populations.

Data availability

The datasets generated during and/or analysed during the current study are available in the Zenodo repository, <https://doi.org/10.5281/zenodo.3521082>.

Received: 20 November 2019; Accepted: 1 April 2020;

Published online: 06 May 2020

Notes

1 Whilst we adopt the definition of temporary populations by Smith (1989) for reasons of inclusiveness and simplicity, we recognise there exists a number of alternate definitions. This is because human spatial mobility is continuous through space and time (de Gans, 1994). For practical purposes it is necessary to delimit temporal and spatial bounds to facilitate measurement and analysis (Gober and Mings, 1984).

References

- Adamiak C, Pitkänen K, Lehtonen O (2017) Seasonal residence and counter-urbanization: the role of second homes in population redistribution in Finland. *Geojournal* 82:1035–1050. <https://doi.org/10.1007/s10708-016-9727-x>. Seasonal
- Ara S (2014) Impact of temporal population distribution on earthquake loss estimation: a case study on Sylhet, Bangladesh. *Int J Disaster Risk Sci* 5:296–312. <https://doi.org/10.1007/s13753-014-0033-2>
- Batista e Silva F, Rosina K, Schiavina Marcello M et al. (2017) Spatiotemporal mapping of population in Europe: The “ENACT” project in a nutshell. In: 57th European Regional Science Association (ERSA) Congress, Groningen, The Netherlands
- Batran M, Mejia M, Kanasugi H et al. (2018) Inferencing human spatiotemporal mobility in Greater Maputo via mobile phone big data mining. *ISPRS Int J Geo-Inf* 7:259. <https://doi.org/10.3390/ijgi7070259>
- Bell M (2000) Data, theory, method: understanding circulation in Australia. In: 10th Biennial conference of the Australian Population Association, Melbourne, Australia
- Bell M (2004) Measuring temporary mobility: dimensions and issues. Queensland Centre for Population Research School of Geography, Planning and Architecture, The University of Queensland, Brisbane, Australia
- Bell M, Ward G (1998) Patterns of temporary mobility in Australia: evidence from the 1991 census. *Austral Geogr Stud* 36:58–81. <https://doi.org/10.1111/1467-8470.00039>
- Bell M, Ward G (2000) Comparing temporary mobility with permanent migration. *Tour Geogr* 2:87–107. <https://doi.org/10.1080/146166800363466>
- Bengtsson L, Lu X, Thorson A et al. (2011) Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a Post-Earthquake Geospatial Study in Haiti. *PLoS Med* 8:e1001083. <https://doi.org/10.1371/journal.pmed.1001083>
- Bhaduri B, Bright E, Coleman P, Urban ML (2007) LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *Geojournal* 69:103–117. <https://doi.org/10.1007/s10708-007-9105-9>
- Boeing G (2018) Estimating local daytime population density from census and payroll data. *Reg Stud Reg Sci* 5:179–182. <https://doi.org/10.1080/21681376.2018.1455535>
- Breese GW (1947) Daytime population of the central business district of Chicago with particular reference to the factor of transportation. University of Chicago
- Campbell N (2010) Developing an extensible solution for modelling daytime population movement. MSc Thesis, University of Southampton
- Census of England and Wales 1921 (1925) Workplaces. HMSO, London
- Charles-Edwards E (2011) Modelling flux: towards the estimation of small area temporary populations in Australia. PhD Thesis, The University of Queensland
- Charles-Edwards E, Bell M (2013) Estimating the service population of a large metropolitan university campus. *Appl Spat Anal Policy* 6:209–228. <https://doi.org/10.1007/s12061-012-9079-y>
- Charles-Edwards E, Panczak R (2018) Elsewhere in Australia: a snapshot of temporary mobility on the night of the 2016 census. *Aust Popul Stud* 2:14–25
- Charles-Edwards E, Panczak R, Corcoran J, Bell M (2020) A framework for official temporary population statistics. *J Off Stat* 36(1):1–24. <https://doi.org/10.2478/jos-2020-0001>
- Chen J, Pei T, Shaw S-L et al. (2018) Fine-grained prediction of urban population using mobile phone location data. *Int J Geogr Inf Sci* 32:1770–1786. <https://doi.org/10.1080/13658816.2018.1460753>
- Collins A, Greaves S (2007) Daytime population tracking for planning and pollution exposure assessment. *Road Transp Res J* 16:55–68
- Crols T, Malleon N (2019) Quantifying the ambient population using hourly population footfall data and an agent-based model of daily mobility. *GeoInformatica* 23:201–220. <https://doi.org/10.1007/s10707-019-00346-1>
- de Gans H (1994) Some reflections on the concept of migration. In: Hooimeijer P, van der Knapp B, van Weesep J, Woods RI (eds) Population dynamics in Europe. Royal Netherlands Geographical Society, Utrecht
- De Montjoye YA, Radaelli L, Singh VK (2015) Unique in the shopping mall: on the reidentifiability of credit card metadata. *Science* 347(6221):536–539. <https://doi.org/10.1126/science.1256297>
- Deville P, Linard C, Martin S et al. (2014) Dynamic population mapping using mobile phone data. *Proc Natl Acad Sci USA* 111:15888–15893. <https://doi.org/10.1073/pnas.1408439111>
- Edmondson B. Nantucket Data Platform team (2019) Making it count. A data-driven look at Nantucket’s dynamic population. Nantucket Data Platform
- Elliott JH, Turner T, Clavisi O et al. (2014) Living systematic reviews: an emerging opportunity to narrow the evidence–practice gap. *PLoS Med* 11:1–6. <https://doi.org/10.1371/journal.pmed.1001603>
- Erickson RB (1961) Measuring resort population increases. *Prof Geogr* 13:16–19. https://doi.org/10.1111/j.0033-0124.1961.134_16.x
- Esri (2018) Methodology statement: 2018 Esri daytime population. Esri, Redlands
- Fehr & Peers (2014) Napa county travel behavior study. Fehr & Peers
- Foley DL (1952) The daily movement of population into central business districts. *Am Sociol Rev* 17:538–543
- Foley DL (1954) Urban daytime population: a field for demographic-ecological analysis. *Soc Forces* 32:323–330. <https://doi.org/10.2307/2574113>
- Francois R (2017) Bibtex: Bibtex parser
- Fulton PN (1984) Estimating the daytime population with the urban transportation planning package. *Transp Res Rec* 981:25–27
- Gao X, Yuan H, Qi W, Liu S (2014) Assessing the social and economic vulnerability of urban areas to disasters: a case study in Beijing, China. *Int Rev Spat Plan Sustain Dev* 2:42–62. https://doi.org/10.14246/irspsd.2.1_42
- Gober P, Mings RC (1984) A geography of nonpermanent residence in the U.S. *Prof Geogr* 36:164–173. <https://doi.org/10.1111/j.0033-0124.1984.00164.x>
- Goldschmidt PG, Dahl AW (1976) Demoflush estimating population in seasonal resort communities. *Growth Change* 7:44–48. <https://doi.org/10.1111/j.1468-2257.1976.tb00305.x>
- Gough D (2015) Qualitative and mixed methods in systematic reviews. *Syst Rev* 4:181. <https://doi.org/10.1186/s13643-015-0151-y>
- Graebert MB, Wyckoff M, Bretz L (2014) Northwest Michigan seasonal population analysis. Michigan State University Land Policy Institute
- Greger K (2015) Spatio-temporal building population estimation for highly urbanized areas using GIS. *Trans GIS* 19:129–150. <https://doi.org/10.1111/tgis.12086>
- Happel SK, Hogan TD (2002) Counting snowbirds: the importance of and the problems with estimating seasonal populations. *Popul Res Policy Rev* 21:227–240. <https://doi.org/10.1023/A:1019560405310>
- Harris E, Leaver J, Young N, Perlich PS (2019) Washington county temporary resident population estimates, 2017. Kem C. Gardner Policy Institute, Salt Lake City
- Himoto K, Kimata J (2014) A Model for the Spatio-temporal Distribution of Population using Country-Wide Statistical Data and Its Application to the Estimation of Human Exposure to Disasters. In: Teodorescu HN, Kirshenbaum A, Cojocaru S, Bruderlein C (eds) Improving Disaster Resilience and Mitigation – IT Means and Tools. NATO Science for Peace and Security Series C: Environmental Security. Springer, Dordrecht, pp. 73–87. https://doi.org/10.1007/978-94-017-9136-6_5
- Hodur NM, Bangsund (2015) Service population estimates for city of Williston. North Dakota State University, Fargo
- Horanont T, Shibasaki R (2010) Estimate Ambient Population Density: Discovering The Current Flow Of The City. Asian Conference On Remote Sensing (Acrs). https://www.academia.edu/2004297/estimate_ambient_population_density_discovering_the_current_flow_of_the_city
- Institute for Research in Social Science University of North Carolina (1952) Population distribution, spatial and temporal study of daytime–nighttime differentials in the proportional distribution of the total population of selected urban areas. Chapel Hill
- Järv O, Tenkanen H, Toivonen T (2017) Enhancing spatial accuracy of mobile phone data using multi-temporal dasymetric interpolation. *Int J Geogr Inf Sci* 31:1630–1651. <https://doi.org/10.1080/13658816.2017.1287369>
- Kashiyama T, Pang Y, Sekimoto Y (2017) Open PFLOW: creation and evaluation of an open dataset for typical people mass movement in urban areas. *Transp Res Part C* 85:249–267. <https://doi.org/10.1016/j.trc.2017.09.016>
- Kavanaugh P (1990) A method for estimating daytime population by small area geography. In: Proceedings of the 1990 Urban and Regional Information Systems Association conference, Edmonton, Alberta, pp. 150–164
- Khakpour M, Rød JK (2016) An attraction-based cellular automaton model for generating spatiotemporal population maps in urban areas. *Environ Plan B* 43:297–319. <https://doi.org/10.1177/0265813515604262>
- Khodabandelou G, Gauthier V, El-Yacoubi M, Fiore M (2016) Population estimation from mobile network traffic metadata. In: 2016 IEEE 17th International symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM). IEEE, Coimbra, Portugal, pp. 1–9
- Kim H, Ahn J-S (2017) Modelling daytime population distribution for emergency response and social vulnerability assessment. *J Korean Geogr Soc* 51:335–346
- Kobayashi T, Medina RM, Cova TJ (2011) Visualizing diurnal population change in urban areas for emergency management. *Prof Geogr* 63:113–130. <https://doi.org/10.1080/00330124.2010.533565>
- Kontokosta CE, Johnson N (2017) Urban phenology: toward a real-time census of the city using Wi-Fi data. *Comput Environ Urban Syst* 64:144–153. <https://doi.org/10.1016/j.compenvurbysys.2017.01.011>
- Lamb G (1999) Estimating the seasonal population of Door county. <https://www.doorcountybusiness.com/wp-content/uploads/2010/10/SeasonalPopulation.pdf>

- Lau KH (2009) A GIS-based stochastic approach to generating daytime population distributions for vehicle route planning. *Trans GIS* 13:481–502. <https://doi.org/10.1111/j.1467-9671.2009.01175.x>
- Lee S (1999) Service Population Pilot Study: an investigation to assess the feasibility of producing service population estimates for selected LGAs. Demography Section, Australian Bureau of Statistics, Perth
- Liu Z, Ma T, Du Y et al. (2018) Mapping hourly dynamics of urban population using trajectories reconstructed from mobile phone records. *Trans GIS* 1–20. <https://doi.org/10.1111/tgis.12323>
- Luedecke D (2020) sjmisc: Data and Variable Transformation Functions. R package version 2.8.4. <https://cran.r-project.org/web/packages/sjmisc/index.html>
- Lwin KK, Sugiura K, Zettsu K (2016) SpaceTime multiple regression model for grid-based population estimation in urban areas. *Int J Geogr Inf Sci* 30:1579–1593. <https://doi.org/10.1080/13658816.2016.1143099>
- Ma YJ, Xu W, Zhao XJ, Li Y (2017) Modeling the hourly distribution of population at a high spatiotemporal resolution using subway smart card data: a case study in the central area of Beijing. *ISPRS Int J Geo-Inf* 6: <https://doi.org/10.3390/ijgi6050128>
- Markham F, Bath J, Taylor J (2013) New directions in Indigenous service population estimation. Centre for Aboriginal Economic Policy Research, The Australian National University
- Martin D, Cockings S, Leung S (2015) Developing a flexible framework for spatiotemporal population modeling. *Ann Assoc Am Geogr* 105:754–772. <https://doi.org/10.1080/00045608.2015.1022089>
- McKenzie F, Canterford S (2016) Demographics for fire risk analysis regional Victoria and peri-urban Melbourne. Department of Environment, Land, Water and Planning, Melbourne, Australia
- McKenzie B, Koerber W, Fields A et al. (2013) Commuter-adjusted population estimates: ACS 2006–10. Journey to Work and Migration Statistics Branch, U.S. Census Bureau
- McPherson T, Brown M (2003) U.S. Day and Night Population Database (Revision 2.0)—description of methodology. Los Alamos National Laboratory
- Menzler FAA (1952) An estimate of the day-time population of London. *J Town Plan Inst* 38:116–120
- Moher D, Liberati A, Tetzlaff J et al. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6: <https://doi.org/10.1371/journal.pmed.1000097>
- Monmouth County Planning Board (2008) Summer coastal population study. Monmouth County Planning Board
- Nelson AC, Nicholas JC (1992) Estimating functional population for facility planning. *J Urban Plan Dev* 118:45–58
- New York Regional Plan Association (1949) Persons and vehicles entering Manhattan South of 61st Street, 1924–1948. *Reg Plan Bull* 74:1–6
- Office for National Statistics (2013) 2011 census: the workday population of England and Wales—an alternative 2011 census output base. Office for National Statistics
- Picornell M, Ruiz T, Borge R et al. (2018) Population dynamics based on mobile phone data to improve air pollution exposure assessments. *J Expo Sci Environ Epidemiol* 29:278–291. <https://doi.org/10.1038/s41370-018-0058-5>
- Qi W, Liu S, Gao X, Zhao M (2015) Modeling the spatial distribution of urban population during the daytime and at night based on land use: a case study in Beijing, China. *J Geogr Sci* 25:756–768. <https://doi.org/10.1007/s11442-015-1200-0>
- Ratti C, Frenchman D, Pulselli RM, Williams S (2006) Mobile landscapes: using location data from cell phones for urban analysis. *Environ Plan B* 33:727–748. <https://doi.org/10.1068/b32047>
- R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Reades J, Calabrese F, Sevtsuk A, Ratti C (2007) Cellular census: explorations in urban data collection. *Pervasive Comput* 6:30–38. <https://doi.org/10.1109/MPRV.2007.53>
- Rigall-I-Torrent R (2010) Estimating overnight de facto population by forecasting symptomatic variables: an integrated framework. *J Forecast* 29:635–654. <https://doi.org/10.1002/for.1166>
- Roddis S (1996) Metropolitan activity profiles—a new perspective on demand. In: 20th Australasian Transport Research Forum conference, Auckland, New Zealand
- Rose LS, Kingma HL (1989) Seasonal migration of retired persons: estimating its extent and its implications for the state of Florida. *J Econ Soc Meas* 15:91–104. <https://doi.org/10.3233/JEM-1989-15106>
- Schmitt RC (1956) Estimating daytime populations. *J Am Plan Assoc* 22:83–85. <https://doi.org/10.1080/01944365608979227>
- Seattle City Planning Commission (1951) Daytime and night-time population distribution in metropolitan Seattle: April 1950. Seattle City Planning Commission
- Sekimoto Y, Shibasaki R, Kanasugi H et al. (2011) PFlow: reconstructing people flow recycling large-scale social survey data. *IEEE Pervasive Comput* 10:27–35. <https://doi.org/10.1109/MPRV.2011.43>
- Sharp H (1955) The non-residential population of the central business district. *Land Econ* 31:378–381. <https://doi.org/10.2307/3159307>
- Silm S, Ahas R (2010) The seasonal variability of population in Estonian municipalities. *Environ Plan A* 42:2527–2546. <https://doi.org/10.1068/a43139>
- Smith SK (1987) How to tally temporary populations. *Am Demogr* 9:44–45
- Smith SK (1989) Toward a methodology for estimating temporary residents. *J Am Stat Assoc* 84:430–436
- Smith G, Arnot C, Fairburn J, Walker G (2005) A National Population Data Base for major accident hazard modelling. Health and Safety Executive
- Smith SK, House M (2007) Temporary migration: a case study of Florida. *Popul Res Policy Rev* 26:437–454
- Sundararaj B, Cheshire J, Longley P (2019) Estimating real-time high-street footfall from Wi-Fi probe requests. *Int J Geogr Inf Sci* 0:1–19. <https://doi.org/10.1080/13658816.2019.1587616>
- Stathakis D, Baltas P (2018) Seasonal population estimates based on night-time lights. *Comput Environ Urban Syst* 68:133–141. <https://doi.org/10.1016/j.compenvurbysys.2017.12.001>
- Stutz FP, Parrott R, Kavanaugh P (1992) Charting urban space-time population shifts with trip generation models. *Urban Geogr* 13:468–474. <https://doi.org/10.2747/0272-3638.13.5.468>
- Swanson DA, Tayman J (2011) On estimating a de facto population and its components. *Rev Econ Finance* 1:17–31. <http://www.bapress.ca/Journal-5/On%20Estimating%20a%20De%20Facto%20Population%20and%20Its%20Components%20By%20David%20A.%20Swanson.pdf>
- Taubenböck H, Roth A, Dech S (2007) Linking structural urban characteristics derived from high resolution satellite data to population distribution. In: Coors V, Rumor M, Fendel EM, Zlatanova (eds) Urban and regional data management: UDMS 2007 annual conference. Taylor & Francis Group, London, Stuttgart, Germany, pp. 35–46
- Taylor J (1998) Measuring short-term population mobility among indigenous Australians: options and implications. *Aust Geogr* 29:125–137. <https://doi.org/10.1080/00049189808703207>
- Terada M, Nagata T, Kobayashi M (2013) Population estimation technology for mobile spatial statistics. *NTT DOCOMO Tech J* 14:10–15
- Thakur R (2018) A methodology to account for seasonal population shifts in North Carolina. MSc Thesis, The University of North Carolina at Charlotte
- Thomas KV, Amador A, Baz-Lomba JA, Reid M (2017) Use of mobile device data to better estimate dynamic population size for wastewater-based epidemiology. *Environ Sci Technol* 51:11363–11370. <https://doi.org/10.1021/acs.est.7b02538>
- Thornthwaite C (1929) Louisville, Kentucky: a study in urban geography. PhD Thesis, Department of Geography, University of California
- Toole JL, Colak S, Sturt B et al. (2015) The path most traveled: Travel demand estimation using big data resources. *Transportation Res Part C* 58:162–177. <https://doi.org/10.1016/j.trc.2015.04.022>
- Walker L, Barros J (2012) An agent-based population model for Wolverhampton, UK: a spatio-temporal activity based approach to population modelling. In: GISRUK, Lancaster
- Warchivker I, Tjapangati T, Wakerman J (2000) The turmoil of aboriginal enumeration: mobility and service population analysis in a Central Australian community. *Aust NZ J Public Health* 24:444–449. <https://doi.org/10.1111/j.1467-842X.2000.tb01610.x>
- Weir TR (1960) A survey of the daytime population of Winnipeg. *Queen's Q* 67:654–664
- Wheeler JO, Brunn SD (2002) An urban geographer before his time: C. Warren Thornthwaite's 1930 doctoral dissertation. *Prog Hum Geogr* 26:463–486. <https://doi.org/10.1191/0309132502ph381ra>
- Whitaker SD (2019) 12 Facts about temporary urbanists. Economic Commentary. <https://www.clevelandfed.org/en/newsroom-and-events/publications/economic-commentary/2019-economic-commentaries/ec-201907-12-facts-about-temporary-urbanists.aspx> Last accessed 26 Jan 2020
- Wickham H (2016) ggplot2: Create elegant data visualisations using the grammar of graphics. R package version 3.3.0. <https://cran.r-project.org/web/packages/ggplot2/index.html>
- Wickham H (2018) Scales: scale functions for visualization. R package version 1.1.0. <https://cran.r-project.org/web/packages/scales/index.html>
- Wickham H (2019a) Forcats: tools for working with categorical variables (factors). R package version 0.5.0. <https://cran.r-project.org/web/packages/forcats/index.html>
- Wickham H (2019b) Stringr: simple, consistent wrappers for common string operations. R package version 1.4.0. <https://cran.r-project.org/web/packages/stringr/index.html>
- Wickham H, Bryan J (2019) Readxl: read excel files. R package version 1.3.1. <https://cran.r-project.org/web/packages/readxl/index.html>
- Wickham H, Francois R, Henry L, Mueller K (2019) Dplyr: a grammar of data manipulation. R package version 0.8.5. <https://cran.r-project.org/web/packages/dplyr/index.html>
- Wickham H, Henry L (2019) Tidy: tidy messy data. R package version 1.0.2. <https://cran.r-project.org/web/packages/tidy/index.html>

- Wilson R, zu Erbach-Schoenberg E, Albert M et al. (2016) Rapid and near real-time assessments of population displacement using mobile phone data following disasters: the 2015 Nepal earthquake. *PLoS Currents Disasters* <https://doi.org/10.1371/currents.dis.d073fbee328e4c39087bc086d694b5c>
- Wurtele ZS, Wellisch JB (1968) Population dynamics: final report. System Development Corporation, Santa Monica
- Xie Y (2019) Knitr: A general-purpose package for dynamic report generation in r. R package version 1.28. <https://cran.r-project.org/web/packages/knitr/index.html>
- Yong li M (1998) Research in statistical models for mobile population (SMMP). *Cartogr J* 35:155–164. <https://doi.org/10.1179/caj.1998.35.2.155>

Acknowledgements

The authors wish to thank the Australian Bureau of Statistics (ABS), in particular - Denise Carlton, Andrew Howe, Anthony Grubb, Glen West, Beidar Cho and Karen Ecclestone of ABS for their input on various aspects of the paper. However, the interpretations of the analysis are solely those of the authors and do not necessarily reflect the views and opinions of the ABS or any of their employees. We would like to also thank the authors that responded to queries about details of their work: Christoph Aubrecht, Filipe Batista e Silva, Andrew Collins, Donald G. Janelle, Nik Lomax, David Martin, Chris Needham, Andy Newing, Paola Pucci, Stanley Smith and Martijn Tennekes. This research was funded by the Australian Government through the Australian Research Council Linkage project scheme, LP160100305.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1057/s41599-020-0455-y>.

Correspondence and requests for materials should be addressed to R.P.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020