Journal of Transport & Health 6 (2017) 128–142

Contents lists available at ScienceDirect

Journal of Transport & Health

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

# Investigating geospatial data usability from a health geography perspective using sensitivity analysis: The example of potential accessibility to primary healthcare

Robin Frew[a,*], Gary Higgs[b], Jenny Harding[c], Mitchel Langford[b]

[a] Faculty of Computing, Engineering and Science, University of South Wales, Pontypridd CF37 1DL, United Kingdom
[b] GIS Research Centre, Wales Institute of Socio-Economic Research, Data and Methods (WISERD), University of South Wales, Pontypridd CF37 1DL, United Kingdom
[c] Ordnance Survey, Explorer House, Adanac Drive, Southampton SO16 0AS, United Kingdom

ARTICLE INFO

*Keywords:*
Accessibility analysis
Floating catchment area (FCA) methodologies
Alternative sources of network data
Data usability
Sensitivity analysis

ABSTRACT

Network distance and travel times are two popular methods of measuring potential geographic accessibility and networks are also used in gravity model-based approaches such as floating catchment area (FCA) techniques. Although some research has been conducted to assess the effectiveness of the representation of demand- (population) or supply- (destinations) side characteristics within such models, there have been few attempts to assess the implications of using alternative sources of network data. This study employs a sensitivity analysis approach to assess accessibility to GP surgeries in south Wales using proprietary and open sources of network data. Results suggest that there are significant differences between access scores derived from the use of networks which purport to portray the same features. Furthermore, the pattern of differences varies between urban and rural areas. Case studies are used to show that the actual representation of network-based features, often overlooked in previous research, can have important implications for the findings from such studies. We conclude by suggesting that the use of sensitivity analysis to assess geospatial data usability has a wider relevance for studies that involve the use of a range of GIS-based techniques in different application areas.

## 1. Introduction

This paper argues for the need to examine the usability of geospatial data sources when applied to 'typical' GIS-based analytical tasks, such as those undertaken in health geography studies. Our primary focus concerns the use of network data sets in examining spatial variations in potential accessibility to General Practitioners (GP) surgeries. By taking the novel approach of applying sensitivity analysis and comparing results obtained from using different sources of spatial data within models typically used to assess geographical variation in access to health facilities, assessments can be made on the usability, or appropriateness, of such data sets in context. The number and variety of sources of spatial data has increased in recent years leading to wider debates regarding the quality and usability of such data, particularly in the light of the increased availability of Free and Open Source (FOS) or free-to-use data including volunteered geographic information (VGI) for GIS modelling applications (Goodchild and Li, 2012; Haklay, 2010a, Senaratne et al., 2016).

Data from national mapping agencies, typically well-documented and assumed to be of the highest quality available, is often

* Corresponding author.
*E-mail address:* robin.frew@southwales.ac.uk (R. Frew).

http://dx.doi.org/10.1016/j.jth.2017.03.013
Received 2 August 2016; Accepted 17 March 2017
Available online 28 April 2017

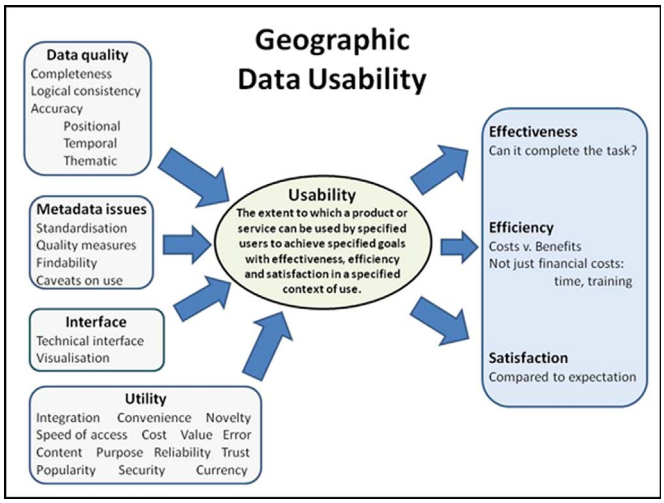


**Fig. 1.** Usability at the centre of quality, metadata, interface and utility assessments. Based on ISO 9241-11 (ISO, 2010) and ISO 19157:2013 (ISO, 2013).

expensive which makes the option of cost-free data sets tempting for many users; particularly those working in the public and third sectors in periods of austerity and reduced IT financial budgets. Concerns over VGI data quality and trust may be higher than those relating to proprietary GI (Goodchild, 2007), but advantages of other usability aspects such as currency and cost have meant that crowd-sourced and VGI geospatial data are increasingly used in GIS studies. This has led to a number of recent studies comparing usability issues of VGI products, such as OpenStreetMap (OSM), with those of 'official' sources of digital data in GIS applications (e.g. Brovelli et al., forthcoming; Du et al., forthcoming).

The concept of geospatial data usability is closely related to that of data quality, with Cai and Zhu (2015) amongst others actually classifying usability as a data quality element. However, with the user experience and the context of the use recognised as key, both in the publication of International Standards (ISO 9241-210, 2010) and the academic literature (see for example Haklay (2010b) and Brown et al. (2012)), a wider range of usability factors beyond that of a data-centric view of data quality are increasingly recognised (Fig. 1). Drawing on ISO 9241 (2010) geospatial data usability can be defined as the extent to which geospatial data can be used to achieve specified goals with effectiveness, efficiency and satisfaction, in a specified context of use. These three key characteristics of usability (effectiveness, efficiency and satisfaction) can be split into many component elements, all of which contribute to the usability of the data (as shown in Fig. 1), with the importance of each element varying according to the particular context and task, and with the potential to be grouped and classified in several different ways, again dependent on the particular context.

Previous usability studies involving geospatial data have tended to involve a battery of techniques including: timing how long a task takes to complete, assessing how well a task is completed, collating and assessing the resources needed to complete a task, and gauging user satisfaction compared to expectations (Harding and Pickering, 2007). Much of this is a subjective, qualitative process involving time-consuming interviews and questionnaires (Harding, 2012). This study takes a quantitative approach to address one aspect of usability: namely effectiveness. The characteristics of each dataset all contribute differently to this aspect. By conducting sensitivity analysis to the different permutations of spatial data for the travel network representation in accessibility models, variations in results are highlighted in order to draw attention to the advantages and limitations of such data sources in this context. This objective approach involves examining how uncertainty in the output of a system (numerical or otherwise) can be apportioned to different sources of uncertainty in its inputs. The most significant sources of uncertainty can then become the focus for further research (European Commission, 2015). Though rarely applied in geographical contexts, sensitivity analysis is regularly used in the financial industry, in business planning and in the fields of medicine and health (Czitrom, 1999).

This study draws on the findings from a study of spatial variations in potential accessibility to primary health care facilities. There is a considerable literature on different approaches to measuring accessibility, especially in health studies where the accessibility of a population to a variety of medical facilities has come under considerable scrutiny (Higgs, 2004). Traditionally, such approaches have included relatively straightforward container and coverage methods which produce easily understood results from simple calculations but may be less appropriate at certain spatial scales or for smaller geographical areas. More recently, studies such as those of Burkey (2012), Delamater (2013) and Fransen et al. (2015) have drawn attention to the potential of more sophisticated tools for measuring accessibility using gravity-based approaches which incorporate sources of public transport data and networks (Biba et al., 2010; Mao and Nekorchuk, 2013; Langford et al., 2016).

Despite the relative plethora of studies investigating the application of these techniques, the use of health-based accessibility analysis to assess the usability of GI data is much less common. In particular, whilst previous studies such as Phibbs and Luft (1995), Bertazzon and Olson (2008), Apparicio et al. (2008) and Boscoe et al. (2013) have compared various distance-measurement methods such as Euclidean, Manhattan and true network distance, few have examined the implications of using different sources of network-based data. In one of the few examples to date, Jones (2010) used sensitivity analysis to compare walking times to medical facilities in the West Midlands using networks based on three Ordnance Survey (OS) products (OS MasterMap® Integrated Transport Network™