Translating networked based accessibility measures into an open source environment; challenges and opportunities

Richard Williams[1], Gary Higgs[1], Mitchel Langford[1], Tim Banks[2], Rhian Edwards[2]

1 Department of Computing, Faculty of Science, Computing and Engineering, University of South Wales, Pontypridd, CF37 1DL
2 Tenovus Cancer Care, Gleider House, Ty Glas Road, Cardiff, CF14 5BD

Summary

There is a large literature that has explored the use of 'floating catchment area' (FCA) techniques in measuring access to a range of primary and secondary health services. This paper describes the development of open source tools, which incorporate FCA techniques initially applied in the context of mapping and modelling access to mobile cancer services. Tests completed on the open source tool allow it be used to compare to an existing proprietary tool developed for use in ArcGIS. This study includes reflections on the underlying data used within these models as well as the software itself in order to better understand differences in outputs and computational speed when performing FCA analyses in an open source environment.

KEYWORDS: Healthcare Accessibility; GIS; Two Step Floating Catchment Area (2SFCA) tools; ,. Open Source, Cancer Services

1. Introduction

A large literature now exists which explores the use of 'floating catchment area' (FCA) techniques in measuring access to a range of primary and secondary health services. Despite their popularity, it is still difficult for those who might most benefit from these types of analysis to gain access to them. Thus, there is an urgent need to develop tools that incorporate these techniques into an open source environment so that they might be made freely available to researchers concerned with analysing variations in accessibility to a range of public and private services.

The two-step floating catchment area (2SFCA) algorithm, first introduced by Wang (2000), is used to measure spatial accessibility and is itself a special case of the gravity model (Joseph & Bantock 1982). It has been deployed in many healthcare accessibility studies (McGrail & Humphreys (2009), Ngui & Apparicio (2011)) since Luo & Wang (2003) first compared the results of 2SFCA to those of a standard gravity model. As its name implies, 2SFCA uses two algorithmic steps to create an accessibility score. Luo & Qi (2009) further developed the method in order to address some limitations such as, large steps of accessibility found at the edge of each catchment in the original 2SFCA, primarily by incorporating distance decay functions to create an enhanced 2SFCA (or E2SFCA).

An add-in tool to compute E2SFCA scores, that operates from within ArcMap and leverages the capabilities of ESRI's Network Analyst Extension, has been developed by Langford et al (2015) and made available online. This has been applied in a number of studies to date (Higgs et al (2015), Frew et al (2017), Higgs et al (2017)). This software makes it possible for an end user to input their own datasets and calculate FCA scores in an ArcGIS environment. It is possible for users with access to ArcGIS to conduct sensitivity analyses by using a range of supply-side and demand-side parameters within the network-based accessibility models.

This paper describes on-going research that is being conducted in partnership with the charity organisation Tenovus Cancer Care, which aims to build upon the previous work of Langford et al. by developing broadly parallel tools that can operate in an open source software environment. This will enable Tenovus, and other users, to perform such calculations without the need for proprietary software or indeed significant GIS expertise. Tenovus Cancer Care currently operate a number of mobile treatment units throughout Wales and they would like to adopt these advanced spatial analytical techniques in order to analyse levels of accessibility to these units and attempt to optimise their delivery locations in relation to potential demand for cancer services arising from within the community.

This presentation will provide a demonstration to an analytical tool that has been created using the FOSS spatially-enabled database PostGIS, its pgRouting network analysis extension, and the desktop QGIS software. All of these components are free and open source, and are rapidly establishing relatively a wide user-base within the GIS community. This presentation will also report on aspects of open source data quality and usability, which are needed for establishing a road network dataset. Along with our methodologies for determining potential service demand using data drawn from the UK Census population data along with supplementary information on the current locations and capacities of fixed site cancer service provision in Wales. This play a vital role in determining the most effective locations for the placement of mobile treatment services.

There are both benefits and potentially many challenges to working in the open source environment, and this project will document those that have been encountered to date, as well as looking at the key differences in processes and procedures when undertaking this task within ArcGIS versus the open source environment. A range of techniques will be used to show how the implementation of such tools varies between these two approaches and the consequences that this may have for how such tools operate.

2. Implementing FCA techniques in a proprietary and open source GIS packages

An ArcGIS add-in tool for performing FCA calculations has already been developed using ArcObjects, Visual Studio and the Microsoft .NET platform (Langford et. al., 2015). This tool requires the user to create or have access to a fully configured ArcGIS network data set, along with point coverages that identify the service supply and potential demand points needed to perform the FCA calculations. The tool presents a user-friendly interface that gathers key modelling parameters from the user and then constructs and executes underlying analyses using facilities provided in the Network Analyst extension, before presenting the results back to the uses through standard ArcMap attribute tables. This tool has been used in the Higgs et al (2015) study which investigated access to sports facilities in Wales, as well as a number of other published works (Higgs et al(2015), Frew et al (2017), Higgs et al (2017)). Although this tool has been made freely available to anyone with access to ArcGIS, it still requires the user to have access to proprietary software, hence the current objective of developing a similar solution that can operate in a FOSS environment.

Currently, a proof-of-concept plugin running in QGIS (Figure 1) has been built using a combination of the QGIS plug-in builder, QT Creator and Python programming scripts. This enables the computation of E2SFCA scores but with these being based only on Euclidean distances rather than the more sophisticated networked distances that are incorporated into the ArcGIS add-in tool. Testing of the current tool has highlighted a number of performance issues which will need to be addressed in future versions, as it is currently quite slow and inefficient when very large data sets are processed.

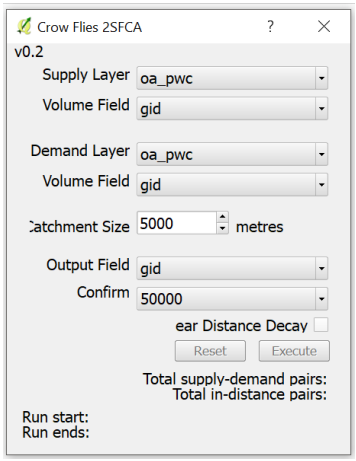

Figure 1: Qgis Plugin

A series of tests have been completed to date using this prototype tool, with the following visualisations created from data provided by Tenovus (Figures 2 & 3). The data received to date has allowed the tool to be exercised on both large and small data sets, and has already enabled Tenovus to get a better understanding of their resources and the efficiency of their operations throughout Wales. Tenovus operate a number of services and as well as their mobile chemotherapy units, and using the same tool it has been possible to visualise access to a range of their activities such as the promotion of self-support choir groups, charity shops, and local counselling services. Visualising this data has proved itself to be important as it allows the decision makers within Tenovus to understand the issues more clearly, and in particular it is able to show those areas with high and low accessibility scores, as well as offering the opportunity to conduct sensitivity analyses on the patterns revealed by using a range of input modelling parameters.

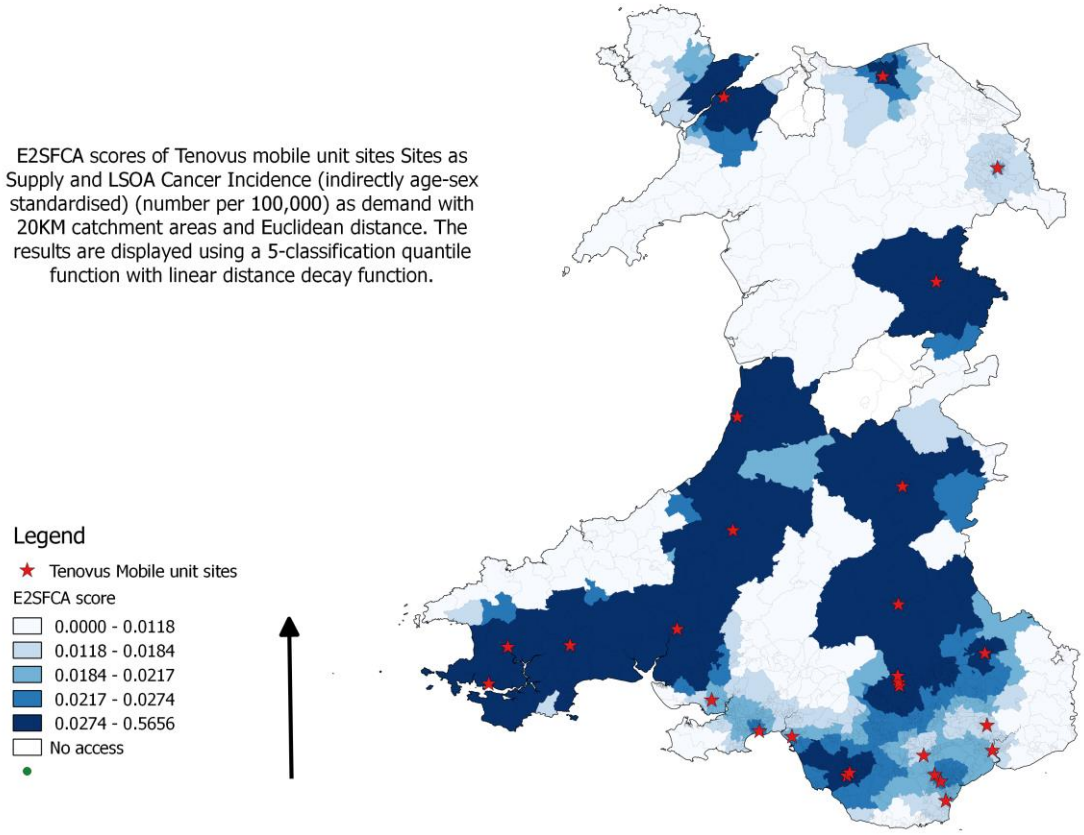

Figure 2: E2SFCA scores of access to mobile cancer treatment (20km catchment size)

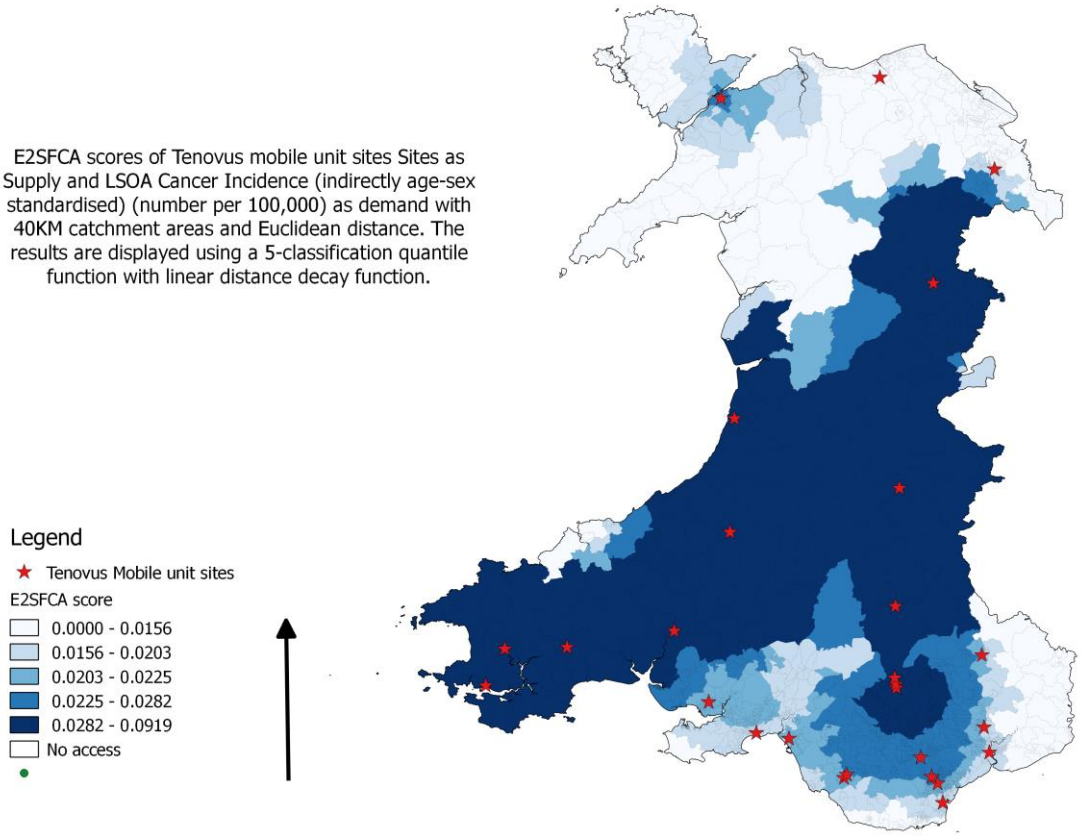

Figure 3: E2SFCA scores of access to mobile cancer treatment (40km catchment size)

3. Important considerations when implementing network-based models in open source environments

In attempting to create a tool such, there are complex decisions to be made, not least because there are potentially a number of different routes that can be taken in order to secure its development. Trying to ensuring that correct choices are made is likely to be the key ultimately to creating a stable and effective analytical tool. This presentation describe some of the key issues encountered to date, and the decisions that have been made along with justifications for these selections. It is hoped that by exploring and discussing these issues it will be possible for future researchers to have a better understanding of how open source and proprietary software can differ in its functions, speed and accuracy. To further support this enquiry a suite of tests will be designed and conducted to look at the following key issues;

- ArcGIS tool vs open source tool – Test and compare both the speed and accuracy of tools developed in each environment and explain any differences in outputs that may occur.
- ArcGIS Network Analyst vs pgRouting – both of these extensions claim to be able to do similar types of network tracing operations. By using both of our tools we will be used to test the speed of their respective operations, but also study whether subtle differences that exist in the way tat they implement network analyses have any significant impact upon the FCA output scores that are reported.
- Road Network Data – we will compare commonly used open source data that might be adopted for constructing the road network, such as Open Street Map and the OS OpenRoads product.
- Networked Distances vs Euclidean Distances – we intend to conduct a detailed analysis of the outputs derived from network distance versus single Euclidean distance as used in FCA models. Under what circumstances should either approach be used? How different are the scores that are produced? What are the cost-benefits of adopting a network solution with its associated overheads of data over the simple straight line solution?

- Network Build Speed – We will compare how long each network takes to be built dependent on size across the platforms and with different levels of build quality and again compare FCA scores to see if lighter networks have any significant difference to the more complex networks.
- Visualisation – We will investigate alternative ways to visualise the results of FCA outputs.
- Usability Testing – It is essential that the open-source tool is highly usable whilst still providing the required analytical capabilities. A range of user tests will be conducted in an attempt to establish whether the provided functionality is fit-for-purpose, whilst also assessing if the intended non-expert users will be able to clearly understand and correctly operate the analyses. A number of different user types will be involved in the testing including operatives at Tenovus Cancer Care who are the intended users of the tool. A number of different methodologies will be used to ensure a full user test is completed including one to one interviews and a small focus group of the users.

The tool will be made available on research gate and github for potential users to access the source code and the program once the PhD has been completed. The authors will also make the tool available if contacted directly.

A range of these computational challenges will be described in relation to the application of real world data on cancer services, using data supplied by Tenovus Cancer Care in order to investigate current provision in Wales and to illustrate how such techniques can be used by such organisations to optimise the provision and operation of their services.

4. Acknowledgments

The authors gratefully acknowledge funding from Knowledge Economy Skills Scholarships (KESS). KESS is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government's European Social Fund (ESF) convergence programme for West Wales and the Valleys.

5. Biography

Richard Williams is a 2nd year PhD candidate working with GIS and has a particular interest in developing open source tools to measure healthcare accessibility.

Gary Higgs is Professor of Geographical Information Science in the Faculty of Computing, Engineering and Science, University of South Wales and a co-Director of the Wales Institute of Social and Economic Research, Data and Methods (WISERD). Over-arching research interests are in the application of GIS in social and environmental studies, most recently in the areas of health geography and emergency planning.

Mitchel Langford is a Reader in the Faculty of Computing, Engineering and Science, University of South Wales. His current research interests include dasymetric mapping, population modelling, and geospatial analysis within the fields of healthcare, social equality and environmental justice.

Tim Banks is the Head of Research at Tenovus Cancer Care where he is responsible for funding and undertaking a variety of research projects aimed at improving the lives of those affected by cancer. Tim's interests and experiences include many areas of health research including, health inequalities, interpretations of disability and social justice.

Rhian Edwards is the Director of Research and Support at Tenovus Cancer Care. Rhian has a varied background and an interest in improving the lives of those affected by cancer.

References

Frew, R., Higgs, G. Harding, J. and Langford, M. (2017) "Investigating geospatial data usability from a health geography perspective using sensitivity analysis: the example of potential accessibility to

primary healthcare" Journal of Transport & Health.

Higgs, G., Langford, M. and Norman, P. (2015) Accessibility to sports facilities in Wales: A GIS-based analysis of socio-economic variations in provision, Geoforum, 62, 105-120.

Higgs, G., Zahnow, R, Corcoran, J. Langford, M. and Fry, R. (2017) "Modelling spatial access to General Practitioner surgeries: Does public transport availability matter?" Journal of Transport & Health. doi:10.1016/j.jth.2017.05.361

Joseph AE, Bantock PR: Measuring potential physical accessibility to general practitioners in rural areas: a method and case study. Social Science and Medicine. 1982, 16 (1): 85-90.

Langford, M., Higgs G. and Fry, R. 2015 "USW-FCA2: An ArcGIS add-In tool to compute Enhanced Two-Step Floating Catchment Area accessibility scores" Software Package – ArcMap add-in Available at:
https://www.researchgate.net/publication/287198887_USW-FCA2_An_ArcGIS_add In_tool_to_compute_Enhanced_Two-Step_Floating_Catchment_Area_accessibility_scores.

Luo W, Wang F: Measures of spatial accessibility to health care in a GIS environment: synthesis and a case in the Chicago region. Environment and Planning B: Planning and Design. 2003, 30: 865-884.

Luo, W.; Qi, Y. An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. Health Place 2009,15, 1100–1107.

McGrail MR, Humphreys JS: Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. Applied Geography. 2009, 29: 533-541. 10.1016/j.apgeog.2008.12.003

Ngui, A.N., Apparicio, P., Optimizing the two-step floating catchment area method for measuring spatial accessibility to medical clinics in Montreal BMC Health Serv. Res., 11 (2011), p. 166

Wang F, 2000 ``Modelling commuting patterns in Chicago in a GIS environment: a job accessibility perspective" The Professional Geographer, 52, 120-133