Geographically Weighted Cronbach's Alpha

Sui Zhang[*1] and Ziqi Li[†2]

[1]School of Social and Political Sciences, University of Glasgow
[2]School of Geographical and Earth Sciences, University of Glasgow

January 20, 2023

Summary

As individual's cognitions and behaviours are affected by where they live, the reliability of responses to tests or scales may vary with location. In this paper, we develop a local version of Cronbach's alpha, geographically weighted Cronbach's alpha, to investigate how the reliability of the measure varies spatially. Two demonstrations of exploratory applications of the three-wave geocoded measure of "social capital index" in the Baltimore metropolitan area, USA, was also performed in this paper.

KEYWORDS: reliability, Cronbach's alpha, GWR

1. Introduction

Cronbach's alpha is the most used reliability coefficient to measure the internal consistency of a test or scale. It provides a global measure of how correctly a certain collection of items can state individual differences for all respondents (Cronbach, 1951). However, as personal understanding of a question may vary with its context in empirical applications, the results of reliability tests are usually not equivalent to the true reliability. Moreover, a substantial body of literature has highlighted how where people live affects their cognitions and behaviours independently of any other personal characteristics over recent decades. Based on such local thinking, geographically weighted (GW) framework has been used to extend a series of algorithms to the spatial dimension, such as geographically weighted regression (GWR) (Brunsdon et al., 1996), geographically weighted principal component analysis (GWPCA) (Harris et al., 2011), and geographically weighted discriminant analysis (GWDA) (Brunsdon et al., 2007), etc. to investigate the potential spatial non-stationarity. Similarly, there is also a need for local test of reliability to better identify spatial variation in the internal consistency of measures. In this paper, we apply the geographically weighted paradigm to Cronbach's alpha to extend it to the local case, Geographically Weighted Cronbach's alpha, to understand how the reliability of measures varies over space. A three-wave geocoded measure of "social capital index" in the Baltimore metropolitan area, USA, was used for demonstration.

2. Methods

Cronbach's alpha is formulated as:

$$\alpha = \frac{k\bar{c}}{\bar{v} + (k-1)\bar{c}} \tag{1}$$

where k is the number of test or scale items, $\bar{c}$ is the average of all covariances between items, $\bar{v}$ and is the average variance of each item. α ranges from 0 to 1. The higher the internal consistency of the measure, the higher the α. Typically, an $\alpha > 0.8$ is the criterion for the reliability of a measurement to be acceptable in applied research, while $\alpha = 0.7$ and $\alpha = 0.9$ are the acceptance thresholds refer to the

[*] 2792981Z@student.gla.ac.uk
[†] ziqi.li@glasgow.ac.uk

early stage of research and important decision making respectively.

On this basis, Geographically Weighted Cronbach's alpha is defined as:

$$\alpha_i = \frac{k\overline{c}_i}{\overline{v}_i + (k-1)\overline{c}_i} \tag{2}$$

where the subscript *i* denotes estimates at location *i*, all other terms are the same as Equation 1. Similar to other geographically weighted algorithms, geographically weighted Cronbach's alpha generates reliability estimates for each respondent by borrowing data from nearby locations and weighting them by distance from the kernel centre. All bandwidths selected for the following demonstrations are adaptive and the data are weighted according to bi-square kernel function.

3. Demonstrations

Baltimore Ecosystem Study (BES) Household Telephone Survey was used for demonstration in this paper. The survey was mainly sampled in the Baltimore metropolitan area (Queen Anne's County was not sampled except in 2011 and is not included in the demonstration area of this paper) and was geocoded in 2003 (n = 1508), 2006 (n = 3312) and 2011 (n = 1636). The survey contains a seven-item five-point Likert scale to measure social capital on the neighbourhood scale (Grove and Locke, 2018). A work conducted by the BES research team in 2011, based on the cross-sectional data from 2003, selected five individual questions from this scale to represent the "social capital index": "People in the neighborhood are willing to help one another", "This is a close knit neighborhood", "People in this neighborhood can be trusted", "There are many opportunities to meet neighbors and work on community problems", and "Churches, temples and other volunteer groups actively support the neighborhood?" (Vemuri et al., 2011). Empirically, these scales are commonly assumed to be treated as interval variables. This paper replicated the reliability tests in this article and explored the spatial variation of the local test results.

After cleaning the data against incomplete responses, the global alpha was estimated to be 0.810, which is generally consistent with that reported (0.805) in Vemuri et.al. (2011). To perform local estimation for alpha, three bandwidths (10%, 30% and 70% of the total respondents) were chosen, which represented three typical scales of spatial process (local, regional, and quasi-global) respectively (**Figure 1**).

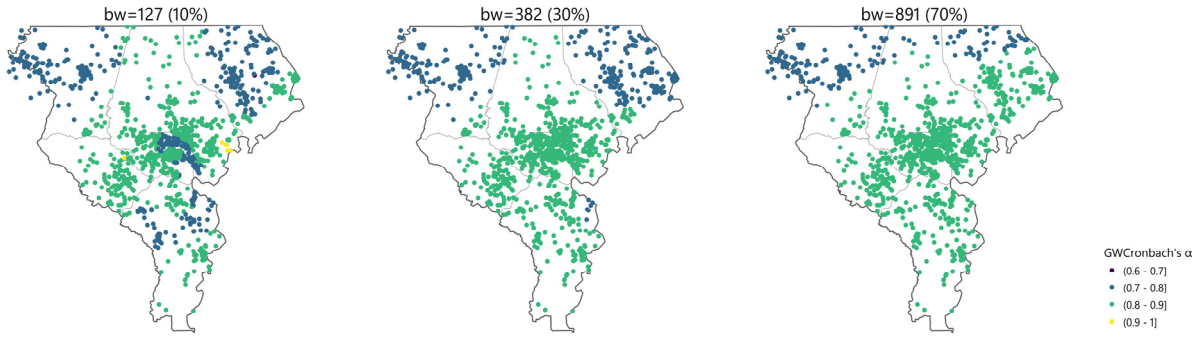

Figure 1 Local Cronbach's alpha estimates with different bandwidths (10%, 30% and 70%).

When the bandwidth was set at 10% of respondents, the local alphas showed significant spatial heterogeneity. A high proportion of responses in Baltimore City, Carroll County in the northwest, Harford County in the northeast, and Anne Arundel County in the south did not meet the reliability criterion of 0.8. Very few responses did not meet the criterion of 0.7, while some of the responses along

the Patapsco River area exceeded 0.9. When the bandwidth was increased to 30%, the spatial variability of the local alpha estimates is significantly reduced, the responses that did not reach 0.8 were mainly in the northwest and northeast, with very few responses from Anne Arundel County residents failing to reach the threshold and all responses from Baltimore City meeting the criterion. As the bandwidth increased to 70%, the number of those not meeting the criterion decreased further and was only found in parts of Carroll County and Harford County.

Additionally, another work published in the same journal as Vemuri et.al. (2011) in 2015 used cross-sectional data from the 2006 BES telephone survey to measure social capital as well (Holtan et al., 2015). However, Holtan et.al. (2015) chose to use the question "There is an active neighborhood association" instead of "Churches, temples and other volunteer groups actively support the neighborhood?". The samples in this study were narrowed to the Baltimore City and the global reliability test was 0.835, which is higher than the previous work conducted in the entire metropolitan area. This paper therefore used local estimates of Cronbach's alpha to test whether the scale used in Holtan et.al. (2015) is preferable in Baltimore City or its surrounding areas compared to the one in Vemuri et.al (2011). The reliability coefficients of the two scales at the same spatial scale were subtracted and the results were interpolated by the inverse distance weighting method (IDW) for better visualization with minimal uncertainty. Furthermore, data from 2003 and 2011 were also tested to examine the solidity of the results (**Figure 2**).

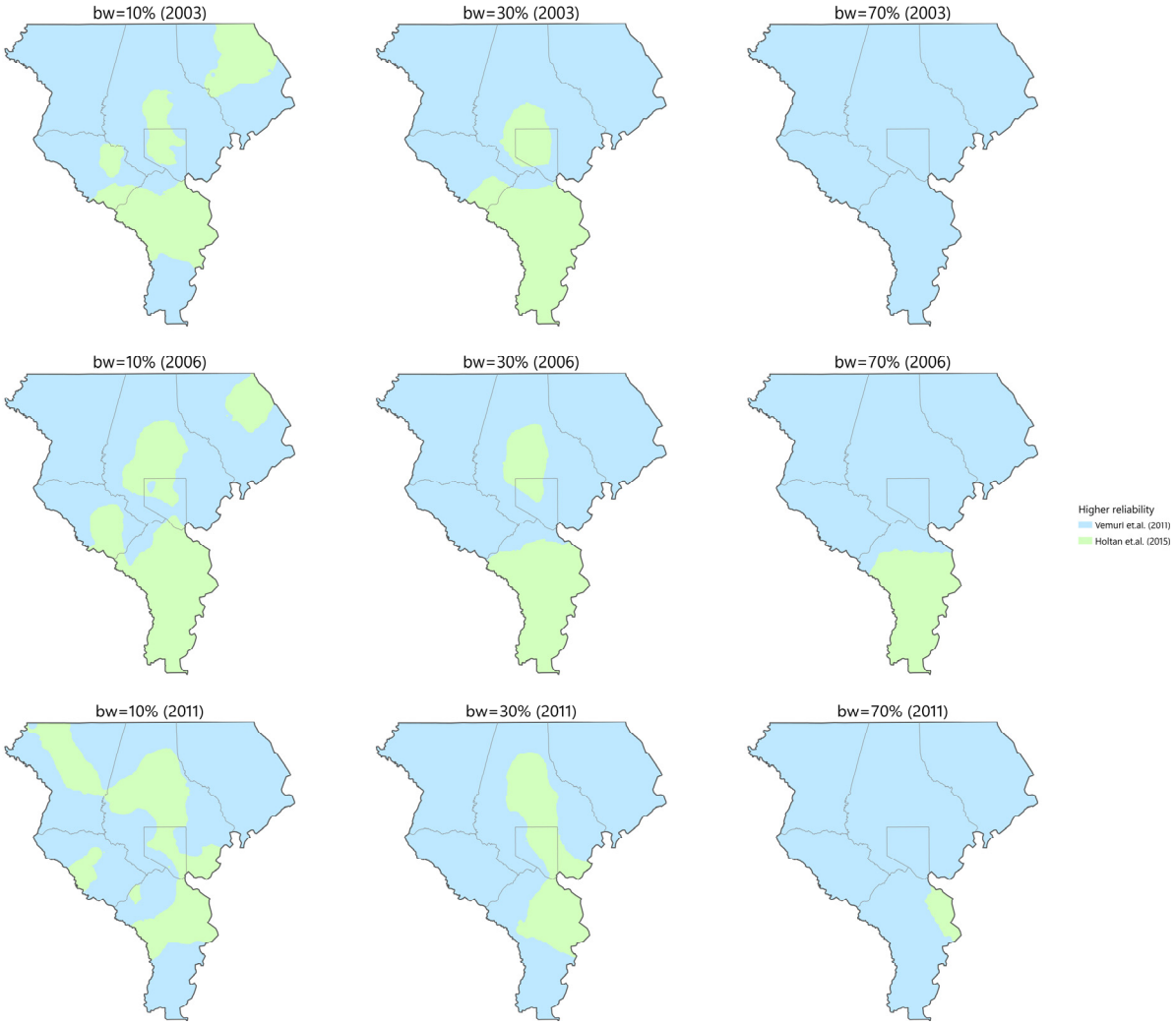

Figure 2 Comparison of local Cronbach's alpha estimates for the "social capital" scale in Vemuri et.al. (2011) and Holtan et.al. (2015).

When the bandwidth was set to 10% of responses, the spatial pattern of results was more variable, but the areas where the reliability of Holtan et.al (2015)'s scale was higher than that of Vemuri et.al (2011) were mainly distributed in Baltimore City and Anne Arundel County. In 2003 and 2006, such areas were found in the northeast, while in 2011 they shifted to the northwest. When the bandwidth is increased to 30%, the results in 2003 and 2006 are very similar both showing that the reliability of the 2015 version of the scale is higher in Baltimore City as well as in Anne Arundel County, which validates the results of Holtan et.al (2015) to a certain extent. In 2011 such areas moved northwards. When bandwidth is quasi-global, Vemuri's scale is more internally consistent in 2003 and 2011 almost globally, while in 2006 the 2015 version of the scale is still the one with more reliability in Anne Arundel County.

4. Discussion

This paper introduced a novel geographically weighted algorithm, Geographically Weighted Cronbach's alpha, to provide local estimates of the internal consistency of a test or scale. The paper also demonstrated in which scenarios and how this algorithm can be applied currently. Future work is expected to extend this algorithm in the following aspects: the method for calibrating the optimal bandwidth should be identified; the interpretability of local alphas will be further explored; and given that the alpha is indicated not to be the optimal estimate of "the true reliability" (Cronbach, 1951), other types of reliability coefficients will be extended to their local versions.

5. Acknowledgements

The demonstration data was supported by the Baltimore Ecosystem Study under National Science Foundation grant DEB-1637661.

References

Brunsdon, C., Fotheringham, A. S. and Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity, *Geographical analysis,* 28(4), 281-298.

Brunsdon, C., Fotheringham, S. and Charlton, M. (2007). Geographically weighted discriminant analysis, *Geographical Analysis,* 39(4), 376-396.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests, *Psychometrika,* 16(3), 297-334.

Grove, M. and Locke, D. (2018). BES Household Telephone Survey (Version 180), Environmental Data Initiative. Available at: https://doi.org/10.6073/pasta/5a4fc7bfa199f3d63748f0853ae073a0 (Accessed: 2023-01).

Harris, P., Brunsdon, C. and Charlton, M. (2011). Geographically weighted principal components analysis, *International Journal of Geographical Information Science,* 25(10), 1717-1736.

Holtan, M. T., Dieterlen, S. L. and Sullivan, W. C. (2015). Social life under cover: tree canopy and social capital in Baltimore, Maryland, *Environment and behavior,* 47(5), 502-525.

Vemuri, A. W., Morgan Grove, J., Wilson, M. A., et al. (2011). A tale of two scales: Evaluating the relationship among life satisfaction, social capital, income, and the natural environment at individual and neighborhood levels in metropolitan Baltimore, *Environment and Behavior,* 43(1), 3-25.

Biographies

Sui Zhang is a postgraduate student in the School of Social and Political Sciences at the University of Glasgow. His research interests mainly focus on spatial statistical methods and their applications in

public health.

Ziqi Li is a Lecturer in GIScience at the University of Glasgow. His research interests broadly include spatial analysis and modelling, spatial statistical learning, interpretable machine learning, and their applications in multidisciplinary fields.