Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lu, Y., Murakami, D., Nakaya, T., Wang, Y. and Harris, P. 2021. A Route Map for Successful Applications of Geographically Weighted Regression. *Geographical Analysis* . https://doi.org/10.1111/gean.12316

geographical analysis

Geographical Analysis (2022) **0**, 1–24

A Route Map for Successful Applications of Geographically Weighted Regression

Alexis Comber[1], Christopher Brunsdon[2], Martin Charlton[2], Guanpeng Dong[3], Richard Harris[4], Binbin Lu[5], Yihe Lü[6], Daisuke Murakami[7], Tomoki Nakaya[8], Yunqiang Wang[9], Paul Harris[10]

[1]School of Geography, Leeds University, Leeds, UK, [2]National Centre for Geocomputation, Maynooth University, Maynooth, Ireland, [3]Key Research Institute of Yellow River Civilization and Sustainable Development, Henan University, Kaifeng, China, [4]School of Geographical Sciences, University of Bristol, Bristol, UK, [5]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, [6]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Joint Center for Global Change Studies, Chinese Academy of Sciences, Beijing, China, [7]Department of Statistical Data Science, Institute of Statistical Mathematics, Tachikawa, Japan, [8]Graduate School of Environmental Studies, Tohoku University, Sendai, Japan, [9]State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an, China, [10]Sustainable Agriculture Sciences, North Wyke, Rothamsted Research, Okehampton, UK

Geographically Weighted Regression (GWR) is increasingly used in spatial analyses of social and environmental data. It allows spatial heterogeneities in processes and relationships to be investigated through a series of local regression models rather than a single global one. Standard GWR assumes that relationships between the response and predictor variables operate at the same spatial scale, which is frequently not the case. To address this, several GWR variants have been proposed. This paper describes a route map to decide whether to use a GWR model or not, and if so which of three core variants to apply: a standard GWR, a mixed GWR or a multiscale GWR (MS-GWR). The route map comprises 3 primary steps that should always be undertaken: (1) a basic linear regression, (2) a MS-GWR, and (3) investigations of the results of these in order to decide whether to use a GWR approach, and if so for determining the appropriate GWR variant. The paper also highlights the importance of investigating a number of secondary issues at global and local scales including collinearity, the influence of outliers, and dependent error terms. Code and data for the case study used to illustrate the route map are provided.

Correspondence: Alexis Comber, School of Geography, Leeds University, Woodhouse Lane, Leeds, LS2 9JT, UK
e-mail: a.comber@leeds.ac.uk
Binbin Lu, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei Province, 430072, China
e-mail: binbinlu@whu.edu.cn

Submitted: 22 June 2021; Revised version accepted: 29 November 2021

doi: 10.1111/gean.12316

Introduction

This article provides guidance for successful applications of geographically weighted regression (GWR), a method of spatial analysis first proposed by Brunsdon and Fotheringham (1996). The aim here is to guide users in how to use GWR, including their choice of GWR model and whether that model is appropriate for their study and data. With a wide range of model options now available, the article provides novice as well as more experienced users with a 'route map' to guide their analysis.

GWR (Brunsdon, Fotheringham, and Charlton 1996; Fotheringham and Brunsdon 2002) investigates if and how relationships between response and predictor variables vary geographically. It is underpinned by the idea that *whole map* (constant-coefficient) regressions such as those estimated by ordinary least squares may make unreasonable assumptions about the stationarity of the regression coefficients under investigation (Openshaw 1996; Fotheringham and Brunsdon 1999)—wrongly assuming that regression relationships are the same no matter where you measure them within the study region. GWR provides measures of process heterogeneity—geographical variation in data relationships—through the generation of mappable and varying regression coefficients, and associated statistical inference. It has been extensively applied in a wide variety of scientific and socio-scientific disciplines, such as environmental health (e.g., Yoneoka and Saito 2016), landscape ecology (e.g., Zhang, Bi, and Cheng 2004), soil quality (e.g., Song et al. 2016), air quality (e.g., You et al. 2015), water quality (e.g., Sun, Guo, and Liu 2014), remote sensing (e.g., Foody 2003), disease patterns (e.g., Brunton et al. 2017), urban studies (e.g., Huang and Yuan 2019), and housing markets (e.g., Yu and Wei 2007).

Linear regression, standard, mixed, and multiscale GWR

Various forms of GWR models are referred to in this article. It is useful first to describe four models that are considered of *primary* importance to a GWR study. These build on the basic linear regression model (LRM), which can be defined as:

$$y_i = \beta_0 + \sum_{k=1}^{m} \beta_k x_{ik} + e_i \tag{1}$$

where for observations indexed by $i = 1 \ldots n$, y_i is the response variable, x_{ik} is the value of the kth predictor variable, m is the number of predictor variables, β_0 is the intercept term, β_k is the regression coefficient for the kth predictor variable, and e_i is the random error term, that is, independently normally distributed with zero mean and variance σ^2. Ordinary least squares (OLS) are commonly used for model estimation in LRMs. Note that the model contains no reference to geography: The relationship between the x and y variables is assumed to be the same, everywhere, with only random (residual) departures from it at any particular location.

Standard GWR is similar to linear regression but calibrates the regression model at each predefined location (u, v)—geographical locations within the study region either from the sampled data or, for example, a grid of locations—using other nearby data falling within a moving window or kernel at the center of each discrete location. Standard GWR can be defined as:

$$y_i = \beta_0 \left(u_i, v_i\right) + \sum_{k=1}^{m} \beta_k \left(u_i, v_i\right) x_{ik} + e_i \tag{2}$$

where (u_i, v_i) are the spatial coordinates of the observations i, and $\beta_k(u_i, v_i)$ are the coefficients estimated at those locations. Thus, in contrast to *global* LRMs, GWR conducts *local* regression at a series of locations to estimate local coefficients (the *geographical* part of GWR), using observations weighted by their distances to the location at the center of the moving window/kernel (the *weighted* part). As with the LRM, the set of e_i obey an independent normal distribution with zero mean and variance σ^2. Equations for calculating the local coefficient standard errors for GWR can be found in Fotheringham, Brunsdon, and Charlton (2002) and Harris, Fotheringham, and Juggins (2010). Critically, GWR creates many local models at locations within the study region. This enables the coefficients of those models to be compared to see if the regression relationships vary spatially.

The weightings in GWR are determined by a kernel-based distance decay function and its bandwidth. Bandwidth can be a fixed distance or a fixed number of nearest data points (i.e., an adaptive radius depending on the local density of points). Automated routines exist to determine an optimal bandwidth by minimizing some measure of model performance such as the Akaike information criterion (AIC) and its corrected version (AICc) (Fotheringham, Brunsdon, and Charlton 2002) or a leave-one-out cross-validation (CV) score (Brunsdon, Fotheringham, and Charlton 1996). The result of larger bandwidths is that the GWR model tends toward the OLS estimator of the *whole map* LRM.

In the standard form, a single bandwidth is used in GWR under the assumption that the response-to-predictor relationships operate over the same scales for all of the variables contained in the model. This may be unrealistic because some relationships can operate at larger scales and others at smaller ones. A standard GWR will nullify these differences and find a “best-on-average” scale of relationship non-stationarity (geographical variation).

A mixed (or semiparametric) GWR (MX-GWR) (Brunsdon and Fotheringham 1999; Mei and Xu 2016) was proposed to allow for a mix of local (non-stationary) and global (stationary) relationships between predictor and response variables. However, MX-GWR only in part addresses the limitation of standard GWR, as the set of locally varying relationships are all assumed to operate at the same spatial scale as each other. In other words, a single local bandwidth is applied to them all.

To fully address this, multiscale GWR (MS-GWR) (Yang 2014; Fotheringham and Yang 2017; Oshan et al. 2019) can be used. In this, the bandwidth for each relationship is determined separately, allowing the scale of individual response-to-predictor relationships to vary. Useful comparisons between MS-GWR and alternative multiscale spatially varying coefficient frameworks can be found in Wolf and Oshan (2018) and Murakami et al. (2019). LRM, standard GWR, and MX-GWR can all be considered as special cases of MS-GWR, with increasing complexity (i.e., flexibility in the specification of the spatial relationships) moving from LRM to MS-GWR.

The article considers a Gaussian response case in its implementations of GWR, MX-GWR, and MS-GWR. A bi-square weighting kernel is used (see, *inter alia*, Gollini et al. 2015) where a single bandwidth b is found for standard GWR and also for the pre-specified local or non-stationary relationships in MX-GWR, whereas $m + 1$ bandwidths are found for MS-GWR. All bandwidths are optimized by minimizing the AICc.

Motivation

The motivation for this article is because GWR increasingly is used for different spatial analyses across a range of disciplines, with sharp increases in the number of applications in recent years. This proliferation has been driven by a number of factors. First is the increasingly spatial nature of data, which are now routinely collected with location attached, facilitated by the many GPS-enabled monitoring devices and the tagging of, for example, administrative data with census geographies. Second, there is a broader cross-disciplinary demand outside of geography for methods to quantify spatial patterns in data, commonly through some kind of hotspot estimation, spatial cluster analysis, or spatially informed regression technique. This has been accompanied by recognition of the need to cater for spatial dependencies in the data or the model parameters themselves, reflecting Tobler's first law of geography that observes how data measured in close proximity to each other tend to display similar characteristics (Tobler 1970). GWR is a method that enables this and builds on the simple LRM with which many students and researchers are familiar. Third, it has been implemented in a number of software packages including the ESRI ArcGIS suite of tools, five R packages (*spgwr* (Bivand, Yu, and Nakaya 2013), *gwrr* (Wheeler 2013), *GWmodel* (Lu, Harris, and Charlton 2014; Gollini et al. 2015), *McSpatial* (McMillen 2013), and *lctools* (Kalogirou 2019)), one Python implementation (the *mgwr* package (Oshan et al. 2019) in the PySAL project (Rey and Anselin 2010)) and standalone implementations such as *GWR3* (Charlton and Fotheringham 2003), *GWR4* (Nakaya 2015) and *MGWR 1.0* (Li, Fotheringham, and Li 2019). Each software package has a standard GWR option complemented by a variety of alternative GWR forms and associated tools.

A consequence of this proliferation is the danger that new users of GWR do not adequately consider whether the GWR form they chose is appropriate for their application. With that in mind, this article aims to provide a route map to help inform best practices and decision-making.

A GWR route map

The GWR route map is described using a soil case study in the Loess Plateau of China. It seeks to guide the reader through different modeling scenarios that are of *primary* importance to a GWR analysis. These main arteries of the route map take the reader to *GWR Basecamp*. Strategies for *secondary* model considerations and decisions (*scaling the summit*) are described in the Discussion section. Not all *secondary* issues may appear in a specific GWR analysis, and some may interact, including interactions with those considered of *primary* importance. Although the GWR route map is presented as a linear workflow, it should be recognized that, in practice, it is often an iterative process, as may be the case in any regression study. The implications in this respect are that, for some spatial processes, a GWR analysis can be a relatively straightforward decision, while for others, decisions can be problematic and complex. Ultimately, the result of this two-stage *primary-to-secondary* strategy should lead to an informed, sensible, and appropriate GWR implementation, from which reliable and robust inferences can be made about regression relationships and their spatially varying nature.

Further considerations and guidance are given in the appendices of the pre-print of this article (Comber et al. 2020) with respect to: Sample data characteristics, influences on the weighting schemes, inference in GWR and alternative spatially varying coefficient models, GWR as a spatial predictor and GWR development through simulation experiments. These cover many important issues that are not fully covered here in order to provide more focus to the exposition. The route map is presented using real rather than simulated data. This is deliberate, as the intention is to provide 'real world' practical guidance to a GWR analysis. Although this article focuses on a GWR application in the environmental sciences where data are commonly measured on a point basis, the main messages of the article are similarly relevant to social applications using data measured for areas such as census geographies and neighborhoods that can be represented by a point centroid.

Case study

Data

The case study consists of a single soil dataset of 689 observations, spaced at approximately 100 m intervals in a small watershed in the Loess Plateau, China (110.32821 E and 38.83433 N). The data locations are shown in Figure 1 and described in Wang and Zhang (2009) who undertook a linear regression analysis complemented with a geostatistical analysis of the data. The data are also described in Comber et al. (2018). They include soil total nitrogen (*STN*), here taken as the response variable, and six predictor variables; soil organic carbon (*SOCgkg*), nitrate-nitrogen (*NO3Ngkg*), ammonium (*NH4Ngkg*), and percentage clay (*ClayPC*), silt (*SiltPC*), sand (*SandPC*) content. In both Wang, Zhang, and Huang (2009) and Comber et al. (2018), the data were transformed, and this operation is retained here: *STN*, *SOCgkg*, *NO3Ngkg,* and *NH4Ngkg* are transformed using natural logs and *ClayPC* is square root transformed. As with any regression analysis, due consideration should be given to the nature of data relationships, the use of data transforms, and associated model specification tasks prior to the main model fits.

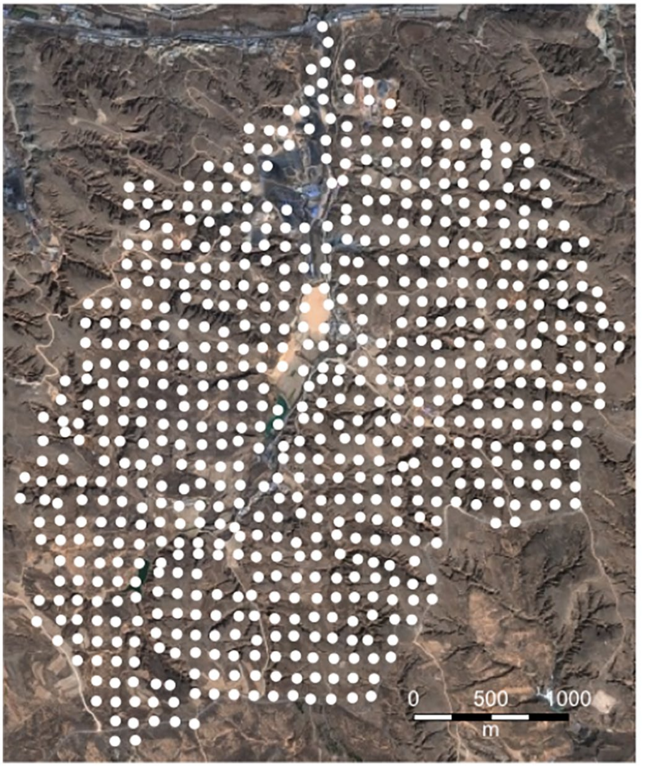

Figure 1. The case study data locations.

Each analysis in the GWR route map below predicts *STN* using different predictor variable subsets to illustrate specific points. At no point is the intention to conduct a nuanced regression analysis that attempts to fully characterize and interpret the soil processes. Rather the different data set scenarios are used to illustrate the route map. For reproducibility (Brunsdon and Comber 2020), the data set and the R code used to undertake the analyses are available from https://github.com/lexcomber/GWRroutemap.

Scenarios

Four data set scenarios were chosen to illustrate the route map decisions. These are given in Table 1, each with *STN* as the response but with different predictors. The compositional nature of the clay/sand/silt data is catered for by omitting at least one from each analysis. Importantly, the intention is to treat each scenario as a distinct and independent data set and not as a linked model specification exercise with respect to predictive variable selection. In this respect, "Analysts" are assigned to each data set, where Analysts B-D are entirely unaware that more predictors of *STN* exist. As such, model performance statistics (here AICc) are only compared for those models relating to each Analyst's scenario, and not across different analysts.

Primary model decisions

The fundamental consideration for undertaking a GWR analysis is that it should be justified in terms of the aims of the analysis and the characteristics of the data. If spatial effects are evident in the data then a standard GWR can be considered but this requires demonstrating that alternate models are not suitable. This is discussed further, below. To achieve this, the following steps for any GWR analysis are recommended.

1. A LRM should be fitted and the results investigated.
2. A MS-GWR should be calibrated and the estimated bandwidths interrogated.
3. Based on findings (1) and (2), one from a standard GWR, MX-GWR, and MS-GWR should be considered for further analysis *provided* a spatially varying coefficient model such as GWR is considered suitable in the first place.

The LRM assumes fixed data relationships and provides the baseline against which all forms of GWR can be compared. The MS-GWR model estimates the bandwidths for each predictor-to-response relationship. Evaluating these directly quantifies the nature of any spatially varying relationships and at what spatial scale they each operate at. This in turn informs whether to pursue a GWR analysis and if so, which of three different GWR forms to follow. That is, given the MS-GWR results, can a simpler model in an LRM, standard GWR, or MX-GWR provide a viable and pragmatic alternative? Or is MS-GWR the only viable option?

Table 1. Data Set Scenarios in Terms of Four Different "Analysts"

	SOCgkg	*ClayPC*	*SiltPC*	*SandPC*	*NO3Ngkg*	*NH4Ngkg*
Analyst A	Yes	Yes	Yes	–	Yes	Yes
Analyst B	–	–	–	Yes	Yes	–
Analyst C	Yes	–	–	–	–	Yes
Analyst D	Yes	–	–	Yes	Yes	–

This approach to *primary* model choice is recommended first because the theory for the standard LRM is well developed, while theoretical developments reduce from standard GWR, to MX-GWR, and to MS-GWR. For example, robust versions exist for standard GWR, but not for MX-GWR or MS-GWR, while autoregressive versions exist for standard GWR and MX-GWR but not for MS-GWR (see *secondary* model decisions section). Second, other considerations of model complexity, sample size, sample configuration, and sample variation play key and intertwined roles, which cannot be entirely resolved through a comparison of a model performance statistic such as AICc. Sometimes, choosing a more basic regression over MS-GWR is to be preferred, even when AICc values suggest otherwise, but where this decision is informed by following the proposed route map. Arguments for not relying on information theory alone for model choice can be found, for example, in Guthery, Brennan, and Peterson (2005) in the context of wildlife science.

Both the LRM and MS-GWR analysis should also be investigated for the presence of spatially autocorrelated model residuals, say through Moran's I (Moran 1950; Cliff and Ord 1973). Thus, further to the four model choices (of LRM, GWR, MX-GWR, and MS-GWR), a fifth model is considered where an alternative fixed coefficient regression is fitted but with a spatially autocorrelated error term (i.e., a spatially autocorrelated model, SAM). For this study, the spatially autocorrelated error term is modeled by the parameterization of its covariance using an exponential function decaying with respect to the Euclidean distance separating sample sites. The restricted maximum likelihood (REML) method (e.g., Lark and Cullis 2006) is used for the estimation. The SAM will warrant consideration depending on the nature of spatially autocorrelated residuals from the LRM fit and also if the MS-GWR fit indicates that only the intercept is found to be spatially varying (Nakaya, Fotheringham, and Brunsdon 2005). The theory for the SAM and related models is also well developed (e.g., Waller and Gotway 2004; Schabenberger and Gotway 2005), where for this study the SAM is a linear mixed model.

The kernel bandwidth identification is *the* critical consideration in GWR as it determines how many data points are included in the data subset passed to each local regression and how these data points are spatially weighted. Bandwidths describe the scale of the predictor-to-response relationships. They dictate the degree of smoothing or variation in the local regression coefficient estimates, and interpretation and inference around process heterogeneity (geographical variation in data relationships) thereafter.

Determining the scale at which data relationships operate is not a straightforward task. In this study, bandwidths are found objectively via AICc, but this should not discount user-specified bandwidths when there exists some strong prior belief, theoretical justification, or expert knowledge for their use. Similar discussions can be found in related kernel weighting paradigms, such as kernel density estimation, where automated bandwidth approaches are not necessarily viewed as a panacea for bandwidth selection (Silverman 1986). There are strong benefits in conducting an extensive bandwidth investigation so that final model outputs can be more assured. Bandwidths found by AICc are preferred to those found by CV and uncorrected AIC because it reflects model parsimony adjusted for small degrees of freedom and its use tends to avoid overfitting GWR models (AICc bandwidths tend to be larger than bandwidths found using CV and AIC). AICc is similarly found for the LRMs and the SAMs enabling objective comparison with the GWR models. For further guidance, but without reflecting model parsimony, R^2 values are reported for all models.

For the *primary* analyses, only rudimentary assessments of statistical (relationship) significance are undertaken using coefficient standard errors, t-values, and p-values from standard GWR, MX-GWR, and MS-GWR models. Caveats on their use with all forms of GWR are discussed in the appendices in Comber et al. (2020). Global assessments from the LRM and SAM fits are also reported, where for all study models only the coefficient estimates and their p-values are given.

Step 1: LRM and autocorrelated residuals

The first step is to undertake a global regression (i.e., the LRM). The aim here is to try to understand how the predictors relate to the response variable. Specifically: (a) which relationships are statistically *significant*, (b) whether there is evidence for specifying an *autocorrelated error* term, and (c) *the fit* of the LRM itself. Table 2 summarizes the LRM coefficient estimates and their significance from zero, for all four Analysts. The LRMs from Analysts A and C provide a mixture of significant and insignificant predictors of *STN* at the 5% level, while all predictors are significant for the LRMs from Analysts B and D.

To assess spatial autocorrelation of the LRM residuals, a spatial weight matrix was defined and unbiased estimates of Moran's I and their significance were determined (Table 3), under the expectation of random and independent residual distributions. Moran's I for all four models are significant, where the spatial structure in the LRM residuals varies from relatively weak (Analyst A and D) to relatively strong (Analyst C), as reported by the magnitude of the estimates. In this case, all four data set scenarios indicate that a fixed coefficient regression with a spatially autocorrelated error term could be suitable (i.e., a SAM via REML estimation). Table 3 also provides summaries of LRM fit for the four scenarios with AICc and R^2 values. To re-iterate an early comment, the purpose of providing these measures of model fit is not to compare across the analysts but to use them as a benchmark for model improvement for each analyst.

For all four scenarios, Table 3 indicates that a GWR analysis may be appropriate as the Moran's I measures are moderately positive and statistically significant. The existence of autocorrelated residuals from a LRM fit commonly *suggests* that a GWR analysis may be useful but Moran's I of residuals gives no indication of the presence of spatially varying relationships between the response and predictor variables. Determining this is important for understanding spatial regression modeling in general but separating autocorrelation effects in the residual term from relationship heterogeneity effects is difficult. Useful discussions on this can be found in Harris (2019) and references therein.

Table 2. LRM Coefficient Estimates and their Significance (P-value)

	Analyst A		Analyst B		Analyst C		Analyst D	
	Estimate	***P*-value**	Estimate	***P*-value**	Estimate	***P*-value**	Estimate	***P*-value**
Intercept	2.220	0.000	0.723	0.000	2.130	0.000	1.437	0.000
SOCgkg	0.690	0.000	–	–	0.918	0.000	0.683	0.000
ClayPC	0.011	0.843	–	–	–	–	–	–
SiltPC	0.015	0.000	–	–	–	–	–	–
SandPC	–	–	0.021	0.000	–	–	0.012	0.000
NO3Ngkg	0.126	0.000	0.355	0.000	–	–	0.112	0.000
NH4Ngkg	0.146	0.047	–	–	0.011	0.884	–	–

Table 3. Residual Autocorrelation Measures Using Moran's *I* (Estimate and *P*-value) and Performance Statistics (AICc and R^2) from the Four LRMs

	Moran's ***I***	***P***-value	AICc	R^2
Analyst A	0.142	0.000	1,124.0	0.609
Analyst B	0.174	0.000	1,377.4	0.430
Analyst C	0.219	0.000	1,223.1	0.545
Analyst D	0.144	0.000	1,131.0	0.603

Table 4. The Fixed Bandwidths in Meters (max = 3,742 m) for Different Models Arising from an MS-GWR

	Intercept	***SOCgkg***	***ClayPC***	***SiltPC***	***SandPC***	***NO3Ngkg***	***NH4Ngkg***
Analyst A	555.9	2,483.9	3,741.7	1,080.8	–	382.5	3,741.7
Analyst A*	57	631	685	306	–	55	685
Analyst B	445.8	–	–	–	1,232.9	731.9	–
Analyst C	424.9	3,741.4	–	–	–	–	3,741.8
Analyst D	573.6	2,214.6	–	–	1,066.5	378.4	–

Note For Analyst A, * indicates an adaptive bandwidth (max = 689 observations).

In general, but not a rule, measures of strong model fit (e.g., $R^2 > 0.8$), coupled with weak or insignificant levels of spatial autocorrelation in the residuals, suggest that a LRM would be appropriate. Such cases suggest that the model is well specified and includes all of the likely factors driving the response variable. If the fit is poor and exhibits significant levels of residual spatial autocorrelation, then this might suggest that the model is mis-specified, perhaps missing key factors driving the response variable. Although the better specification of the model should be pursued, a GWR analysis is still an option, as is a SAM. This is returned to in the Discussion section.

In summary, this first step fits a LRM to identify which relationships are globally significant and to determine whether spatial autocorrelation effects may potentially exert an important influence on these findings. Model performance statistics (AICc and R^2) enable comparison between SAM and GWR models.

Step 2: MS-GWR and bandwidth estimation

The second step is to undertake an MS-GWR analysis. The MS-GWR bandwidths explicitly describe the degree of spatial heterogeneity associated with each variable's relationship to the response. This provides information about the different scales of predictor-to-response relationships, where some may be local and others global, and those that are local may have different scale effects from one another. The focus of the MS-GWR analysis at this stage is to examine (i) the *estimated bandwidths* (ii) evidence for *residual autocorrelation*, and (iii) *the fit* of MS-GWR itself.

The estimated MS-GWR fixed distance bandwidths are shown in Table 4, with adaptive distance bandwidths illustrated for the MS-GWR model of Analyst A only. In this study, the maximum number of data points that can be included under an adaptive bandwidth is 689 (the total number of observations in the data) and the maximum fixed bandwidth is 3,742 m (the maximum distance between any pair of data points). The bandwidths in Table 4 should be interpreted in light of these values. For Analyst A, the relative bandwidth sizes are consistent between fixed and adaptive forms as was the case for the other three data set scenarios. This similarity is re-assuring but to a certain extent reflects that the study data were sampled on a loosely regular grid. Studies with data over an irregular sample configuration may need to experiment more in this respect.

The fixed bandwidths show clear patterns for each predictor variable and the scale of its spatially varying relationship to the response, *STN*.

For Analyst A, the MS-GWR bandwidths for *ClayPC* and *NH4Ngkg* both strongly tend towards the maximum, global bandwidth of 3,742 m, while *SOCgkg* and *SiltPC* have bandwidths of about two-thirds and one-third of the global one, respectively. The bandwidths for the *intercept* and *NO3Ngkg* for Analyst A are both strongly local.

For Analyst B, the bandwidths for the *intercept, SandPC and NO3Ngkg* are all local.

For Analyst C, the bandwidths for *SOCgkg* and *NH4Ngkg* are essentially global, while the *intercept* is local.

For Analyst D, none of the bandwidths are global and those for the *intercept, SOCgkg, SandPC, and NO3Ngkg* vary locally, but appear quite different in magnitude.

To assess model residual spatial autocorrelation for the MS-GWR fits, estimates of Moran's *I* and their significance are given in Table 5 along with MS-GWR performance statistics. These can be compared with the corresponding results in Table 3 for the LRMs. Clearly, in all scenarios, residual autocorrelation is now negligible, while model performance improves (lower AICc and higher R^2) over the corresponding linear regressions.

Step 3: Choice of the *primary* model

The results of the initial LRM and MS-GWR analyses can guide *primary* model choice. First, from the LRM analysis in Step 1, it appears that for all four Analysts, a fixed coefficient model could be considered but only if calibrated with an autocorrelated error term, (i.e., SAM fits). Second, from the MS-GWR analysis in Step 2, some form of GWR is similarly worth considering, for all Analysts. This is because residual autocorrelation disappears with a MS-GWR fit, while at least one predictor bandwidth in addition to the intercept is clearly locally varying.

The following sub-sections provide guidance on how deciding on the *primary* model should be undertaken, considering the five regression possibilities (LRM, SAM, standard GWR, MX-GWR, and MS-GWR) and the four data set scenarios. Critical to Step 3 is the presentation and interpretation of the estimated coefficients and associated uncertainties from competing models, and not just choosing a *primary* model by referring to model fit statistics alone.

Table 5. MS-GWR Residual Autocorrelation Measures Using Moran's *I* and Error Statistics

	Moran's ***I***	***P***-value	AICc	R^2
Analyst A	0.007	0.604	1,050.4	0.713
Analyst B	0.013	0.700	1,264.4	0.580
Analyst C	0.005	0.381	1,106.8	0.662
Analyst D	0.009	0.636	1,057.4	0.708

Table 6. Coefficient Estimates and their Significance Arising from LRM and SAM Fits for Analyst C

	LRM		SAM	
	Estimate	***P***-value	Estimate	***P***-value
Intercept	2.130	0.000	1.817	0.000
SOCgkg	0.918	0.000	0.816	0.000
NH4Ngkg	0.011	0.884	0.086	0.284

Investigating LRM and SAM for Analyst C

A LRM should be considered as a potential final model when all bandwidths from MS-GWR are large (i.e., tend toward the global situation), including the intercept. As a rule of thumb, this is when they are broadly greater than 80% of the maximum distance between data points (or 80% of the data points in the adaptive bandwidth case). In this respect, none of the Analysts have a data set that clearly suggests a LRM to be appropriate. However, from the above, it is stated that all analysts could consider a SAM fit (because all indicated autocorrelated residuals from their LRMs), and in this respect, a SAM can be further endorsed if all predictor variable bandwidths from MS-GWR tend to the global, but the *intercept* is local. This is clearly the case for Analyst C's data set (from Table 4).

Thus, in this instance, the *primary* route map has guided Analyst C to a SAM. It is prudent to compare SAM outputs to the LRM outputs because only the intercept term is locally varying from the MS-GWR. The coefficient summaries in Table 6 indicate only marginal gains in process interpretation with the SAM fit, despite the AICc improvement with the SAM (1,148.4 compared to 1,223.1 for the LRM). In this instance, there is only marginal improvements with the inclusion of second-order spatial effects via a SAM, as reflected by the broadly similar coefficient estimates. Thus, in summary, Analyst C could proceed with a fixed coefficient regression, where a LRM suffices, although choosing a SAM is also reasonable.

Note that Analyst C could have considered an MX-GWR with only the *intercept* locally varying, but as a rule, spatial effects via a SAM should be preferred due to its stronger inferential properties (e.g., see LeSage and Pace 2009). This is because inference in any GWR model is somewhat compromised by there being no-one single model, but a collection of models re-using sample data at multiple locations. This entails that a valid probability model is unavailable with GWR, making inference biased and problematic. This is not the case for a LRM or this study's SAM, and is similarly not the case for many alternative spatially varying coefficient models that are based on linear mixed model constructs (e.g., see Wolf, Oshan, and Fotheringham 2018; Murakami et al. 2019).

Investigating MX-GWR and MS-GWR for Analyst A

An MX-GWR can be explored when the MS-GWR analysis suggests two distinct sets of bandwidths, with one set tending to the global and the other set tending to a similar local scale. This scenario appears likely for Analyst A (from Table 4), where the MS-GWR bandwidths for *SOCgkg, ClayPC,* and *NH4Ngkg* can be viewed as global, while those for the *intercept*, *SiltPC,* and *NO3Ngkg* can be viewed as local.

To explore this further, an MX-GWR was applied with a user-specified local bandwidth of 700 m, as an approximate average of the local scales from Table 4, although optimal local bandwidth could have been determined with *GWR4* (Nakaya 2015). Figure 2 shows the spatial distribution

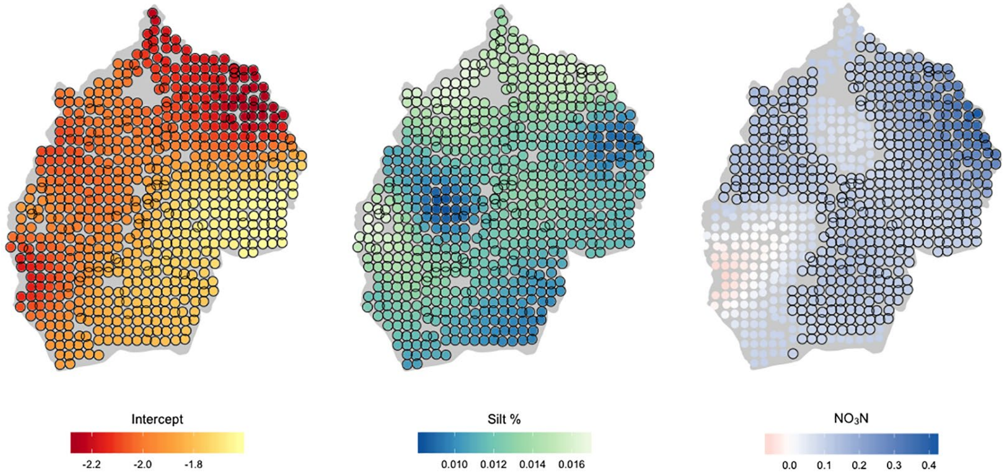

Figure 2. The spatial variation of the local coefficient estimates given with P-values < 0.05 highlighted from the MX-GWR analysis of Analyst A.

of the significant local coefficient estimates and the spatially varying relationships between *STN* and the *intercept*, *SiltPC* and *NO3Ngkg*. The global coefficient estimates for *SOCgkg ClayPC* and *NH4Ngkg* from the MX-GWR fit were 0.677, 0.016, and 0.193, respectively, similar to the coefficient estimates in Table 2, with *ClayPC*, again not found to be significant. The MS-GWR coefficient estimates should also be mapped for comparison and are given in Figure 3. The AICc fit of the MX-GWR model was 1,065.9, worse than that found for the MS-GWR (1,050.4).

Here, information from the LRM (Table 2), MX-GWR (Figure 2), and MS-GWR (Table 4, Figure 3) need to be jointly considered to fully interpret the nature of the relationships in Analyst's A data set. On balance, *STN*'s relationships with *SOCgkg*, *ClayPC,* and *NH4Ngkg* are clearly global and constant across space (Table 4), where *STN*'s relationships with *ClayPC* and *NH4Ngkg* are not viewed as significant (Table 2) and noting that the *NH4Ngkg* relationship to *STN* is borderline significant/insignificant in all fits (LRM, MX-GWR, and MS-GWR). Conversely, *STN*'s relationship with the *intercept*, *SiltPC,* and *NO3Ngkg* are local, where the local behavior varies little between the MX-GWR and MS-GWR forms. Only for *NO3Ngkg* do differences occur, where more distinct and significant areas of negative coefficient estimates were generated with MS-GWR, but not seen in MX-GWR. If the differences were more pronounced, then Analyst A should consider re-specifying the MS-GWR model with bandwidths for *SOCgkg*, *ClayPC,* and *NH4Ngkg* as global, while those for the *intercept*, *SiltPC,* and *NO3Ngkg* are re-estimated so that each relationship varies at its own local scale. However, given the similarity in the coefficient distributions, Analyst A could justifiably and pragmatically proceed with a MX-GWR fit, even with its marginally worse AICc.

Investigating MS-GWR only for Analyst D

The MS-GWR fit should be retained when the variable-specific bandwidths clearly suggest each data relationship is operating at its own unique spatial scale, as in the case of the data set for Analyst D (see Table 4). Figure 4 maps the distribution of the local coefficient estimates. Here, only the relationship for *NO3Ngkg* with *STN* changes in sign, and is the only relationship that varies between significant and insignificant in different locations.

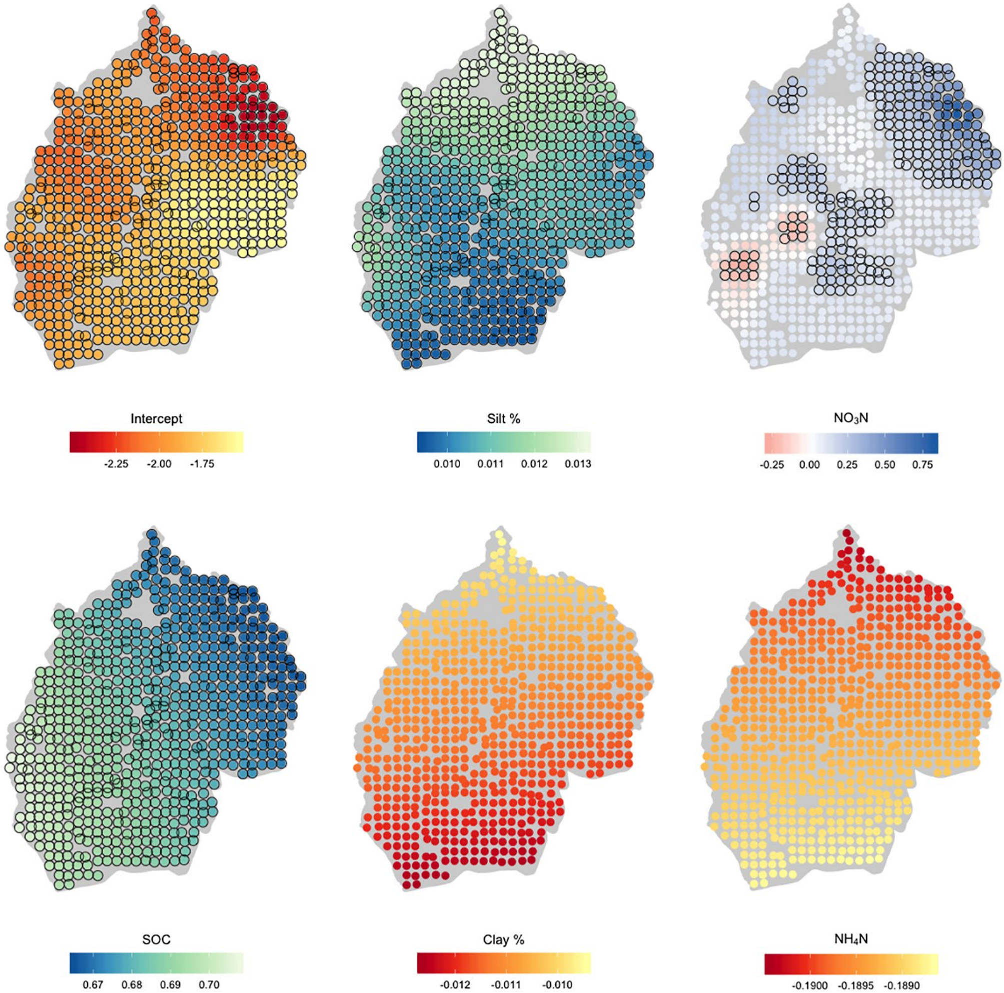

Figure 3. The spatial variation of the local coefficient estimates given with P-values < 0.05 highlighted from the MS-GWR analysis of Analyst A.

Investigating standard GWR and MS-GWR for Analyst B

A standard GWR is generally not an adequate model. It can be chosen over an MS-GWR only on the rare occasions when the intercept and all predictors have broadly similar MS-GWR estimated bandwidths, which is potentially found for Analyst B (Table 4). This scenario predicts *STN* using just *SandPC* and *NO3Ngkg*, for which a single local bandwidth appears reasonable. In this instance, the single bandwidth can be optimally determined through a standard GWR calibration, where it was found via AICc to be 597.5 m.

Where possible, the bandwidth function in a standard GWR should be investigated. This confirms that the bandwidth optimization search has not settled on a local minimum. Figure 5 shows the bandwidth function for an AICc minimization, which is well-behaved with a clear minimum. Note that if the bandwidth function was very shallow and plateaued, then a LRM would likely suffice. Also, small bandwidths (say, <2% of the data points when using an

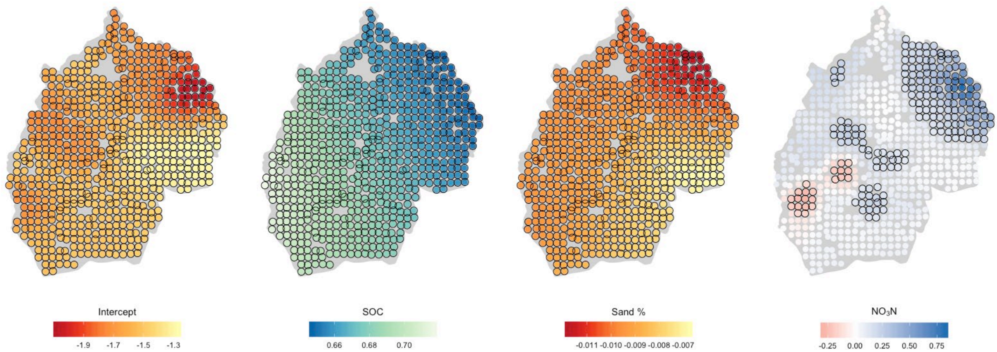

Figure 4. The spatial variation of the local coefficient estimates given with *P*-values < 0.05 highlighted from the MS-GWR analysis of Analyst D.

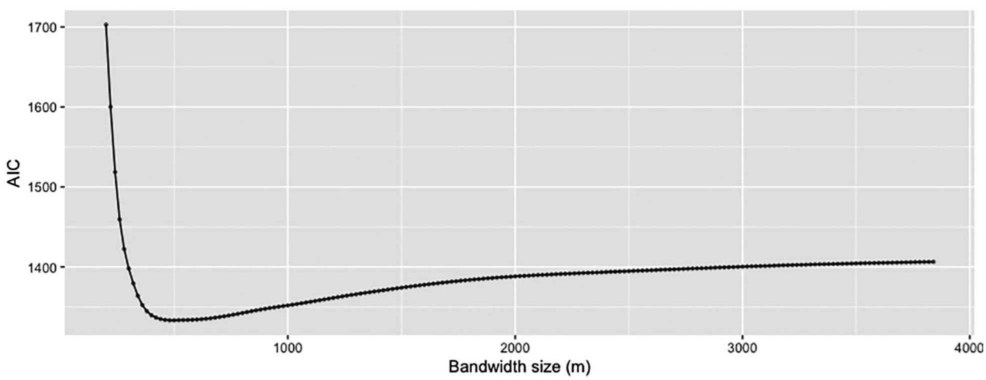

Figure 5. The bandwidth optimization search for standard GWR.

adaptive bandwidth) are indicative of over-fitting (to random perturbations, for example), and that a standard GWR is suggesting geographical patterns when none exists. In such cases, the GWR analysis should cease.

Figure 6 maps the distribution of the local coefficient estimates from standard GWR. Here, the relationships for the *intercept* and *SandPC* with *STN* can change in sign. Again, the MS-GWR coefficient estimates are mapped for comparison (Figure 7), indicating clear spatial differences between standard GWR and MS-GWR coefficients. In general, MS-GWR indicates smaller ranges of coefficient variation (as would be expected since the estimated bandwidths for *SandPC* and *NO3Ngkg* are much larger with MS-GWR), but regression relationships are consistently significant across space. Thus, given these differences and that the AICc for standard GWR is 1,272.3 compared to that found with MS-GWR (1,264.4), here, it is prudent to retain the MS-GWR model rather than simplifying the analysis with standard GWR.

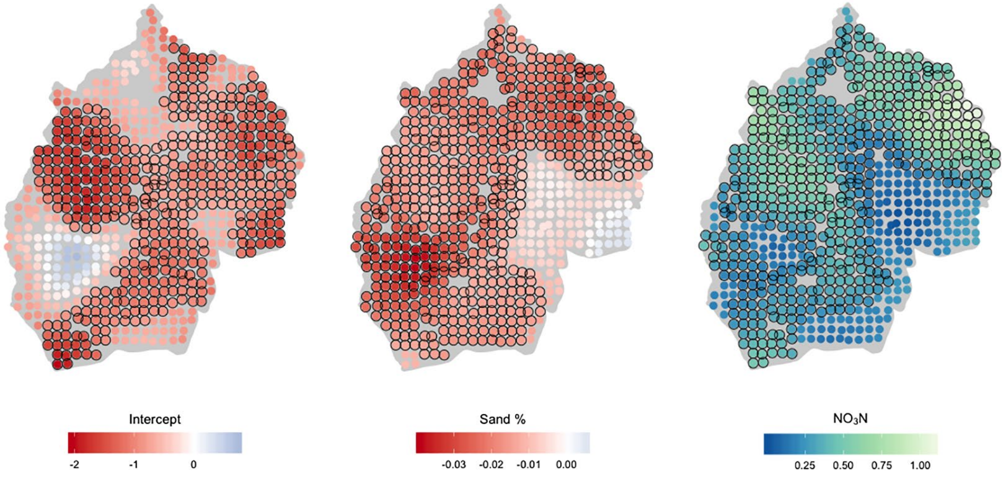

Figure 6. The spatial variation of the local coefficient estimates given with *P*-values < 0.05 highlighted from the standard GWR analysis of Analyst B.

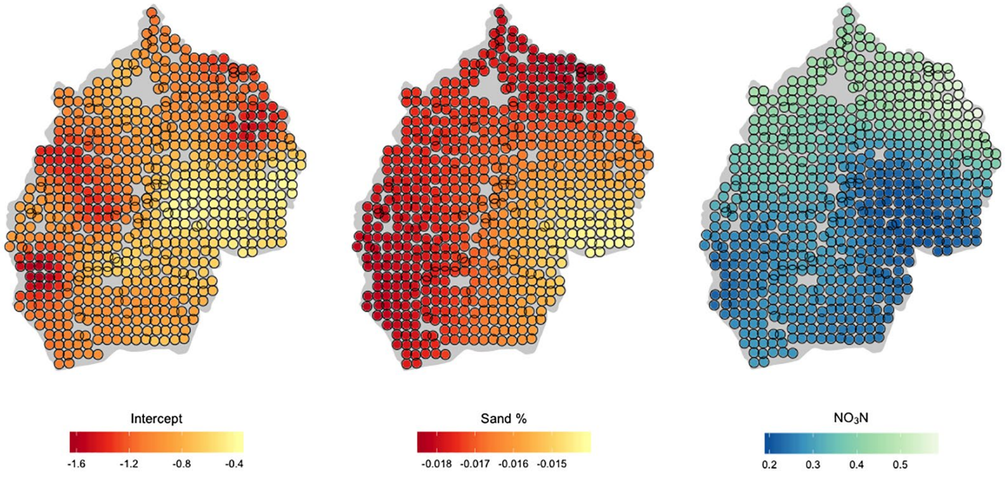

Figure 7. The spatial variation of the local coefficient estimates given with *P*-values < 0.05 highlighted from the MS-GWR analysis of Analyst B.

Summary in terms of AICc

Table 7 summarizes the AICc results for each Analyst, where for all scenarios the MS-GWR model provides the most parsimonious fit in terms of AICc. The choice of any *primary* model is always an improvement in fit over the LRM (except when LRM itself is chosen) but does not necessarily provide an improvement in predictive performance over the corresponding MS-GWR model in terms of AICc. This is because: (a) The measures of relationship non-stationarity (via the coefficient maps, above) can sometimes remain broadly unaltered despite a poorer model performance (AICc) measure, as was the case when MX-GWR was chosen over MS-GWR for

Table 7. AICc Values Arising from the *Primary* Model Analyses

	LRM	MS-GWR	Chosen primary model
Analyst A	1,124.0	1,050.4	1,065.9 (MX-GWR)
Analyst B	1,377.4	1,264.4	1,264.4 (MS-GWR)
Analyst C	1,223.1	1,106.8	1,223.1 (LRM)
Analyst D	1,131.0	1,057.4	1,057.4 (MS-GWR)

Analyst A, or (b) relationships are in essence stationary as was the case for Analyst C when a LRM suffices.

Discussion

GWR model choice

The route map involves balancing evidence of residual autocorrelation with coefficient non-stationarity in deciding between GWR alternatives (standard, MX-GWR, MS-GWR), while also considering stationary coefficient models (LRM, SAM). It is evident from the above that there are few hard and fast rules about primary GWR model choice. Rather, choices are practical, based on an iterative process, where spatial structures in the data are considered alongside measures of model performance (AICc) and coefficient significance:

- Analyst C could have chosen to undertake a SAM rather than a LRM (and SAM approaches frequently do result in larger differences in coefficients compared to those of LRM) but in this case, the inferential gains in the regression relationships were marginal despite a strong improvement in AICc.
- Analyst A could have chosen a MS-GWR but the local differences in regression relationships with those of a MX-GWR were small, which were similarly reflected in the AICc measures in this case (1,065.9 for MX-GWR and 1,050.4 for MS-GWR).
- Analyst B experimented with a standard GWR based on the loose similarity of the MS-GWR bandwidths and the small differences in AICc between GWR forms (1,272.3 for standard GWR and 1,264.4 for MS-GWR), but standard GWR resulted in different local inferences where it identified fewer significant regression relationships (comparing Figures 6 and 7) and thus MS-GWR was retained.
- Only for Analyst D was the model choice relatively straightforward, where the MS-GWR (Table 4) fit indicated that each predictor-response relationship was operating at a different spatial scale.

This practical rather than theoretical focus potentially indicates ambiguity in the GWR route map. For example, there are, of course, other spatial regression models that could be used, including the spatially lagged models. However, what is described in this route map is a largely data-led approach which reflects a reality of statistical modeling—there is rarely a perfect model and even selecting one amongst competing choices relies on some subjective judgment, informed by the analyst's purpose and point-of-view. Here, the route map describes a process of investigations, which we have described with some broad guidelines or rules mitigated by practical considerations. However, these are embedded within wider philosophical considerations associated with the construction of statistical models and the analytical decisions that are made. Put simply, what the route map implies is that for any given user, the strength of the evidence provided by the measures of spatial structure, model performance, coefficient significance, etc. will depend on how well users see the data as representing the process being considered. Weak model fits, spatial autocorrelation of residuals, local coefficients whose signs flip across space, spatially varying significance of the coefficients, could all be as a result of insufficient, or even incorrect evidence. Depending on your philosophical position and the way you see data, statistical models, and so forth, any relationship heterogeneity observed through a GWR analysis may simply be a consequence of missing predictors, i.e., global, LRM, or SAM *misspecification*. These are long-standing objections to GWR (Brunsdon and Fotheringham 1998) rooted in a view of spatial dependencies in data as a statistical “nuisance” to be fixed (Harris 2019). However, an alternative and more productive view is that they arise due to environmental, social, or other processes that are spatially contingent, spatially varying, scale-dependent and poorly described by one size fits all types of modeling. Even if the apparent spatial effects are due to a mis-specified model, the value of GWR remains as an exploratory tool for helping to identify that mis-specification. However, the greater mis-specification may actually lie in a presumption that regression relationships would be constant in their effects across a study region. That is a bold and unnecessarily limiting claim.

Secondary model decisions

Having arrived at *GWR Basecamp* through a *primary* analysis, where one from a standard GWR, MX-GWR, or MS-GWR form is considered suited to the observed spatially varying relationships, the next part of the GWR route map is the consideration of *secondary* GWR model issues. To avoid overly lengthening the article, strategies for *secondary* model decisions (*scaling the summit*) are described below but not investigated for their effects on the preceding models. In order of importance, the following issues should be investigated: (a) *predictor collinearity*, (b) the *influence of outliers*, and (c) evidence of a *dependent error term*. These should be examined globally using standard approaches and locally through the associated GWR form. These issues can be detrimental to a reliable GWR analysis, giving rise to say, spurious local changes in the sign of the coefficient estimates between positive and negative and local changes in significance. They can also compromise bandwidth estimation, where GWR fits of a *secondary* analysis will often give rise to different (optimized) bandwidths or a change in the behavior of the bandwidth function to that found with the *primary* analysis, and thus, potentially changing the chosen GWR form (e.g. see respectively, Cho and Lambert 2010; Harris, Fotheringham, and Juggins, 2010; Gollini et al. 2015).

Collinearity

For any global regression, collinearity occurs when pairs of predictors have a strong linear relationship between each other, either positive or negative. Broadly, collinearity may be a problem when correlation coefficients for a predictor pair are >0.8 or < 0.8 as these can affect model reliability and precision: It becomes increasingly difficult to reliably estimate the specific effect of a predictor on a response the more that predictor becomes indistinguishable from others in the model. Diagnostics such as matrix condition numbers (CNs), predictor variance inflation factors (VIFs), and variance decomposition factors (VDPs) can be found and rules of thumb applied (CNs > 30, VIFs > 10, and VDPs > 0.5) to indicate worrying levels of collinearity (Belsey, Kuh, and Welsch 1980). Often a simple remedy is to remove one or more predictors. The difficultly is in deciding which predictor(s) to remove, especially when all are considered important to describing the study process. Here, a penalized regression can provide a solution, as by design it includes a model specification (i.e., variable selection) capability (Zou and Hastie 2005; Friedman and Hastie 2010; Dormann et al. 2013).

Collinearity may also be present in some local predictor data subsets of GWR even when not observed globally (Wheeler and Tiefelsdorf 2005). Compositional, categorical, and ordinal predictors can be particularly problematic, often resulting in exact local collinearity making bandwidth optimization impossible. Geographically weighted collinearity diagnostics (CNs, VIFs, and VDPs) are available for standard GWR (Wheeler 2007, 2013; Lu et al. 2014) and if any observed collinearity is considered a concern, a standard GWR can be replaced with a penalized GWR form (Wheeler 2007, 2009; Gollini et al. 2015; Li and Lam 2018). Mapping geographically weighted correlation coefficients (Fotheringham, Brunsdon, and Charlton 2002) between predictor variable pairs can also be useful to identify areas of local collinearity.

Outliers

For outliers, it is first useful to examine the LRM and MS-GWR residuals of the *primary* analysis for evidence of outliers that may influence the validity of their fits. This should be done spatially (for example, with maps of standardized residuals), to determine where any GWR analysis may be compromised. Again, robust (outlier resistant) theory in the global regression case (e.g., Huber 1981; Marazzi 1993) has been transferred to the local case with robust extensions to standard forms of GWR only (Fotheringham, Brunsdon, and Charlton 2002; Farber and Páez 2007; Harris, Fotheringham, and Juggins, 2010; Zhang and Mei 2011). These handle influential outliers arising globally, but also locally in each individual regression, which may go undetected in any global assessment (i.e., via the standardized residual maps, above).

Dependence in the error data

As with the LRM, most forms of GWR assume that the errors, e_i are independently normally distributed with zero mean and constant variance σ^2. To examine for a non-constant error variance (in a non-spatial, global manner), the LRM's fitted values can be plotted against its residuals. A funnel shape indicates that a heteroscedastic regression should be considered, such as through some consistent estimator (see Davidson and MacKinnon 1993) or a weighted least squares (WLS) estimator. A direct heteroscedastic extension to standard GWR is given in Fotheringham, Brunsdon, and Charlton (2002) and for MX-GWR in Mei et al. (2021) wherein both instances the error variance varies geographically.

Although it is common for any GWR fit to reduce error spatial autocorrelation over that found with a LRM fit (as demonstrated in Primary model decisions section), it is likely that error autocorrelation will also occur for each local regression in a GWR. GWR models that account for local autocorrelation effects have been proposed including an extension to standard GWR (Brunsdon, Fotheringham, and Charlton 1998) and an extension to MX-GWR (Geniaux and Martinetti 2018) through autoregressive GWR model forms.

Prediction and other GW models

Finally, the clear objective of the analyses described in this study is relationship inference. If GWR were to be used as a spatial predictor (Harris, Fotheringham, Crespo, et al. 2010) or the geographically weighted (GW) framework used to construct a spatial classifier (Brunsdon and Fotheringham 2007; Comber and Harris 2016) then a different route map would result (Comber et al. 2020). For this case, prediction/classification accuracy and its inference would be to the fore with respect to model choice. Further, the route map would include GW hybrids with kriging (Harris, Charlton,

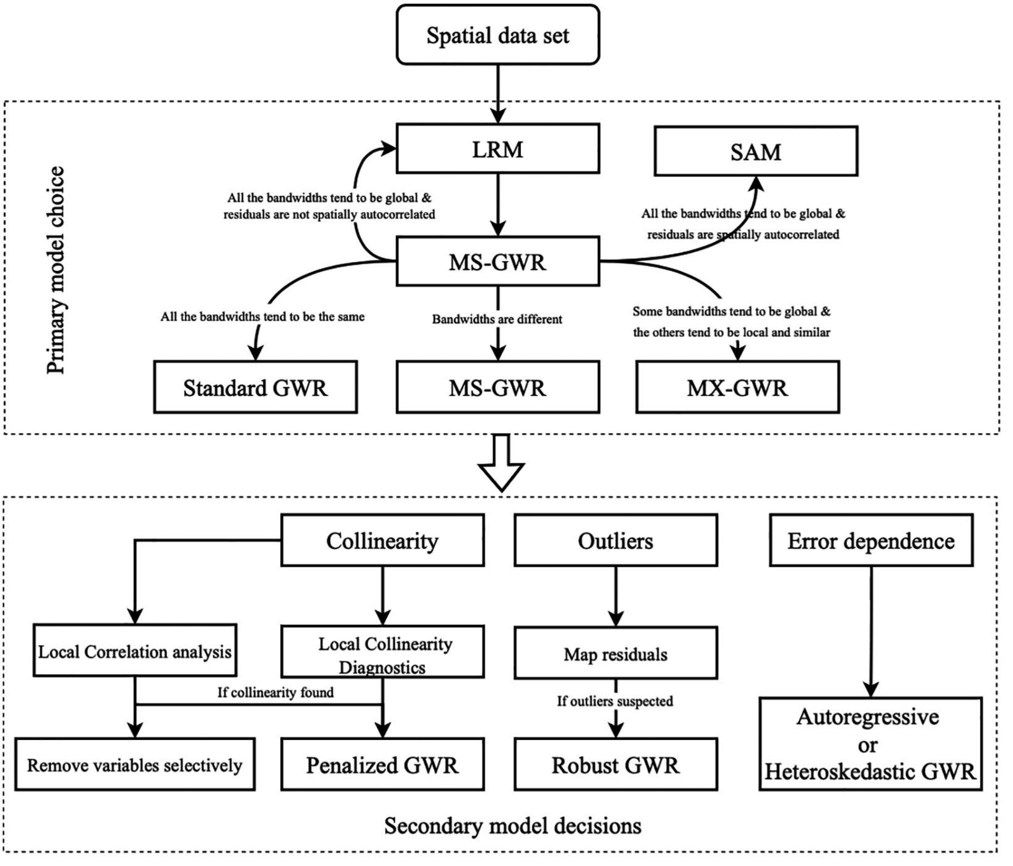

Figure 8. Flowchart of the GWR route map.

and Fotheringham 2010; Harris, Fotheringham, Crespo, et al. 2010) and with machine learning (Hagenauer and Helbich 2021; Quiñones and Goyal 2021), together with GW tools for assessing accuracy (Comber, Brunsdon, and Charlton 2017; Tsutsumida et al. 2019).

Concluding remarks

Geographically Weighted Regression provides a framework to investigate spatial relationships in data, how their effects vary geographically, and their varying scales of interaction. Its use in analyses of environmental and socio-economic data continues to grow and is easily undertaken in a number of software implementations. This article describes a GWR route map of *primary* and *secondary* considerations to ensure the GWR analysis is justified in terms of the aims of the analysis and the characteristics of the data, over alternate models, with fixed regression coefficients. As summarized in Figure 8, the route map has the following *primary* steps:

1. A LRM (basic linear regression model) should always be undertaken and the results investigated.
2. A MS-GWR (multi-scale GWR) should always be calibrated and the estimated bandwidths interrogated.

3. Following the investigations of steps (1) and (2), the analysis should proceed with a standard GWR, or a core variant in MX-GWR (mixed GWR) or MS-GWR, only if a spatially varying coefficient model is considered appropriate. Otherwise, a LRM or a SAM (spatially autocorrelated model) should be chosen.

The LRM (step 1) provides global insight into how the predictors relate to the response, which relationships are significant, and measures of model fit. This step includes evidence of global spatial autocorrelation in the residuals, for example, through a Moran's *I* analysis.

The MS-GWR (step 2) provides information through the MS-GWR bandwidths about the different scales of relationships in the data, where some may be local and others global. The MS-GWR bandwidths describe the degree of spatial dependency associated with each variable's relationship to the response, and thus an indication of process heterogeneity. Insignificant Moran's *I* estimates of the spatial autocorrelation of the MS-GWR residuals provide evidence that accounting for relationship spatial heterogeneity using MS-GWR is capturing most of the structural variation in the data.

Investigations of the LRM and MS-GWR results (step 3), along with other candidate GWR models guide the choice of the final *primary* model (i.e., a LRM or SAM, standard GWR, MX-GWR, or MS-GWR). A LRM should be used when all bandwidths from MS-GWR tend towards the global situation, including the intercept (i.e., are greater than ~80% of the maximum distance between data points or 80% of the data points in the adaptive bandwidth case), and where spatial autocorrelation in the residuals is either absent or if present does not significantly effect process interpretation (as the case for Analyst C, above).

If spatial autocorrelation in the LRM residuals is present *and* MS-GWR bandwidths are not all large, then a GWR variant can be considered:

- A standard GWR should be considered in the rare situation when all of the MS-GWR bandwidths tend to the same value.
- A MX-GWR should be considered when the MS-GWR bandwidths indicate two distinct sets of bandwidths, with one set tending to the global and with the other set tending to a similar local scale.
- A MS-GWR should be considered when the all of bandwidths vary, suggesting that each data relationship operates at different spatial scales.

Only now can *secondary* GWR model decisions be considered (Figure 8), noting that not all *primary* GWR forms have all *secondary* model options in current implementations and that steps are currently being made to address this (Comber et al. 2021).

It is important to stress that the final *primary/secondary* model choice should not be guided by simply selecting the model with the lowest AICc value, especially as the aim of any GWR analysis is to explore relationship spatial heterogeneity and spatial variations in the process. Rather, interrogation of the coefficient estimates and their uncertainty arising from the different models, in the context of the analysis aims is paramount.

Acknowledgments

This research was supported by the Natural Environment Research Council Newton Fund Grant (NE/N007433/1 and NE/S009124/1), the Biotechnology and Biological Sciences Research Council Grants (BBS/E/C/000J0100, BBS/E/C/000I0320 and BBS/E/C/000I0330), the National Natural Science Foundation of China (No. 41571130083 and No. 42071368) and the National Key Research and Development Program of China (No. 2016YFC0501601). All of the analyses and mapping were undertaken in R, the open source statistical software using the GWmodel package. The code and data used in this paper can be downloaded from https://github.com/lexcomber/GWRroutemap.

References

Belsley, D. A., E. Kuh, and R. E. Welsch. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: Wiley.

Bivand, R., D. Yu, T. Nakaya, and M. A. Garcia-Lopez. (2013). spgwr: Functions for Computing Geographically Weighted Regressions. R package version 0.6-22. http://CRAN.R-project.org/package=spgwr

Brunsdon, C., and A. Comber. (2020). "Opening Practice: Supporting Reproducibility and Critical Spatial Data Science." *Journal of Geographical Systems* 23(4), 477–96. https://doi.org/10.1007/s10109-020-00334-2

Brunsdon, C., A. S. Fotheringham, and M. Charlton. (1996). "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity." *Geographical Analysis* 28, 281–98.

Brunsdon, C., A. S. Fotheringham, and M. Charlton. (1998). "Spatial Nonstationarity and Autoregressive Models." *Environment and Planning A* 306, 957–93.

Brunsdon, C., A. S. Fotheringham, and M. Charlton. (1999). "Some Notes on Parametric Significance Tests for Geographically Weighted Regression." *Journal of Regional Science* 39, 497–524.

Brunsdon, C., A. S. Fotheringham, and M. Charlton. (2007). "Geographically Weighted Discriminant Analysis." *Geographical Analysis* 39(4), 376–96.

Brunton, L. A., N. Alexander, W. Wint, A. Ashton, and J. M. Broughan. (2017). "Using Geographically Weighted Regression to Explore the Spatially Heterogeneous Spread of Bovine Tuberculosis in England and Wales." *Stochastic Environmental Research and Risk Assessment* 31(2), 339–52.

Charlton, M., A. S. Fotheringham, and C. Brunsdon. (2003). GWR 3: Software for Geographically Weighted Regression. Maynooth: NCG, National University of Ireland Maynooth.

Cho, S.-H., D. M. Lambert, and Z. Chen. (2010). "Geographically Weighted Regression Bandwidth Selection and Spatial Autocorrelation: An Empirical Example Using Chinese Agriculture Data." *Applied Economics Letters* 17, 767–72.

Cliff, A. D., and J. K. Ord. (1973). Spatial Autocorrelation. London: Pion.

Comber, A., C. Brunsdon, M. Charlton, G. Dong, R. Harris, B. Lu, Y. Lu, D. Murakami, T. Nakaya, Y. Wang, and P. Harris. (2020). "The GWR Route Map: A Guide to the Informed Application of Geographically Weighted Regression." *arXiv: 2004.06070* https://arxiv.org/abs/2004.06070

Comber, A., C. Brunsdon, M. Charlton, and P. Harris. (2017). "Geographically Weighted Correspondence Matrices for Local Change Analyses and Error Reporting: Mapping the Spatial Distribution of Errors and Change." *Remote Sensing Letters* 8(3), 234–43.

Comber, A., P. Harris, and N. Tsutsumida. (2016). "Improving Land Cover Classification Using Texture Variables Outputted from a Geographically Weighted Principal Components Analysis." *ISPRS Journal of Photogrammetry and Remote Sensing* 119, 347–60.

Comber, A., N. Malleson, H. Nguyen Thi Thuy, T. Bui Quang, M. Kieu, H. Huu Phe, and P. Harris. (2021). "Multiscale Geographically Weighted Discriminant Analysis." GIScience 2021 Short Paper Proceedings. https://doi.org/10.25436/E2PP4F

Comber, A., Y. Wang, Y. Lü, X. Zhang, and P. Harris. (2018). "Hyper-local Geographically Weighted Regression: Extending GWR through Local Model Selection and Local Bandwidth Optimization." *Journal of Spatial Information Science* 17, 63–84. https://doi.org/10.5311/JOSIS.2018.17.422

Davidson, R., and J. MacKinnon. (1993). Estimation and Inference in Econometrics, 531. New York: Oxford University Press.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach. (2013). "Collinearity: A Review of Methods to Deal with it and a Simulation Study Evaluating their Performance." *Ecography* 36, 27–46.

Farber, S., and A. Páez. (2007). "A Systematic Investigation of Cross-validation in GWR Model Estimation: Empirical Analysis and Monte Carlo Simulations." *Journal of Geographical Systems* 9, 371–96.

Foody, G. M. (2003). "Geographical Weighting as a Further Refinement to Regression Modelling: An Example Focused on the NDVI-Rainfall Relationship." *Remote Sensing of Environment* 88, 283–93.

Fotheringham, A. S., and C. Brunsdon. (1999). "Local Forms of Spatial Analysis." *Geographical Analysis* 314, 340–58.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. New York: Wiley.

Fotheringham, A. S., W. Yang, and W. Kang. (2017). "Multiscale Geographically Weighted Regression mgwr." *Annals of the American Association of Geographers* 107(6), 1247–65.

Friedman, J., T. Hastie, and R. Tibshirani. (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software* 331, 1–22.

Geniaux, G., and D. Martinetti. (2018). "A New Method for Dealing Simultaneously with Spatial Autocorrelation and Spatial Heterogeneity in Regression Models." *Regional Science and Urban Economics* 72, 74–85.

Gollini, I., B. Lu, M. Charlton, C. Brunsdon, and P. Harris. (2015). "GWmodel: an R Package for Exploring Spatial Heterogeneity using Geographically Weighted Models." *Journal of Statistical Software* 63, 1–50.

Guthery, F. S., L. A. Brennan, M. J. Peterson, and J. J. Lusk. (2005). "Information Theory in Wildlife Science: Critique and Viewpoint." *The Journal of Wildlife Management* 69, 457–65.

Hagenauer, J., and M. Helbich. (2021). "A Geographically Weighted Artificial Neural Network." *International Journal of Geographical Information Science*, 1–21. https://doi.org/10.1080/13658816.2021.1871618

Harris, P. (2019). "A Simulation Study on Specifying a Regression Model for Spatial Data: Choosing Between Autocorrelation and Heterogeneity Effects." *Geographical Analysis* 51(2), 151–81.

Harris, P., M. Charlton, and A. S. Fotheringham. (2010). "Moving Window Kriging with Geographically Weighted Variograms." *Stochastic Environmental Research and Risk Assessment* 24, 1193–209.

Harris, P., A. S. Fotheringham, R. Crespo, and M. Charlton. (2010). "The Use of Geographically Weighted Regression for Spatial Prediction: An Evaluation of Models Using Simulated Data Sets." *Mathematical Geosciences* 42, 657–80.

Harris, P., A. S. Fotheringham, and S. Juggins. (2010). "Robust Geographically Weighted Regression: A Technique for Quantifying Spatial Relationships Between Freshwater Acidification Critical Loads and Catchment Attributes." *Annals of the Association of American Geographers* 100, 286–306.

Harris, R. (2019). "Not Just Nuisance: Spatialising Social Statistics." In Towards a Spatial Social Policy: Bridging the Gap Between Geography and Social Policy, 149–68, edited by A. Whitworth. Bristol: Policy Press.

Huang, Y., M. Yuan, and Y. Lu. (2019). "Spatially Varying Relationships Between Surface Urban Heat Islands and Driving Factors Across Cities in China." *Environment and Planning B* 462, 377–94.

Huber, P. J. (1981). Robust Statistics. New York: John Wiley and Sons.

Kalogirou, S. (2019). lctools: Local Correlation, Spatial Inequalities, Geographically Weighted Regression and Other Tools. R package version 0.2-7. https://CRAN.R-project.org/package=lctools

Lark, R. M., B. R. Cullis, and S. J. Welham. (2006). "On Spatial Prediction of Soil Properties in the Presence of a Spatial Trend: The Empirical Best Linear Unbiased Predictor E-BLUP with REML." *European Journal of Soil Science* 57, 787–99.

LeSage, J. P., and R. K. Pace. (2009). Introduction to Spatial Econometrics. Boca Raton: Chapman and Hall/CRC.

Li, K., and N. S. Lam. (2018). "Geographically Weighted Elastic Net: A Variable-Selection and Modeling Method Under the Spatially Nonstationary Condition." *Annals of the American Association of Geographers* 108(6), 1582–600.

Li, Z., A. S. Fotheringham, W. Li, and T. Oshan. (2019). "Fast Geographically Weighted Regression FastGWR: A Scalable Algorithm to Investigate Spatial Process Heterogeneity in Millions of Observations." *International Journal of Geographical Information Science* 33(1), 155–75.

Lu, B., P. Harris, M. Charlton, and C. Brunsdon. (2014). "The GWmodel R Package: Further Topics for Exploring Spatial Heterogeneity Using Geographically Weighted Models." *Geo-spatial Information Science* 17, 85–101.

Marazzi, A. (1993). Algorithms, Routines and S Functions for Robust Statistics. Pacific Grove, CA: Wadsworth and Brooks/Cole.

McMillen, D. (2013). McSpatial: Nonparametric Spatial DataAnalysis. R package version 2.0. http://CRAN.R-project.org/package=McSpatial

Mei, C.-L., F. Chen, W.-T. Wang, P.-C. Yang, and S.-L. Shen. (2021). "Efficient Estimation of Heteroscedastic Mixed Geographically Weighted Regression Models." *Annals of Regional Science* 66, 185–206.

Mei, C.-L., M. Xu, and N. Wang. (2016). "A Bootstrap Test for Constant Coefficients in Geographically Weighted Regression Models." *International Journal of Geographical Information Science* 30(8), 1622–43.

Moran, P. (1950). "A Test for the Serial Independence of Residuals." *Biometrika* 37, 178–81.

Murakami, D., B. Lu, P. Harris, C. Brunsdon, M. Charlton, T. Nakaya, and D. A. Griffith. (2019). "The Importance of Scale in Spatially Varying Coefficient Modelling." *Annals of the American Association of Geographers* 109, 1–21.

Nakaya, T. (2015). "Geographically Weighted Generalised Linear Modelling." In Geocomputation: A Practical Primer, 201–20, edited by C. Brunsdon and A. Singleton. London: Sage Publication.

Nakaya, T., A. S. Fotheringham, C. Brunsdon, and M. Charlton. (2005). "Geographically Weighted Poisson Regression for Disease Association Mapping." *Statistics in Medicine* 24, 2695–717.

Openshaw, S. (1996). "Developing GIS-Relevant Zone-Based Spatial Analysis Methods." In Spatial Analysis: Modelling in a GIS Environment, 55–73, edited by P. A. Longley and M. Batty. London: John Wiley and Sons.

Oshan, T. M., Z. Li, W. Kang, L. J. Wolf, and A. S. Fotheringham. (2019). "mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale." *ISPRS International Journal of Geo-Information* 86, 269.

Quiñones, S., A. Goyal, and Z. U. Ahmed. (2021). "Geographically Weighted Machine Learning Model for Untangling Spatial Heterogeneity of Type 2 Diabetes Mellitus (T2d) Prevalence in the USA." *Scientific Reports* 11(1), 1–13.

Rey, S. J., and L. Anselin. (2010). "PySAL: A Python Library of Spatial Analytical Methods." In Handbook of Applied Spatial Analysis, 175–93, edited by M. Fischer and A. Getis. Berlin: Springer.

Schabenberger, O., and C. Gotway. (2005). Statistical Methods for Spatial Data Analysis. London: Chapman & Hall.

Silverman, B. (1986). Density Estimation for Statistics and Data Analysis. London: Chapman & Hall.

Song, D., D. J. Brus, F. Liu, D. C. Li, Y. G. Zhao, J. L. Yang, and G. L. Zhang. (2016). "Mapping Soil Organic Carbon Content by Geographically Weighted Regression: A Case Study in the Heihe River Basin, China." *Geoderma* 261, 11–22.

Sun, Y. W., Q. H. Guo, J. Liu, and R. Wang. (2014). "Scale Effects on Spatially Varying Relationships Between Urban Landscape Patterns and Water Quality." *Environmental Management* 54(2), 272–87.

Tobler, W. (1970). "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46, 234–40.

Tsutsumida, N., P. Rodrigues-Veiga, P. Harris, H. Baltzer, and A. Comber. (2019). "Investigating Spatial Errors Structures in Continuous Raster Data." *International Journal of Applied Earth Observation and Geoinformation* 74, 259–68.

Waller, L. A., and C. A. Gotway. (2004). Applied Spatial Statistics for Public Health Data. Hoboken, NJ: Wiley & Sons.

Wang, Y., X. Zhang, and C. Huang. (2009). "Spatial Variability of Soil Total Nitrogen and Soil Otal Phosphorus Under Different Land Uses in a Small Watershed on the Loess Plateau, China." *Geoderma* 150(1–2), 141–9.

Wheeler, D. C. (2007). "Diagnostic Tools and a Remedial Method for Collinearity in Geographically Weighted Regression." *Environment and Planning A* 39, 2464–81.

Wheeler, D. (2009). "Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression: The Geographically Weighted Lasso." *Environment and Planning A* 41, 722–42.

Wheeler, D. (2013). "Geographically Weighted Regression." In Handbook of Regional Science, 1895–921, edited by M. Fischer and P. Nijkamp. Berlin: Springer.

Wheeler, D., and M. Tiefelsdorf. (2005). "Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression." *Journal of Geographical Systems* 7(2), 161–87.

Wolf, L. J., T. M. Oshan, and A. S. Fotheringham. (2018). "Single and Multiscale Models of Process Spatial Heterogeneity." *Geographical Analysis* 50(3), 223–46.

Yang, W. (2014). An Extension of Geographically Weighted Regression with Flexible Bandwidths. Doctoral dissertation, University of St Andrews.

Yoneoka, D., E. Saito, and S. Nakaoka. (2016). "New Algorithm for Constructing Area-Based Index with Geographical Heterogeneities and Variable Selection: An Application to Gastric Cancer Screening." *Scientific Reports* 6, 26582.

You, W., Z. Zang, L. Zhang, Z. Li, D. Chen, and G. Zhang. (2015). "Estimating Ground-level PM10 Concentration in North Western China Using Geographically Weighted Regression Based on Satellite AOD Combined with Calipso and MODIS Fire Count." *Remote Sensing of Environment* 168, 276–85.

Yu, D., Y. Wei, and C. Wu. (2007). "Modelling Spatial Dimensions of Housing Prices in Milwaukee, WI." *Environment and Planning B* 34, 1085–102.

Zhang, H., and C. Mei. (2011). "Local Least Absolute Deviation Estimation of Spatially Varying Coefficient Models: Robust Geographically Weighted Regression Approaches." *International Journal of Geographical Information Science* 25, 1467–89.

Zhang, L., H. Bi, P. Cheng, and C. J. Davis. (2004). "Modeling Spatial Variation in Tree Diameter-Height Relationships." *Forest Ecology and Management* 189, 317–29.

Zou, H., and T. Hastie. (2005). "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society: Series B* 672, 301–20.