

Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data

Binbin Lu^{a*}, Martin Charlton^a, Paul Harris^a and A. Stewart Fotheringham^b

^aNational Centre for Geocomputation, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland; ^bSchool of Geography and Geosciences, University of St. Andrews, St. Andrews, Scotland, UK

(Received 30 July 2012; final version received 8 November 2013)

Geographically weighted regression (GWR) is an important local technique for exploring spatial heterogeneity in data relationships. In fitting with Tobler's first law of geography, each local regression of GWR is estimated with data whose influence decays with distance, distances that are commonly defined as straight line or Euclidean. However, the complexity of our real world ensures that the scope of possible distance metrics is far larger than the traditional Euclidean choice. Thus in this article, the GWR model is investigated by applying it with alternative, non-Euclidean distance (non-ED) metrics. Here we use as a case study, a London house price data set coupled with hedonic independent variables, where GWR models are calibrated with Euclidean distance (ED), road network distance and travel time metrics. The results indicate that GWR calibrated with a non-Euclidean metric can not only improve model fit, but also provide additional and useful insights into the nature of varying relationships within the house price data set.

Keywords: local regression; non-stationarity; road network distance; travel time; real estate

1. Introduction

Waldo Tobler's celebrated dictum, to which he referred as the first law of geography, has been widely adopted as a basic principle in Geographic Information Science (Goodchild 1992). It is at the core of many spatial techniques, such as those found in spatial interpolation and spatial interaction paradigms (Goodchild 2004, Miller 2004, Sui 2004). In most cases, Tobler's first law (Tobler 1970, p. 236) relates to the data, but it can similarly relate to a statistic or model. Here, local or non-stationary models are used in the exploration of some form of spatial heterogeneity. For example, local indicators of spatial association (Getis and Ord 1992, Anselin 1995) are used to investigate heterogeneity in spatial autocorrelation. For this study, our focus lies with the geographically weighted regression (GWR) model – a technique that investigates heterogeneity in data relationships across space (Brunsdon *et al.* 1996). In particular, we empirically assess the value of applying a non-Euclidean distance (non-ED) metric to GWR, with respect to its calibration and performance (in terms of model fit and the interpretation of its results).

For analysing or visualising geographical phenomena, non-ED metrics suit situations when the spatial interaction or spatial behaviour of the process does not follow straight lines or the standard properties of a metric space (Miller 2000). For instance, non-ED

*Corresponding author. Email: binbin.lu.2009@nuim.ie

metrics have routinely been applied in (1) hydrology studies using kriging, where water distances are used for spatial prediction along a stream or river network (e.g. [Curriero 2006](#), [Money et al. 2009](#)); (2) landscape studies using kriging, where landscape-based distance metrics are derived (e.g. [Jensen et al. 2006](#), [Lyon et al. 2010](#)); and (3) socio-economic studies, where Minkowski distances have been used ([Kent et al. 2006](#), [Shahid et al. 2009](#)). In general, empirical work indicates that there is no ‘one-fit-all’ distance metric, and the scope of possible distance metrics in spatial analysis is far larger than a single Euclidean choice. Non-ED metrics are also routinely integrated into Geographic Information Systems. For example, ArcGIS 10 provides options for non-ED metrics in its ‘Geostatistical Analyst’ extension ([ESRI 2011](#)). This includes options for cost and network distances (NDs). [Smith et al. \(2008\)](#) provide an R (<http://www.r-project.org>) package **ramps** to analyse spatial or spatio-temporal correlation structures, where great circle, maximum and absolute distances are available.

Our article is constructed in the following four sections. We start by describing the basics of the GWR technique. Second, the potential of using a non-ED metric in GWR is reviewed. Third as a case study, we model a London house price data set using GWR with, respectively, Euclidean distance (ED) and two non-ED metrics (i.e. ND and travel time (TT)). Properties of our case study data are such that it is likely to benefit from the use of a non-ED metric when modelling with GWR (and many other spatial models). Finally, we summarise this research and suggest future directions. The modelling functions used in this article are included in the **GWmodel** R package ([Lu et al. 2013](#)), which is an integrated framework for handling spatially varying structures via a wide range of geographically weighted (GW) models.

2. Geographically weighted regression

GWR is a non-stationary technique that models spatially varying relationships. Compared with a basic (global) regression, the coefficients in GWR are functions of spatial location. [Fotheringham et al. \(1998, 2002\)](#) give a general form of a basic GWR model as:

$$y_i = \beta_{i0} + \sum_{k=1}^m \beta_{ik} x_{ik} + \varepsilon_i \quad (1)$$

where y_i is the dependent variable at location i ; x_{ik} is the k th independent variable at location i ; m is the number of independent variables; β_{i0} is the intercept parameter at location i ; β_{ik} is the local regression coefficient for the k th independent variable at location i ; and ε_i is the random error at location i .

GWR allows coefficients to vary continuously over the study area, and a set of coefficients can be estimated at any location – typically on a grid so that a coefficient surface can be visualised and interrogated for relationship heterogeneity. GWR makes a point-wise calibration concerning a ‘bump of influence’: around each regression point where nearer observations have more influence in estimating the local set of coefficients than observations farther away ([Fotheringham et al. 1998](#)). In essence, GWR measures the inherent relationships around each regression point i , where each set of regression coefficients is estimated by weighted least squares. The matrix expression for this estimation is

$$\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i y \quad (2)$$

where X is the matrix of the independent variables with a column of 1s for the intercept; y is the dependent variable vector; $\hat{\beta}_i = (\beta_{i_0}, \dots, \beta_{i_m})^T$ is the vector of $m + 1$ local regression coefficients; and W_i is the diagonal matrix denoting the geographical weighting of each observed data for regression point i .

Here, the weighting scheme W_i is calculated with a kernel function based on the proximities between regression point i and the N data points around it. A number of kernel functions can be used for the weighting scheme, where for this study a Gaussian kernel is specified, which in its usual continuous form can be defined as

$$\text{Gaussian : } w_{ij} = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right] \quad (3)$$

where d_{ij} is the distance between observation point j and regression point i , for which the ED is generally employed with planar coordinates, and b is the kernel bandwidth. The bandwidth is the key controlling parameter and can be specified either by a fixed distance (i.e. a fixed bandwidth) or by a fixed number of nearest neighbours (i.e. an adaptive bandwidth). GWR defaults to the corresponding global regression if a very large bandwidth is specified such that all geographical weights tend to unity.

An optimum bandwidth can be found by minimising some model goodness-of-fit diagnostic (Loader 1999), such as the cross-validation (CV) score (Cleveland 1979, Bowman 1984), which only accounts for model prediction accuracy, or the Akaike Information Criterion (AIC) (Akaike 1973), which accounts for model parsimony (i.e. a trade-off between prediction accuracy and complexity). In practice, a corrected version of the AIC is used, which unlike basic AIC is a function of sample size (Hurvich *et al.* 1998). Thus for a GWR model with a bandwidth b , its AIC_c can be found from

$$AIC_c(b) = 2n \ln(\hat{\sigma}) + n \ln(2\pi) + n \left\{ \frac{n = \text{tr}(S)}{n - 2 - \text{tr}(S)} \right\} \quad (4)$$

where n is the sample size; $\hat{\sigma}$ is the estimated standard deviation of the error term; and $\text{tr}(S)$ denotes the trace of the hat matrix S . The hat matrix is the projection matrix from the observed y to the fitted values \hat{y} (Hoaglin and Welsch 1978), where for GWR each row r_i of this hat matrix is

$$r_i = X_i (X^T W_i X)^{-1} X^T W_i \quad (5)$$

where X_i is its i th row of the matrix of independent variables X . In this study, we calibrate all GWR models using the AIC_c approach. Observe that (local) sample size will decrease with bandwidth size (i.e. n becomes N), thus AIC_c increases the relative penalty for model complexity with small bandwidth sizes. Furthermore, as we have specified a continuous kernel, the value of N should be regarded as an *effective N* (i.e. the Gaussian kernel is only asymptotic to its bandwidth).

3. Why use a non-ED metric in GWR?

GWR provides an intuitive and technically accessible tool to explore spatially varying relationships (Páez and Wheeler 2009). With attention to improving its performance,

numerous contributions have been made. For example, different kernel functions have been suggested (e.g. [Brunsdon et al. 1996](#), [Fotheringham et al. 1998](#), [Yrigoyen et al. 2007](#)) and different rules to select an optimum bandwidth have also been proposed (e.g. [Fotheringham et al. 2002](#), [Páez et al. 2002](#), [Farber and Páez 2007](#)). Methods to address local collinearity issues that are inherent in local regression modelling have also been developed ([Wheeler 2007](#), [2009](#)).

However, the choice of distance metric has been largely neglected,¹ and in practice, GWR uses ED to measure ‘spatial proximity’, although great circle measurements for unprojected geographical coordinates are possible ([Charlton et al. 2007](#)). Of the few attempts to use a non-ED metric in GWR, [Lloyd and Shuttleworth \(2005\)](#) modified a ward-to-ward distance matrix by calculating ward-to-ward distance with straight-line distances between ward centroids and within-ward distance as the mean distance between enumeration district centroids; [Huang et al. \(2010\)](#) introduced a spatio-temporal distance under an ellipsoidal coordinate system for their geographically and temporally weighted regression model.

Our real world space is not only an absolute geometric container for human activities, but is also a cognitive space which is mentally constructed by spatial relations or phenomena ([Kitchin 2009](#)). A geographical space is usually too complicated to be measured simply by ED metrics. In practice, a geographical space often refers to an environmental space, which is characterised by the knowledge of the environment that surrounds us ([Montello 1992](#)). This knowledge is often known as environmental, geographical or spatial context. Different spatial contexts may provide distinctive cognitions of the surrounding space. Accordingly, distance metrics are perceived in a concrete spatial context, even for the same person or application ([Worboys 1996](#)).

As addressed empirically by [Longley et al. \(2005\)](#), the distance metric in a distance-decay scheme is potentially dependent on a number of factors, including: physical factors (e.g. rivers, road infrastructure and associated conditions of accessibility), socio-economic factors (e.g. preferences to hospitals, schools and stores) and administrative geographies. In this sense, a number of distance metrics have been used for geographical space, such as (1) Manhattan distance based on the taxicab geometry ([Krause 1987](#)); (2) chamfer distance designed for a lattice or grid space ([Leymarie and Levine 1992](#), [De Smith 2004](#)); (3) shortest path distance ([Smith 1989](#)); and (4) qualitative distance by translating an absolute distance metric to linguistic terms ([Gahegan 1995](#), [Hernández et al. 1995](#), [Guesgen and Albrecht 2000](#), [Yao and Thill 2005](#)).

Considering the many complex and diverse situations that GWR can be applied to, it is reasonable that the option of specifying a non-ED metric, in addition to the Euclidean one, is available. Subject areas that GWR has been applied to include (1) housing market modelling (e.g. [Bitter et al. 2007](#), [Páez et al. 2008](#)); (2) regional economics (e.g. [Yu 2006](#), [Ócal and Yildirim 2010](#)); (3) urban and regional analysis (e.g. [Ali et al. 2007](#), [Noresah and Ruslan 2009](#)); (4) sociology (e.g. [Jordan 2006](#), [Cho and Gimpel 2009](#), [Lee et al. 2009](#)); and (5) ecology and environmental science (e.g. [Svenning et al. 2009](#), [Windle 2010](#)). However, for GWR studies involving transportation, river/stream networks or some complex terrain conditions, ED metrics may fail to reflect true spatial proximity, and instead, non-ED metrics should be considered, such as road ND, TT, water distance or landscape distance. In these situations, the ED metric may lead to inaccurate coefficient estimates, and in turn, a spatial pattern in estimates that is wrongly interpreted because of artefacts from the straight-line measure between the local regression calibration points and the data points. Intuitively, the more appropriate the distance metric is, the better GWR should perform.

Observe that the distance metric is an essential but separate component of the GWR technique. There is no special statement that the distance metric d_{ij} has to be Euclidean. Thus, the ED metric can be directly replaced by an appropriate non-ED measure in the basic GWR model. Furthermore as the weights matrix W_i is diagonal, there is no need to check for positive definiteness, as with a (full) covariance matrix needed in kriging (Diamond and Armstrong 1984). Thus, the theoretical framework of GWR and any related GW model (e.g. GW principal components analysis (Harris *et al.* 2011a)) can still be followed with a generalised distance metric (Euclidean or non-Euclidean).

4. Case study: GWR using a non-ED metric

4.1. Case study data

4.1.1. London house price and hedonic data

As a case study, a house price data set for London, UK, is used to assess and compare GWR models with different distance metrics. This data set is sampled from a house price data set provided by the Nationwide Building Society of the United Kingdom and was combined with various hedonic independent variables (Fotheringham *et al.* 2002). The data consists of 2108 properties sold during the 2001 calendar year, which are geo-coded using the postcode of each property. The locations of these properties are visualised in Figure 1. We choose to model house price data from 2001, as it matches available road network data and also coincides with a UK census that provides some of the hedonic variables.

Thus, the regression for this study is a hedonic house price model (Goodman 1978), a model that has been widely applied in exploring a housing market, where

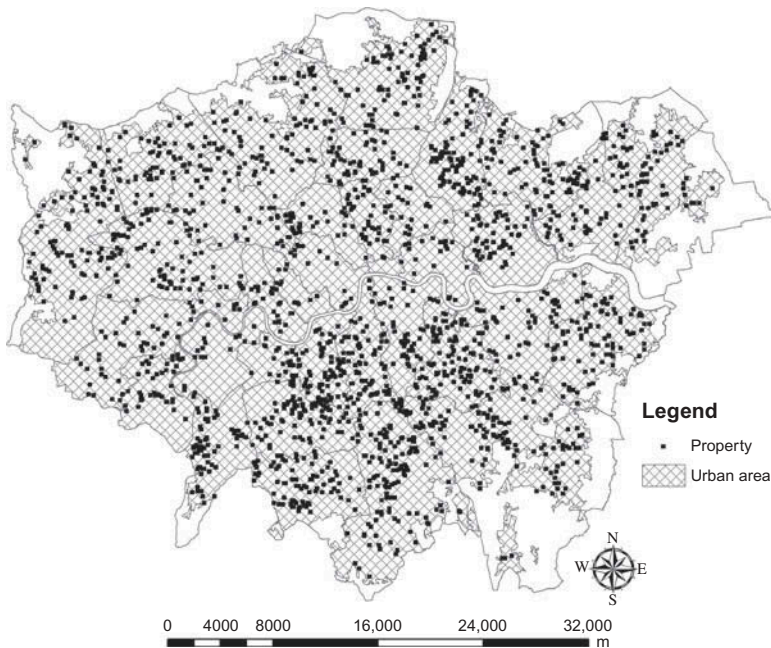


Figure 1. Locations of sampled properties in the London area.

hedonic characteristics are typically divided into locational attributes, structural attributes, neighbourhood attributes and other features (Goodman 1989, Chin and Chau 2003). Accordingly for our study, the sale price, *PURCHASE*, the dependent variable, is related to the following 15 independent variables that reflect structural characteristics, construction time, property type and local household income conditions:

- *FLOORSZ* is the floor size of the property in square metres;
- *BATH2* is 1 if the property has 2 or more bathrooms, 0 otherwise;
- *BEDS2* is 1 if the property has 2 or more bedrooms, 0 otherwise;
- *CENTHEAT* is 1 if the property has central heating, 0 otherwise;
- *GARAGE1* is 1 if the property has one or more garages, 0 otherwise;
- *BLDPWW1* is 1 if the property was built prior to 1914, 0 otherwise;
- *BLDINTW* is 1 if the property was built between 1914 and 1939, 0 otherwise;
- *BLD60S* is 1 if the property was built between 1960 and 1969, 0 otherwise;
- *BLD70S* is 1 if the property was built between 1970 and 1979, 0 otherwise;
- *BLD80S* is 1 if the property was built between 1980 and 1989, 0 otherwise;
- *BLD90S* is 1 if the property was built between 1990 and 2000, 0 otherwise;
- *TYPEDETCH* is 1 if the property is detached (i.e. it is stand-alone), 0 otherwise;
- *TYPETRRD* is 1 if the property is in a terrace of similar houses, 0 otherwise;
- *TYPFLAT* is 1 if the property is a flat or apartment, 0 otherwise;
- *PROF* is the percentage of the workforce in professional or managerial occupations in the census enumeration district in which the house is located.

In many housing market studies, spatial dependence in house price and spatial heterogeneity in its relationships to hedonic independent variables have been explored. For the latter, GWR has been extensively applied to identify spatially varying influences on house price. Here GWR has often improved the goodness-of-fit diagnostics in comparison with (1) the global regression (e.g. [Gao and Asami 2005](#), [Yu 2007](#), [Cho et al. 2009](#), [Díaz-Garayúa 2009](#), [Vichiensan and Miyamoto 2010](#)), (2) the expansion method ([Kestens et al. 2006](#), [Bitter et al. 2007](#)), (3) multivariate moving window kriging ([Páez et al. 2008](#)) and (4) spatial simultaneous autoregressive regression ([Löchl and Axhausen 2010](#)), where models (2) and (3) can similarly account for non-stationary relationships. Non-stationary models, related to GWR, have also been used in this context, such as the area-to-point local kriging with external drift model presented in [Yoo and Kyriakidis \(2009\)](#).

In all such studies, the spatial models were calibrated using an ED metric. However, a house is a place for people to live or work from, which constitutes a main space for their activities. Clusters of houses connected by roads or paths form an organism. Thereby, it is reasonable to consider *connected* relationships underlying a housing market. As such, we investigate the use of ND and TT metrics with GWR and compare their respective performances with basic GWR in exploring the London house price data. Further work could investigate the use of non-EDs for all of the listed spatial models, not just GWR. Observe that preliminary work on GWR with a ND metric (only) can be found in [Lu et al. \(2011\)](#), using a small subset of the London house price data of this study. Here, the resultant goodness-of-fit statistics indicated an initial promise in the ND model, which leads us to this study's considerably extended and more insightful presentation.

4.1.2. London road network data

Road network data produced by the UK Ordnance Survey (OS) in 2001 is used to calculate the ND and TT metrics for our GWR models. To get a relatively accurate TT, the road speed limits are used as the average speeds for each road link. The locations of the speed limits are provided by Transport for London and are used for intelligent speed adaptation technology (Transport for London 2011). Based on these locations, a map of speed limits for the road links over London is produced in Figure 2. Notably, the locations of our house data are not on the road network. As such, we established a topological relationship between the house points and the road network, by allowing each house point to correspond with its nearest point on the network. Observe that this operation, together with the errors attached to geo-coding properties at the post-code level, and known quality issues with the road network data itself will introduce a degree of bias in the ND and TT distances d_{ij} used in all GWR models.² The resultant ND and TT matrices for our study data are calculated using the R packages, **shp2graph** (Lu and Charlton 2011) and **igraph** (Csardi and Nepusz 2006).

4.2. Hedonic variable selection

4.2.1. Global regressions

As with any GWR study, it is important to estimate the parameters of the global regression, so that this benchmark model can be compared to its GWR counterpart. As there is no single agreed functional form in hedonic price modelling (Halvorsen and Pollakowski 1981, Fotheringham *et al.* 2002), we use a pseudo stepwise procedure for exploring our data with a limited number of ordinary least squares (OLS) regression

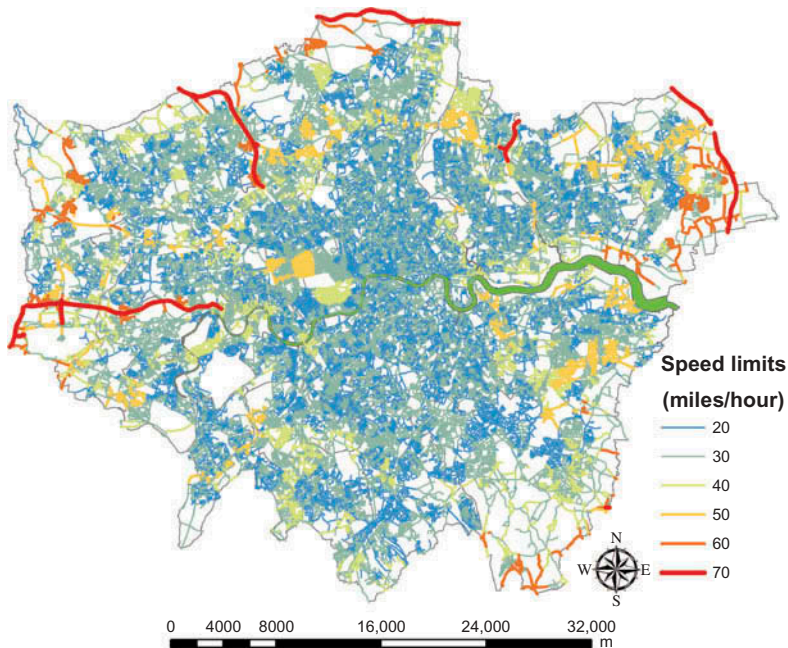


Figure 2. London road network and speed limits of road links.

models with respect to hedonic variable selection. The procedure is described in the following four steps:

- Step 1. Start by fitting all possible bivariate OLS regressions by sequentially regressing a single hedonic variable against the *PURCHASE* variable.
- Step 2. Find the best performing model which produces the minimum AIC_c value and permanently include the corresponding hedonic variable in subsequent models.
- Step 3. Sequentially introduce a variable from the remaining group of hedonic variables to construct new models with the permanently included hedonic variables and determine the next permanently included variable from the best fitting model that has the minimum AIC_c value.
- Step 4. Repeat step 3 until all the hedonic variables are permanently included in the model.

Observe that although we use this pseudo stepwise procedure for hedonic variable selection with the OLS regression, the procedure was actually developed for use with GWR. Here GWR models can be compared using different (1) distance metrics, (2) kernel specifications and (3) hedonic variable subsets. This complexity in model choice for GWR soon becomes unmanageable with a standard stepwise procedure, and our pseudo stepwise procedure tries to address this problem. That said, the GWR model selection process can still be somewhat tedious, even with this simplified procedure. We choose to use the pseudo stepwise procedure for the OLS regression, instead of a standard procedure (which for the OLS model is entirely viable), as it provides a means to compare a range of ordered OLS regression fits (by variable subset) with their corresponding GWR fits with different distance metrics and different kernel specifications.

In our model selection procedure, the hedonic variables are iteratively included into the OLS model in a ‘forward’ direction. The procedure results in 120 OLS regression fits, which is a small subset of the total number of models that are possible (thus, there is no guarantee that most parsimonious OLS regression fit is actually found). The results of this procedure are shown in Figure 3a and b. Figure 3a presents a circle view of the 120 OLS regression fits, where the dependent variable *PURCHASE* is located in the centre of the chart and the independent hedonic variables are represented as nodes differentiated by shapes and colours; Figure 3b presents the corresponding AIC_c values from the same 120 OLS regressions.

For clarity on Figure 3a, a model with *FLOORSZ* as the hedonic variable produces the lowest AIC_c for the first round of bivariate regressions. This variable is thus retained, giving the inner most (partial) circle as a series of red stars, the symbol associated with *FLOORSZ*. Next, for the regressions with two hedonic variables, a model with *FLOORSZ* and *PROF* produces the lowest AIC_c . Thus, *PROF* is the next variable retained, giving the next (partial) circle as a series of orange diamonds, the symbol associated with *PROF*. This pattern continues until model no. 120, which includes all 15 variables. Thus, this chart need only be viewed in context of which variable is retained at each of the 15 key stages (i.e. first *FLOORSZ*, then *PROF*, then *BATH2*, etc., which is the order given in the legend). Tentatively, we can assume this order reflects the relative importance of each hedonic variable in explaining house price. Observe that at each of the 15 stages, the regressions are ordered by their AIC_c , with the regression with the highest AIC_c first. This is clearer in Figure 3b. To assist in the reading of this chart, grey lines are used to separate groups of models with a different number of hedonic variables (intercept not counted);

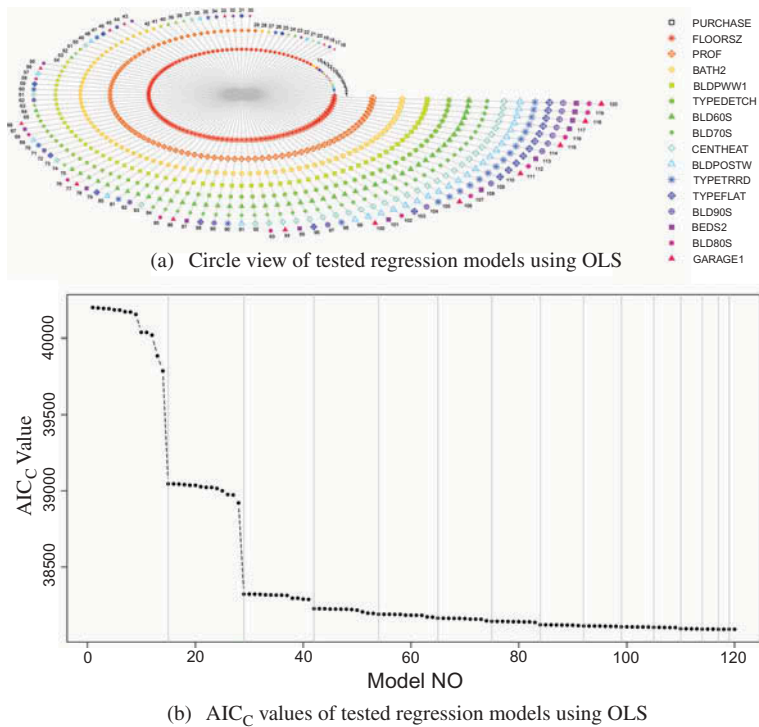


Figure 3. Results of the OLS regression stepwise procedure (120 models in total).

that is starting at 15 models each with one variable, then reducing to one model with all 15 variables.

From Figure 3b, where the same 120 regressions are plotted in the same order as Figure 3a, we can observe that (a) AIC_c reduces as more variables are included; (b) this reduction is relatively small once more than three hedonic variables are included. Observe that for a model to be considered a more parsimonious and improved fit over an alternative, all that is required is a difference in AIC_c (or AIC) of 3 or more units (Burnham and Anderson 2002). The size and sign of the AIC_c values are irrelevant, in this respect. Thus, a reduction in AIC_c from some benchmark model can be used as a key model fit diagnostic.

4.2.2. GWR calibrations with ED, ND and TT metrics

We now calibrate the corresponding GWR models to the same 120 OLS regressions, above. Here we apply GWR using ED (the basic fit), ND and TT metrics, using both fixed and adaptive kernel bandwidths, with each bandwidth found optimally via the minimised AIC_c approach. Thus, $120 \times 6 = 720$ GWR models are assessed in total, and for a given model number (1–120), the hedonic variables will remain the same, whilst the optimal bandwidth and distance metric of the six GWR models will vary.

Observe that as different distance metrics are used, it is not appropriate to compare the three adaptive or the three fixed bandwidth GWR models (for a given model number) using the same bandwidth size, as inevitably this would favour one model over the other

two. For example, if this bandwidth was chosen as that found optimally for the basic GWR fit, then this particular GWR model would have an unfair advantage. Conversely, comparing our GWR models with the three different optimised bandwidths would appear to make model comparisons difficult, as we cannot isolate the effects of a particular distance metric. However, as discussed in Section 2, AIC_c implicitly accounts for bandwidth size, as GWR models with small bandwidths are considered more complex (i.e. small local sample size and less degrees of freedom) than corresponding GWR models with large bandwidths (i.e. large local sample size and more degrees of freedom) (see Fotheringham *et al.* (2002)).

Thus, we argue that provided the ED, ND and TT, metric GWR calibrations are based on an AIC_c approach (or a similar model parsimony statistic) for their optimised bandwidths, then problems of a subjective comparison associated with using different bandwidths should be alleviated. Conversely, if our GWR models were based on a CV approach (or a similar prediction accuracy statistic) for their optimised bandwidths, then it is felt that the differences in bandwidth (and in particular, differences in local sample size) do matter, as they are not accounted for in the minimised statistic. That is, in the latter case, it is not possible to determine whether an improvement in model fit is a consequence of the distance metric used or a consequence of using more or less data in the local calibration.

Figure 4 plots the AIC_c values from all 720 GWR calibrations. From this plot, the following points are observed with respect to goodness-of-fit: (1) GWR models using fixed bandwidths consistently outperform those using adaptive ones; (2) for fixed bandwidths, GWR models using the TT metric consistently perform the best; (3) for fixed bandwidths, GWR models using the ND metric usually outperform basic GWR, except for models with 10–12 hedonic variables when they perform equally or 12–15 hedonic variables when basic GWR performs better; (4) for adaptive bandwidths, GWR models using non-ED metrics usually outperform basic GWR, except for models with one to two hedonic variables; (5) for adaptive bandwidths, GWR models using ND metrics consistently perform better than those using TT metrics; (6) differences between GWR calibrations with respect to the choice of distance metric are greater using fixed bandwidths than adaptive ones; and (7) overall, GWR models using non-ED metrics tend to perform better than basic GWR models.

Observe that similar to the 120 OLS regressions, there is no guarantee that the most parsimonious GWR models are included in our 720 calibrations with respect to hedonic variable selection. Furthermore, the model rankings found for the OLS regressions are not the same as that found for the GWR fits (compare Figure 3b with Figure 4) and there is no reason why they should be. Here the largest reductions in AIC_c for the OLS regressions only occur when the number of hedonic variables is increased, while this is also true for GWR, significant reductions in AIC_c can also occur when a particular hedonic variable is included. For example, significant reductions occur at model nos. 31, 47, 59, 72, 79, 88, 96, 103 and 110 where in all cases, the *TYPEFLAT* variable is introduced (see also Figure 3a). Thus, this particular hedonic variable is clearly more important in a GWR calibration than the OLS regression calibration (and in turn, suggests a noteworthy non-stationary relationship in this respect). Observe, however (and noting points (3) and (4) above), that no particular hedonic variable clearly (or significantly) stands out in response to a change of metric in the GWR fit. If this were the case, then Figure 4 would have lines that on occasion strongly diverged from each other, when this particular variable was introduced.

These issues are not considered an oversight or problem, as our study's prime objective is a comparison of GWR models with different distance metrics, where this

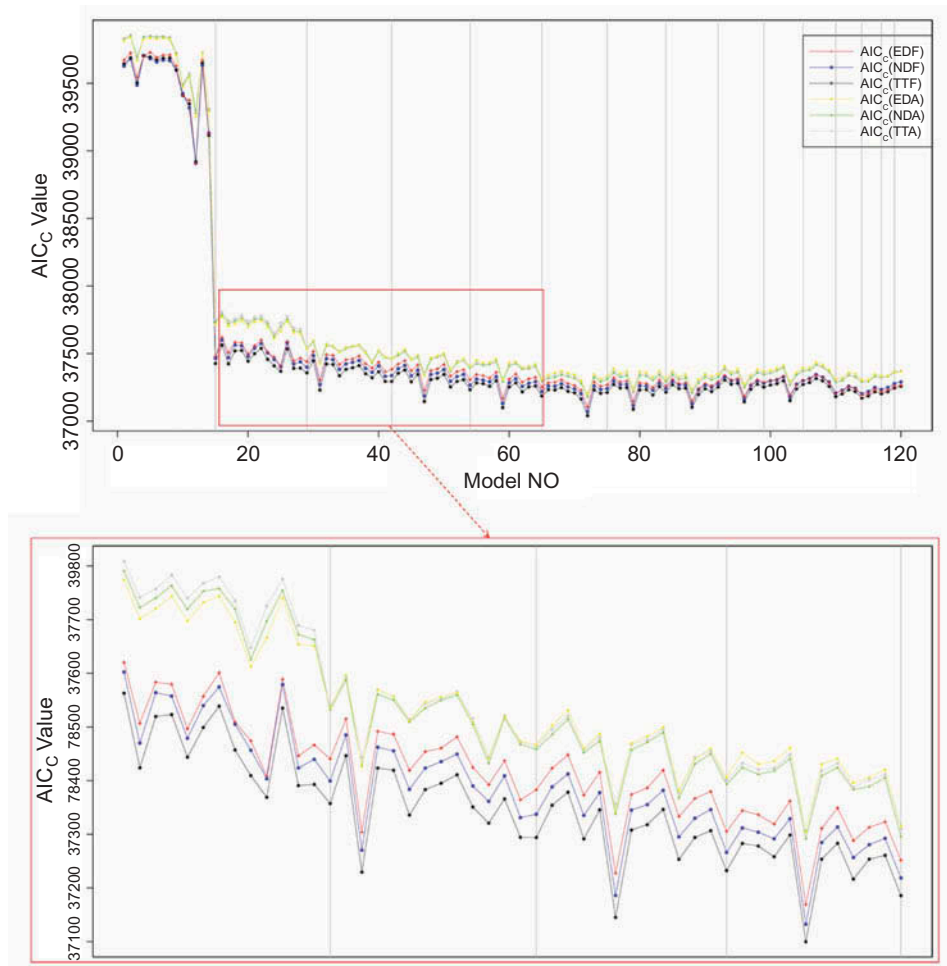


Figure 4. AIC_c values from 720 GWR models where ED, ND and TT metrics are applied for both fixed and adaptive bandwidths. Abbreviations: (1) EDF, ED with a fixed bandwidth; (2) NDF, ND with a fixed bandwidth; (3) TTF, TT with a fixed bandwidth; (4) EDA, ED with an adaptive bandwidth; (5) NDA, ND with an adaptive bandwidth; (6) TTA, TT with an adaptive bandwidth.

study's 120 OLS regressions and 720 GWR fits should be sufficient for this purpose. Variable selection in GWR is a topic in its own right, which requires its own presentation. Our investigation does provide some insight into this issue when distance metrics are allowed to vary. However, for this particular data set, variable choice appears to only marginally mitigate against a poorly chosen distance metric (i.e. points (3) and (4) above), where basic GWR sometimes performs slightly better than a non-ED metric counterpart. An exploratory analysis was conducted in this respect, using the pseudo stepwise procedure for the six different GWR models, where, for a given number of hedonic variables, a slightly different set of hedonic variables can sometimes result for each GWR model form. We do not present these results as (1) a different set of hedonic variables may result from only a small reduction in AIC_c and (2), comparing GWR models with different hedonic variables, in addition to different distance metrics, would again entail that the effects of a

particular distance metric are not isolated (which is our study's focus). For other studies, however, the choice of independent variables in the GWR model may more strongly depend on the chosen distance metric. As variable choice can similarly depend on the bandwidth, then variable selection in this richer class of GWR models is clearly a challenging topic.

4.3. Investigation of a single model specification

It is unrealistic to delve deeper into all 720 GWR models, one by one. Thus, we choose a representative model to illustrate more specific differences in the GWR fits using ED, ND and TT metrics. As shown in [Figure 3](#), there is relatively little reduction in AIC_c for the OLS regressions when more than three hedonic variables are included. Similarly from [Figure 4](#), there is relatively little reduction in AIC_c for the GWR models when more than two hedonic variables are included. Thus, a GWR model with two or three hedonic variables seems an appropriate choice. Further, this model should be specified with a fixed bandwidth, as AIC_c values are consistently lower than that found using an adaptive bandwidth ([Figure 4](#)). From a range of possible candidates (models no. 16 to no. 42), we choose model no. 42. This model has the largest reduction in AIC_c for the basic GWR fit over a non-ED GWR fit using a fixed bandwidth (in this case, using the TT metric). This model also provides the best performing OLS regression with three or less than three hedonic variables. Model no. 42 has the following form:

$$PURCHASE_i = \beta_{0i} + \beta_{1i}FLOORSZ_i + \beta_{2i}PROF_i + \beta_{3i}BATH2_i \quad (6)$$

4.3.1. Summary of the OLS regression and GWR models

Bandwidth, R -squared and AIC_c results for model no. 42, using OLS regression and GWR are given in [Table 1](#). Observe that the bandwidths for the GWR models using ND and TT metrics are actually relatively similar to each other. Here we need to look again at [Figure 2](#), where the average speed is 30 miles/hour in the local areas. From this, a distance bandwidth for the TT GWR model can be approximately found using the following expression: $30 \times 0.44704 \times 174.95 = 2346.29$ m (which is not too dissimilar from the bandwidth of the ND GWR model of 2375.78 m). For the bandwidth of the basic GWR model, it is hard to clarify the relationship of this bandwidth with the other two. However, we can say that this bandwidth is not too dissimilar from the other two, as the measure of ND is always larger than or at least equal to the measure of ED, for the same pair of

Table 1. Calibrations and outputs for the model no. 42 via an OLS regression and via GWR using ED, ND and TT metrics with fixed bandwidths.

	OLS regression	GWR (EDF)	GWR (NDF)	GWR (TTF)
Bandwidth	–	1914.50 (m)	2375.78 (m)	174.95 (s)
R -squared	0.708	0.864	0.875	0.885
AIC_c	38205.29	37382.97	37337.39	37293.96
AIC_c reduction (from OLS regression fit)	0	822.32	867.90	911.33

locations. The three bandwidth functions (not shown) for the GWR models (i.e. AIC_c vs. bandwidth size) were all well-behaved, where in each case, a clear minimum AIC_c was reached at the optimal bandwidth.

As expected, all three GWR models provide better model fit diagnostics than the OLS regression does, where (i) the R -squared values have been increased by around 16% and (ii) the AIC_c values have been reduced by at least 822.32. For the GWR models, the two fits using a non-ED metric provide a significant reduction in AIC_c compared to the ED fit, with the fit using the TT metric (significantly) performing the best. Following the discussion given in Section 4.2.2, it is not considered wise to compare the R -squared values for the three GWR models, since R -squared values only reflect prediction accuracy and different bandwidths have been specified. Also higher R -squared values are likely if bandwidths had been optimised via a CV approach.

4.3.2. Spatial analysis of the GWR residuals

Figure 5a–c depicts discrepancy maps for the absolute residuals (i.e. the absolute value of the actual *PURCHASE* price minus the GWR predicted *PURCHASE* price) from the three GWR models using ED, ND and TT metrics. Here, Figure 5a subtracts the absolute residuals of the ED GWR model from the absolute residuals of the ND GWR model and Figure 5b and c follow accordingly. Thus, our discrepancy maps are only concerned with comparing the magnitude of the residuals from two different GWR models and not the direction of the residuals (i.e. they do not convey instances of over- or under-prediction). For example, for Figure 5a, positive values indicates where ND GWR predicts better, while negative values indicates where ED GWR predicts better. For context and to aid interpretation, Figure 6 provides a map of the London boroughs.

From these maps, it appears that the TT GWR model tends to perform both the best and worst locally, while both non-ED GWR models tend to outperform the basic GWR model in this local assessment of model fit accuracy. From Figure 5a and b, the non-ED GWR models can significantly reduce residual size with respect to the basic GWR model along the River Thames, which runs centrally through London (shown by numerous large positive difference classes, coloured red). This effect is particularly strong in the West London boroughs of Richmond upon Thames, Hammersmith & Fulham and Wandsworth. Such reductions in the residuals are entirely expected, where differences between the distance metrics are likely to be at their highest, such as points neighbouring some barrier, which in this instance is the River Thames. Here the ND and TT metrics would strongly depend on the number of bridges or tunnels (i.e. crossing points).

Furthermore, we calculate the Moran's I -statistic (Moran 1948) to measure the (global) spatial autocorrelation for each set of GWR residuals. Small values of Moran's I indicate a weak autocorrelation structure in the residuals, which in turn suggests a good GWR fit. For each of the three residual data sets, we calculate the Moran's I values using the same three weighting matrices as in the GWR calibrations (i.e. W_{ED} , W_{ND} and W_{TT} as labelled in Figure 5d). Thus, nine Moran's I values are found in total, as plotted in Figure 5d. Here the Moran's I values from the TT GWR model are relatively always the smallest, suggesting the weakest autocorrelation structures in this model's residual data. Conversely, the Moran's I values from the ED GWR model are relatively always the largest, indicating the strongest autocorrelation structures in this model's residual data. Thus, the choice of distance metric for calculating the Moran's I value itself effects the results in a relative fashion, where the weakest and strongest autocorrelation structures relate to using the TT and ED metrics, respectively.

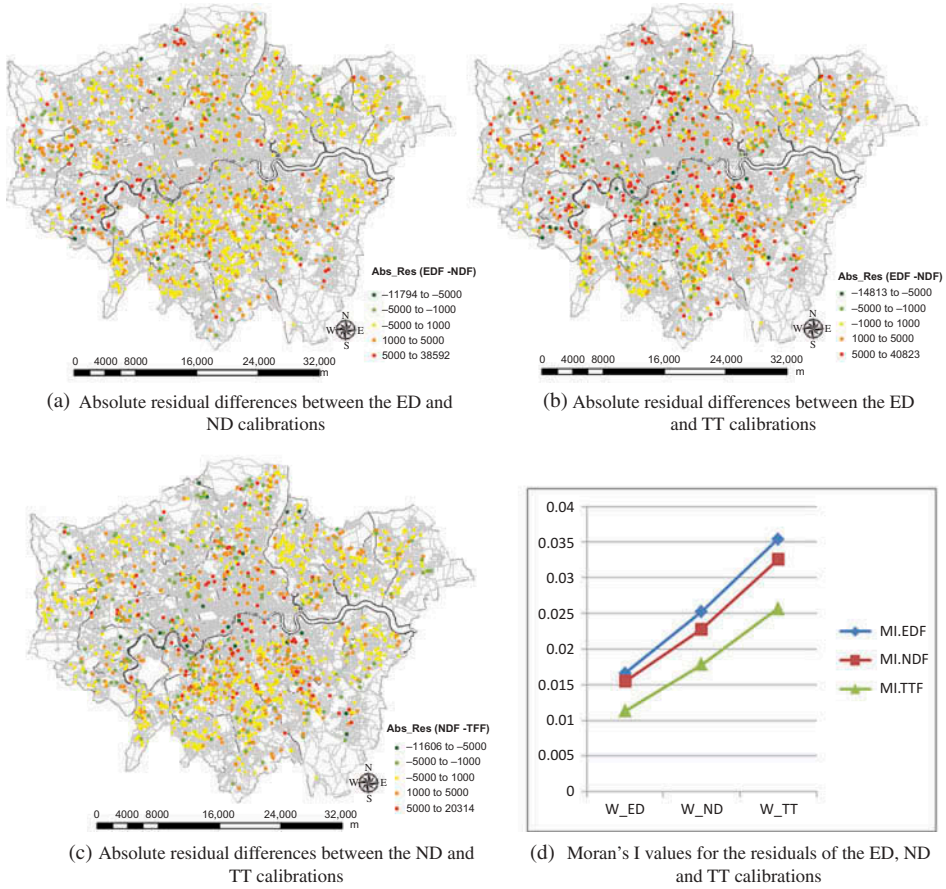


Figure 5. Residual data comparisons for the ED, ND and TT GWR fits for model no. 42. The legend title 'Abs_Res (EDF-NDF)' in (a) means the absolute residuals from the ED GWR model subtracted from the corresponding absolute residuals from the ND GWR model; legend titles in (b) and (c) should be interpreted in the same way. All GWR models are specified with fixed bandwidths.

Observe, however, that if we visualise an extra three point curve, where the GWR calibration and the Moran's I value use the same distance metric (i.e. a plot from left to right that joins the lowest blue diamond to the middle red square to the highest green triangle), the curve is fairly flat, indicating little difference in the autocorrelation status of each residual data set (and conversely, the weakest and strongest autocorrelation structures now relate to using the ED and TT metrics, respectively). Thus, the interpretation of this residual autocorrelation analysis clearly depends on one's viewpoint on which distance metric to use for the Moran's I calculations.

4.3.3. Spatial analysis of the GWR coefficients

GWR is most commonly used in an exploratory fashion, where the local coefficient estimates are mapped to investigate for any change in data relationships across space. As such, it is important to investigate for relative changes amongst the coefficient



Figure 6. London boroughs.

estimates of our GWR models. Here, discrepancy maps can again be produced by subtracting the coefficient estimates for each hedonic variable (and the estimates for the intercept) between each pair of our three GWR models (producing 12 maps in total). For brevity, we only present the three discrepancy maps for the coefficients of *FLOORSZ* (Figure 7a–c). These coefficients reflect the relationship between house price and the size of the property. For context, Table 2 presents the five number summaries for the coefficient estimates of *FLOORSZ* from the three GWR models. Here the (spatial) variation in this coefficient is larger with the two non-ED GWR models than that found with the basic GWR model, tentatively suggesting stronger evidence for relationship heterogeneity.

From the presented coefficient discrepancy maps, it appears that ‘significant’ differences between the basic and the two non-ED GWR models are more widespread when using the TT metric than when using the ND metric (say, differences above 50 or below –50 are considered ‘significant’, i.e. difference classes not coloured yellow in Figure 7a and b). This observation can be attributed to the great diversity in the road network speeds used to calculate the TT metric. For example, speed variations lead to relatively large differences between the ND and TT GWR models, within the boroughs of Waltham Forest, Southwark and Hounslow.

Again the strongest differences can be observed between the basic and either of the two non-ED GWR models along the River Thames, particularly in the West London

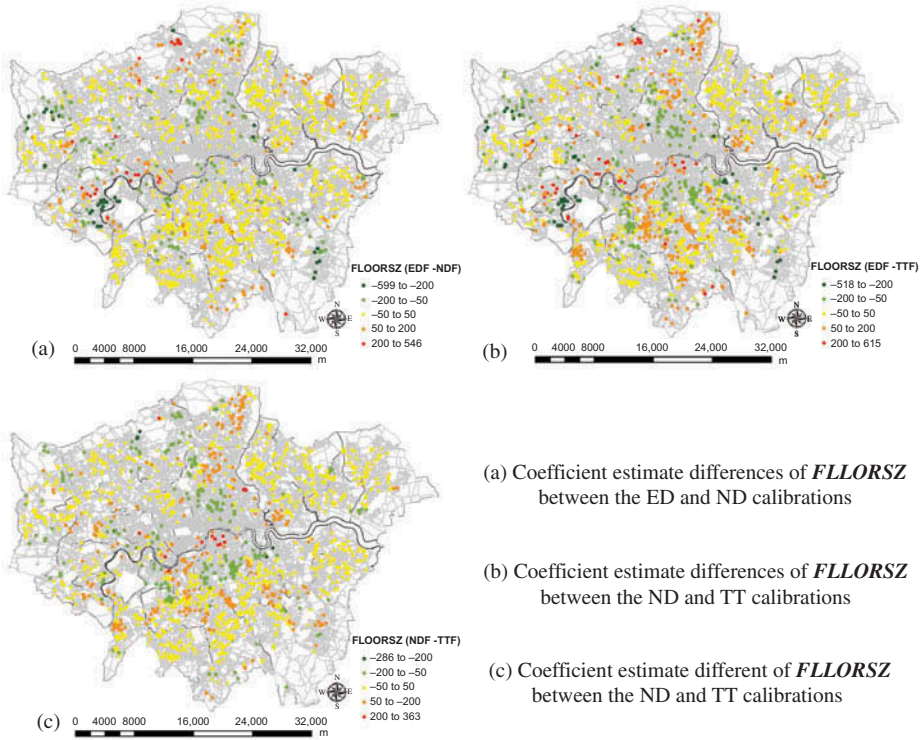


Figure 7. Comparisons of coefficient estimates for *FLOORSZ* from the ED, ND and TT GWR fits for model no. 42.

boroughs. For example in Richmond upon Thames, there is a cluster of large negative differences (classes coloured dark green) in both maps (Figure 7a and b). Thus, the relationship between house price and the size of the property in this area is much stronger using GWR with a non-ED metric than that found using basic GWR (i.e. for non-ED GWR, the results suggest that larger houses are more valuable than that indicated for basic GWR). This may be a consequence of houses north (or west) of the river being more valuable than houses south (or east) of the river. Conversely, weaker relationships can be observed when using GWR with a non-ED metric, such as a few clusters of large positive differences (classes coloured red) down river in the borough of Kensington and Chelsea, where a property’s location can be more important than its size when determining its

Table 2. Five number summaries for the coefficient estimates for *FLOORSZ* for model no. 42 via GWR using ED, ND and TT metrics with fixed bandwidth. The OLS regression gave a stationary coefficient estimate of 1412.9.

Metric	Minimum	1st Quintile	Median	3rd Quintile	Maximum
ED	476.8	1168.0	1348.0	1553.0	3150
ND	503.9	1158.0	1343.0	1552.0	3717
TT	395.6	1133.0	1320.0	1552.0	3668

value. Of course for many areas, this particular data relationship is relatively unaffected by distance metric choice, such as the outer London boroughs of Havering or Sutton (boroughs where difference classes are primarily coloured yellow, in all three maps).

4.3.4. Further observations

Although each GWR model is specified with a different AIC_c defined optimal bandwidth, it is argued that the observed differences in goodness-of-fit and the estimated coefficients (at least for *FLOORSZ*) are fundamentally caused by the distinctive measurements of the ED, ND and TT metrics. Differences in model calibrations and outputs are interrelated where major roles are played by the sample point (housing) density, road density, complexity of road shapes, road speed and the topological relationship between sample points and the road network. A sparse local network with complicated road shapes will lead to heightened differences between ED and ND (or TT) metrics. Examples of such effects can be found in the outer London boroughs of Hillingdon and Bromley (observe difference classes coloured dark green, in [Figure 7a](#) and [b](#)). Natural and man-made barriers like rivers, hills, canals or railways also give rise to major differences between the metrics (and in turn, model outputs). Here, the River Thames plays a dominant role in the calibration and interpretation of our GWR models. In contrast, model outputs tend to be similar where both sample points and road densities are dense and uniformly distributed (and in doing so, provides uniform road speeds for TT measurements). GWR outputs are similar in such instances, as more weight (or importance) is attached to the nearest observations, where the choice of distance metric measuring this ‘nearness’ has little influence. We call this phenomenon a ‘local ED effect’.

5. Summary and discussion

In summary, a GWR model using a TT metric appears most suited to our London house price data set, followed closely by a GWR model using a ND metric. A basic GWR model using an ED metric appears least suited to this data, while all GWR models should use a fixed bandwidth approach in preference to an adaptive one. Thus, for this particular data set, the use of a non-ED metric with GWR has clear merit. In particular, such a model has shown promise in terms of (1) the lowest AIC_c model fit diagnostics; (2) reduced spatial autocorrelation in its residuals; and (3) improved interpretation of relationship heterogeneity via the spatial distribution of its estimated coefficients. Given such model improvements are often small or subtle, the extra effort involved in modelling with the non-ED metrics is still considered worthwhile, since it adds to our understanding of the house price study.

Of course, it is not easy to generalise these results and further empirical work would be useful, possibly combined with Monte Carlo sampling for added clarity. A more objective model assessment is possible within a simulation experiment, where data properties can be controlled. Initial work on this subject is currently underway, where the simulation of local coefficient surfaces that are suited to evaluating GWR models with different distance metrics is proving a challenge. Within a simulation experiment, questions such as (1) which metric provides the most accurate fit and (2) which metric provides the most accurate assessment of relationship heterogeneity can be objectively answered. The simulation experiment needs to be realistic enough to guide distance metric decisions in real case studies.

Future investigations could also experiment with the use of different weighting functions to the Gaussian kernels specified in this study (e.g. bi-square, box-car, etc.). However, it is expected that the overall picture is likely to remain the same, as that found here. Future work could also investigate the use of non-ED metrics, when GWR is used as a spatial predictor. This study's residual analysis suggests that GWR with a non-ED metric may perform well in this respect. However, the use of different optimised bandwidths for GWR with different distance metrics requires close attention, especially if a bandwidth is found via CV. It can be difficult to isolate the specific effects of a non-ED GWR calibration because of the complexity of the non-ED metric calculations. This predictive work could extend that conducted by [Harris et al. \(2010, 2011b\)](#).

Finally, it is possible to take an alternative approach to our study problem by first transforming the coordinate system and then applying the basic (ED) model in this transformed space. The coordinate transform is such that it accounts for the need to use a non-ED metric. Results are then back-transformed to the original coordinates for display and interpretation. Examples for spatial prediction along a river network can be found in [Legleiter and Kyriakidis \(2006, 2008\)](#), and an example for house price prediction can be found in [Brunsdon \(2013\)](#). These alternatives were beyond the scope of our study, as the transformations can be quite involved.³ Future work, however, could compare the approach of [Brunsdon \(2013\)](#), but now, in a GWR context, to those proposed here.

Acknowledgements

The research presented in this article was funded by a Strategic Research Cluster grant (07/SRC/I1168) by the Science Foundation Ireland under the National Development Plan. We thank all the reviewers for their valuable comments and suggestions, which are very important for improving this article.

Notes

1. This may be due to a certain ignorance of its importance or more likely due to a lack of suitable software. In the latter case, this is somewhat addressed with the **GWmodel** R package.
2. Future research could investigate this source of uncertainty more closely, together with its impact on model outputs. Simulation studies may help in this respect.
3. The idea of transforming the coordinate system can also be found in the context of non-stationary spatial covariance estimation. For example, see the coastal water distance study of [Løland and Høst \(2003\)](#) where non-ED metrics are required and the related deformation methods of [Sampson and Guttorp \(1992\)](#).

References

- [Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: *International symposium on information theory, 2nd, Tsahkadsor, Armenian SSR*. Budapest: Akademiai Kiado, 267–281.](#)
- [Ali, K., Partridge, M.D., and Olfert, M.R., 2007. Can geographically weighted regressions improve regional analysis and policy making? *International Regional Science Review*, 30 \(3\), 235–268.](#)
- [Anselin, L., 1995. Local indicators of spatial association–LISA. *Geographical Analysis*, 27 \(2\), 93–115.](#)
- [Bitter, C., Mulligan, G., and Dall'erba, S., 2007. Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, 9 \(1\), 7–27.](#)
- [Bowman, A.W., 1984. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71 \(2\), 353–360.](#)

- Brunsdon, C., 2013. Street value: assessing non-Euclidean metrics in spatial analysis applied to property values. *GISRUK 2013*, 3–5 April 2013 Liverpool.
- 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.
- Burnham, K.P. and Anderson, D.R., 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. 2nd ed. New York: Springer-Verlag.
- Charlton, M., Fotheringham, A.S., and Brunsdon, C., 2007. *Geographically weighted regression: software for GWR*. Maynooth: National Centre for Geocomputation.
- Chin, T.L. and Chau, K., 2003. A critical review of literature on the hedonic price model. *International Journal for Housing Science and Its Applications*, 27 (1), 145–165.
- Cho, S.-H., Jung, S., and Kim, S., 2009. Valuation of spatial configurations and forest types in the southern Appalachian highlands. *Environmental Management*, 43(4), 628–644.
- Cho, W.K.T. and Gimpel, J.G., 2009. Presidential voting and the local variability of economic hardship. *The Forum*, 7 (1), 23.
- Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74 (368), 829–836.
- Csardi, G. and Nepusz, T., 2006. The igraph software package for complex network research. *Complex Systems* [online], 1695. Available from: <http://www.citeulike.org/user/phlow/article/3443126> [Accessed 13 July 2011].
- Curriero, F., 2006. On the use of non-Euclidean distance measures in geostatistics. *Mathematical Geology*, 38 (8), 907–926.
- De Smith, M.J., 2004. Distance transforms as a new tool in spatial analysis, urban planning, and GIS. *Environment and Planning B: Planning and Design*, 31 (1), 85–104.
- Diamond, P. and Armstrong, M., 1984. Robustness of variogram and conditioning of kriging matrices. *Mathematical Geology*, 16 (8), 809–822.
- Díaz-Garayúa, J.R., 2009. Neighborhood characteristics and housing values within the San Juan, MSA, Puerto Rico. *Southeastern Geographer*, 49 (4), 376–393.
- ESRI, 2011. *ArcGIS desktop: release 10* [online]. Redlands, CA: Environmental Systems Research Institute.
- Farber, S. and Páez, A., 2007. A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. *Journal of Geographical Systems*, 9 (4), 371–396.
- Fotheringham, A.S., Brunsdon, C., and Charlton, M., 2002. *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley.
- Fotheringham, A.S., Charlton, M.E., and Brunsdon, C., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30 (11), 1905–1927.
- Gahegan, M., 1995. Proximity operators for qualitative spatial reasoning. In: A. Frank and W. Kuhn, eds. *Spatial information theory a theoretical basis for GIS*. Berlin: Springer, 31–44.
- Gao, X. and Asami, Y., 2005. Influence of spatial features on land and housing prices. *Tsinghua Science & Technology*, 10 (3), 344–353.
- Getis, A. and Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24 (3), 189–206.
- Goodchild, M.F., 1992. Geographical information science. *International Journal of Geographical Information Systems*, 6 (1), 31–45.
- Goodchild, M.F., 2004. The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers*, 94 (2), 300–303.
- Goodman, A.C. and Muth, F., 1978. Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, 5 (4), 471–484.
- Goodman, A.C., 1989. *Topics in empirical urban housing research*. Chur: Harwood Academic, 49–146.
- Guesgen, H.W. and Albrecht, J., 2000. Imprecise reasoning in geographic information systems. *Fuzzy Sets Syst*, 113 (1), 121–131.
- Halvorsen, R. and Pollakowski, H.O., 1981. Choice of functional form for hedonic price equations. *Journal of Urban Economics*, 10 (1), 37–49.
- Harris, P., et al., 2010. The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. *Mathematical Geosciences*, 42 (6), 657–680.

- Harris, P., Brunson, C., and Charlton, M., 2011a. Geographically weighted principal components analysis. *International Journal of Geographical Information Science*, 25 (10), 1717–1736.
- Harris, P., Brunson, C., and Fotheringham, A., 2011b. Links, comparisons and extensions of the geographically weighted regression model when used as a spatial predictor. *Stochastic Environmental Research and Risk Assessment*, 25 (2), 123–138.
- Hernández, D., Clementini, E., and Di Felice, P., 1995. Qualitative distances. In: A. Frank and W. Kuhn, eds. *Spatial information theory a theoretical basis for GIS*. Berlin: Springer, 45–57.
- Hoaglin, D.C. and Welsch, R.E., 1978. The hat matrix in regression and ANOVA. *The American Statistician*, 32 (1), 17–22.
- Huang, B., Wu, B., and Barry, M., 2010. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, 24 (3), 383–401.
- Hurvich, C.M., Simonoff, J.S., and Tsai, C.-L., 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 60 (2), 271–293.
- Jensen, O.P., Christman, M.C., and Miller, T.J., 2006. Landscape-based geostatistics: a case study of the distribution of blue crab in Chesapeake Bay. *Environmetrics*, 17 (6), 605–621.
- Jordan, L., 2006. Religion and fertility in the United States: a geographic analysis. In: *Population association of America 2006 annual meeting*, 30 March–1 April Los Angeles, CA, 56.
- Kent, J., Leitner, M., and Curtis, A., 2006. Evaluating the usefulness of functional distance measures when calibrating journey-to-crime distance decay functions. *Computers, Environment and Urban Systems*, 30 (2), 181–200.
- Kestens, Y., Thériault, M., and Rosiers, F.D., 2006. Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. *Journal of Geographical Systems*, 8 (1), 61–96.
- Kitchin, R., 2009. Space II. In: R. Kitchin and N. Thrift, eds. *International encyclopedia of human geography*. Oxford: Elsevier, 268–275.
- Krause, E.F., 1987. *Taxicab geometry: an adventure in non-Euclidean geometry*. New York: Dover Publications.
- Lee, S., Kang, D., and Kim, M., 2009. Determinants of crime incidence in Korea: a mixed GWR approach. In: *World conference of the spatial econometrics association*, 8–10 July 2009 Barcelona.
- Legleiter, C.J. and Kyriakidis, P.C., 2006. Forward and inverse transformations between Cartesian and channel-fitted coordinate systems for meandering rivers. *Mathematical Geology*, 38 (8), 927–958.
- Legleiter, C.J. and Kyriakidis, P.C., 2008. Spatial prediction of river channel topography by kriging. *Earth Surface Processes and Landforms*, 33 (6), 841–867.
- Leymarie, F. and Levine, M.D., 1992. Fast raster scan distance propagation on the discrete rectangular lattice. *CVGIP: Image Understanding*, 55 (1), 84–94.
- Lloyd, C. and Shuttleworth, I., 2005. Analysing commuting using local regression techniques: scale, sensitivity, and geographical patterning. *Environment and Planning A*, 37 (1), 81–103.
- Loader, C.R., 1999. Bandwidth selection: classical or plug-in? *The Annals of Statistics*, 27 (2), 415–438.
- Löchl, M. and Axhausen, K.W., 2010. Modelling hedonic residential rents for land use and transport simulation while considering spatial effects. *Journal of Transport and Land Use*, 3 (2), 39–63.
- Løland, A. and Høst, G., 2003. Spatial covariance modelling in a complex coastal domain by multidimensional scaling. *Environmetrics*, 14, 307–321.
- Longley, P.A., et al., 2005. *Geographical information systems and science*. 2nd ed. New York: John Wiley.
- Lu, B. and Charlton, M., 2011. Convert a spatial network to a graph in R. In: *The R user conference 2011*, August 16–18 2011 Coventry.
- Lu, B., Charlton, M., and Fotheringham, A.S., 2011. Geographically weighted regression using a non-Euclidean distance metric with a study on London house price data. *Procedia Environmental Sciences*, 7 (0), 92–97.
- Lu, B., et al., 2013. GWmodel: an R package for exploring spatial heterogeneity. *GISRUK 2013*, 3–5 April 2013 Liverpool.
- Lyon, S.W., et al., 2010. Using landscape characteristics to define an adjusted distance metric for improving kriging interpolations. *International Journal of Geographical Information Science*, 24 (5), 723–740.

- Miller, H.J., 2000. Geographic representation in spatial analysis. *Journal of Geographical Systems*, 2 (1), 55–60.
- Miller, H.J., 2004. Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*, 94 (2), 284–289.
- Money, E.S., Carter, G.P., and Serre, M.L., 2009. Modern space/time geostatistics using river distances: data integration of turbidity and e. coli measurements to assess fecal contamination along the Raritan river in New Jersey. *Environmental Science & Technology*, 43 (10), 3736–3742.
- Montello, D.R., 1992. The geometry of environmental knowledge. In: *Proceedings of the international conference GIS – from space to territory: theories and methods of spatio-temporal reasoning on theories and methods of spatio-temporal reasoning in geographic space*. Milan: Springer-Verlag, 136–152.
- Moran, P., 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, 10, 245–251.
- Noresah, M.S. and Ruslan, R., 2009. Modelling urban spatial structure using Geographically Weighted Regression. In: *18th world IMACS/MODSIM Congress*, 13–17 July 2009 Cairns.
- Öcal, N. and Yildirim, J., 2010. Regional effects of terrorism on economic growth in Turkey: A geographically weighted regression approach. *Journal of Peace Research*, 47 (4), 477–489.
- Páez, A., Fei, L., and Farber, S., 2008. Moving window approaches for hedonic price estimation: an empirical comparison of modelling techniques. *Urban Studies*, 45 (8), 1565–1581.
- Páez, A., Uchida, T., and Miyamoto, K., 2002. A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. *Environment and Planning A*, 34 (4), 733–754.
- Páez, A. and Wheeler, D., 2009. Geographically weighted regression. In: R. Kitchin and N. Thrift, eds. *International encyclopedia of human geography*. Oxford: Elsevier, 407–414.
- Sampson, P.D. and Guttorp, P., 1992. Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 87 (417), 108–119.
- Shahid, R., et al., 2009. Comparison of distance measures in spatial analytical modeling for health service planning. *BMC Health Services Research*, 9 (200), 1–14.
- Smith, B.J., Yan, J., and Cowles, M.K., 2008. Unified geostatistical modeling for data fusion and spatial heteroskedasticity with r package ramps. *Journal of Statistical Software*, 25 (10), 1–21.
- Smith, T.E., 1989. Shortest-path distances: an axiomatic approach. *Geographical Analysis*, 21 (1), 1–31.
- Sui, D.Z., 2004. Tobler's first law of geography: a big idea for a small world? *Annals of the Association of American Geographers*, 94 (2), 269–277.
- Svenning, J.-C., Normand, S., and Skov, F., 2009. Plio-Pleistocene climate change and geographic heterogeneity in plant diversity-environment relationships. *Ecography*, 32 (1), 9.
- Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(ArticleType: research-article/issue title: Supplement: Proceedings. International Geographical Union. Commission on quantitative methods/full publication date: Jun., 1970/ Copyright © 1970 Clark University), 234–240.
- Transport for London, 2011. *Intelligent speed adaptation* [online]. Available from: http://www.tfl.gov.uk/corporate/projectsandschemes/7893.aspx#For_developers [Accessed 21 May 2011].
- Vichiensan, V. and Miyamoto, K., 2010. Influence of urban rail transit on house value: spatial hedonic analysis in Bangkok. *Journal of the Eastern Asia Society for Transportation Studies*, 8, 974–984.
- Wheeler, D.C., 2007. Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environment and Planning A*, 39 (10), 2464–2481.
- Wheeler, D.C., 2009. Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. *Environment and Planning A*, 41 (3), 722–742.
- Windle, M.J.S., et al., 2010. Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. *ICES Journal of Marine Science: Journal du Conseil*, 67 (1), 145–154.
- Worboys, M.F., 1996. Metrics and topologies for geographic space. In: *Advances in geographic information systems research II: proceedings of the symposium on spatial data handling*, 12–16 August 1996 Delft. London: Taylor & Francis.
- Yao, X. and Thill, J.-C., 2005. How far is too far? – a statistical approach to context-contingent proximity modeling. *Transactions in GIS*, 9 (2), 157–178.

- [Yoo, E.-H. and Kyriakidis, P.C., 2009. Area-to-point kriging in spatial hedonic pricing models. *Journal of Geographical Systems*, 11 \(4\), 381–406.](#)
- [Yrigoyen, C.C., García, I., and Vicéns, J., 2007. *Modeling spatial variations in household disposable income with geographically weighted regression. MPRA paper.* Munich: University Library of Munich, 1–28.](#)
- [Yu, D., 2007. Modeling owner-occupied single-family house values in the city of Milwaukee: a geographically weighted regression approach. *GIScience & Remote Sensing*, 44 \(3\), 267–282.](#)
- [Yu, D.-L., 2006. Spatially varying development mechanisms in the Greater Beijing Area: a geographically weighted regression investigation. *The Annals of Regional Science*, 40 \(1\), 173–190.](#)