Application of Geographically Weighted Regression to a 19-year set of house price data in London to calibrate local hedonic price models

Ricardo Crespo, Stewart Fotheringham, Martin Charlton

National Centre for Geocomputation
National University of Ireland Maynooth, Co. Kildare,Ireland
Tel. +353 (0)1 708 6196
Fax. +353 (0)1 708 6456
Email: ricardo.crespo@nuim.ie

1. Introduction

In spite of literature on hedonic house price models is quite vast, much more efforts must be made to integrate and to model spatiotemporal data (see Cressie, 1993). Nevertheless, it is worth mentioning that excellent works have been done in recent years to deal with this phenomenon. For example, of paricular interest has been the research on panel data models (Holly et al., 2006) and spatiotemporal autoregressive (STAR) models (Pace et al., 1998).

On the other hand, the development of geographically weighted regression (Brunsdon et al., 1998) has brought new insights into the understanding of spatial dynamics in econometric models. In this respect, we do believe that the incorporation of temporal data into the model and the subsequent development of a spatiotemporal version of GWR will naturally contribute to a better understanding of stochastic processes that vary and interact in space and time.

This is presicely the aim of this study, that is, to develop a spatiotemporal version of the GWR technique, which would enable the forecasting of and eventually the interpolation of local parameters throughout time.

2. The hedonic price model

In this paper we follow the hedonic price model used by Fotheringham et al. (2002). House prices in London are regressed on (1) *house attributes*: floor area, type of the property, date of construction, number of bathrooms, provision of garage and central heating, (2) *socioeconomics characteristics*: proportion of workforce in professional or managerial occupations and rate of unemployment at the census *output area*[1] where the property is located, and (3) *environmental characteristics*: straight-line distance from the property to the centre of London.

Attributes and sale prices of the properties were obtained from a sample of Building Society mortgage records for Greater London from 1980 to 1998. Socioeconomics characteristics, in turn, were obtained from the 2001 census of population in UK.

3. The spatiotemporal bandwidth

In a cross-sectional approach, the hedonic price model is calibrated using GWR at each one of the nineteen years (from 1980 to 1998). Only data points of the same regression-year[2] are included in the model, therefore, bandwidths at each year are calculated independently one another. (see fig. 1).

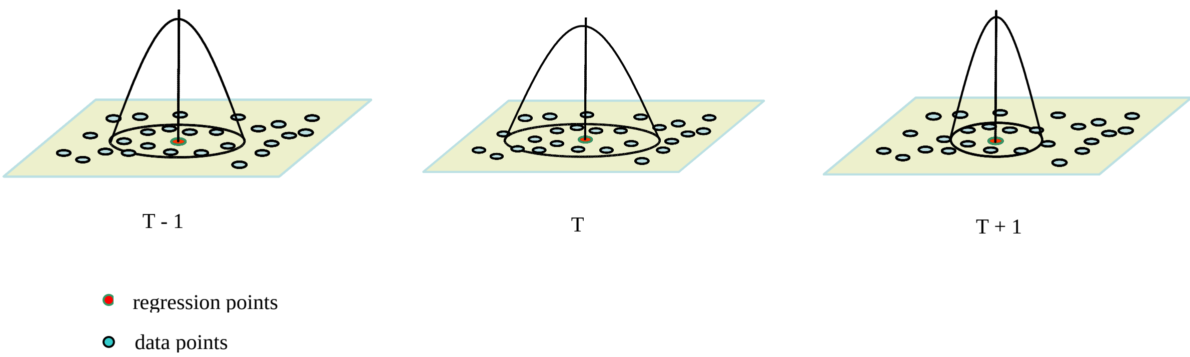

Figure 1. Example of GWR calibrated independently at each year.

[1] Each census output area has an approximate population of 120 residents
[2] We use this term *regression-year* to refer to the year in which GWR is being calibrated

In the case of the spatiotemporal approach, the model is also calibrated at each one of the nineteen years, but in this case, we propose the incorporation of data points from past and future years into the model. From this new perspective, data points are now both spatially and temporally weighted. The combination of spatial and temporal weighting functions leads to what we have called *spatiotemporal bandwidth*, that is, a bandwidth whose size may vary along time and is determined by the selection of appropriate spatial and temporal kernel functions.

In this study we are particularly interested in the analysis of two types of *spatiotemporal bandwidths*: (a) a time-decay bandwidth and (b) an inverse-variance bandwidth.

Fig. 2 shows and example of a possible time-decay spatiotemporal bandwidth. The regression-year is T. As can been observed, spatial bandwidths become smaller as the data points are located farther from the regression-year T.

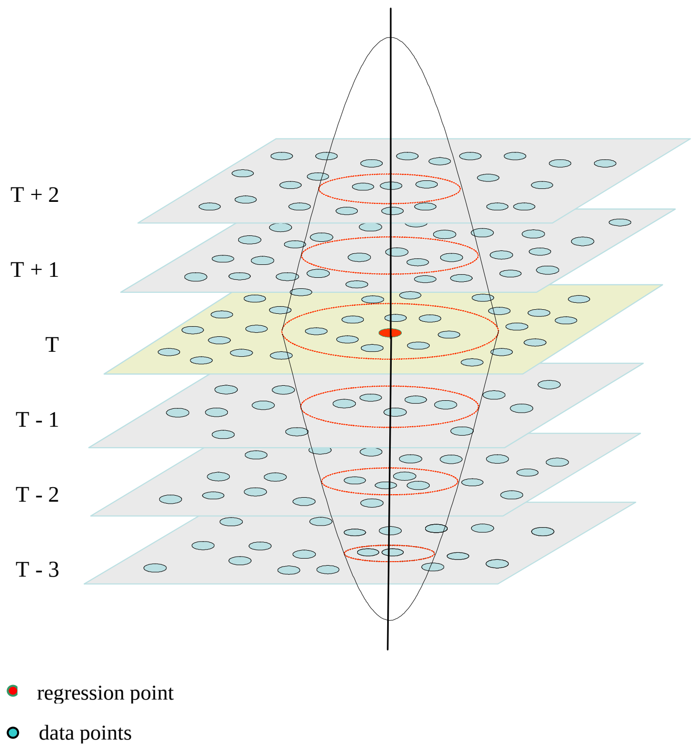

Figure 2. Possible time-decay spatiotemporal bandwidth

Figure 3 shows an example of a possible inverse-variance spatiotemporal bandwidth. In this case, bandwidths vary over time proportionality to the inverse of the variance of the response variable. Thus, the higher the variance of house prices is at each time period, the smaller the size of the bandwidth.

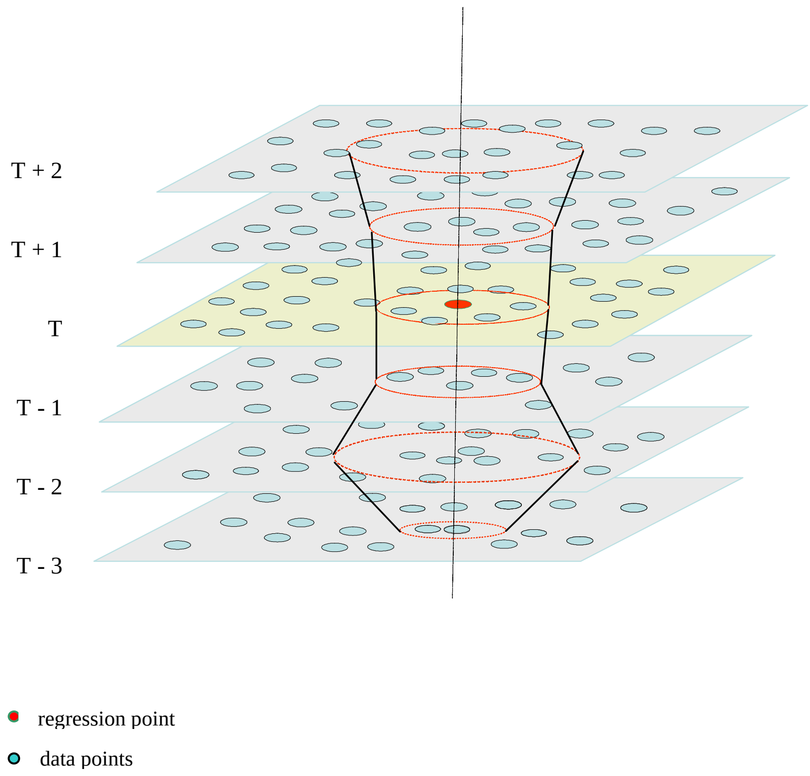

Figure 3. Possible inverse-variance spatiotemporal bandwidth

The interpretation of the parameter estimates obtained by the application of the spatiotemporal bandwidths shown above will naturally differs to one another. For example, if we calibrate the hedonic price model using GWR at the regression-year T by means of both types of spatiotemporal bandwidths, we will be able to analyze the estimated parameters as follows:

a) Time-decay bandwidth:

Since in this case data points located temporally closer to the regression point are weighted more heavily than are data points temporally farther away, it is expected that parameters estimates capture and therefore reflect more locally the spatial and temporal dynamics of the underlying house price process.

b) Inverse-variance bandwidth:

Since in this case data points are partially weighted according to the temporal variability of the response variable, i.e., temporally remote data points from the regression point do not necessarily receive lower weights than do data points located temporally farther away, it is expected that resulting parameters estimates reflect the global trend throughout time and space of the underlying house price process.

3. Final comments

Even though spatiotemporal dynamics might be more difficult to model and probably to interpretate, the incorporation of temporal data into the model will doubtless provide us with useful information about the dynamics of the pricing process. For that reason, we expect to obtain through this new spatiotemporal version of GWR, more statistically robust parameter estimates that would provide not only to geographers, but also to economists with more powerful information for their analysis on hedonics house price models.

4. Acknowledgements

Support for this work is acknowledged to Science Foundation Ireland through the National Centre for Geocomputation.

5. References

Brunsdon C, Fotheringham S, Charlton M, 1998, Geographically weighted regression-modelling spatial non-stationarity. *The Statistician*, 47(3): 431-443

Cressie N, 1993, *Statistics for Spatial Data*. John Wiley & Sons, New York.

Fotheringham S, Brunsdon C, Charlton M, 2002, *Geographically Weighted Regression: the analysis of spatially varying relationships*. John Wiley & Sons, UK

Pace K, Barry R, Clapp J, Rodriguez M, 1998, Spatiotemporal Autoregressive Models of Neighbourhood Effects. *Journal of Real Estate Finance and Economics*, 17(1): 15-33.

Holly S, Hashem M, Yamagata T, 2006, A Spatio-Temporal Model of House Prices in the US. CESifo *Working Paper No. 1826*.