Spatial Statistics 21 (2017) 149–165

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Analysis of residential property sales using space–time point patterns

Lucia Paci [a,*], María Asunción Beamonte [b], Alan E. Gelfand [c], Pilar Gargallo [b], Manuel Salvador [b]

[a] Department of Statistical Sciences, Università Cattolica del Sacro Cuore, Largo Gemelli, 1 - 20123 Milan, Italy

[b] Facultad de Economía y Empresa, Universidad de Zaragoza, C/ María de Luna s/n Campus Rio Ebro, 50018 Zaragoza, Spain

[c] Department of Statistical Science, Duke University, Box 90251, Durham, NC 27708-0251, USA

ARTICLE INFO

Article history:
Received 16 January 2017
Accepted 14 June 2017
Available online 21 June 2017

Keywords:
Empirical coverage
Log-Gaussian Cox process
Nearest neighbor Gaussian process
Nonhomogeneous Poisson process
Ranked probability scores

ABSTRACT

Customarily, for housing markets, interest focuses on selling prices of properties at locations and times. Hedonic models are employed using property-level, neighborhood-level, and economic regressors. However, in hedonic modeling the fact that the locations and times of property transactions are random is ignored. Here, we focus on explanation of the locations of transactions in space and time, viewing them as a point pattern over space and time.

Our contribution is to explain such a point pattern using suitable regressors. We examine two explanatory models, the nonhomogeneous Poisson process and the log Gaussian Cox process. We study a point pattern in the city of Zaragoza, Spain, over the years, 2006–2014. We argue for point level modeling since the process of property sales operates at that scale. We elaborate efficient computation for fitting the foregoing models to the Zaragoza data. We show how the modeling enables rich inference and extraction of novel stories for this market over this time period. In addition, we clarify the potential benefits of this modeling for brokers, buyers, and administrators. To our knowledge, this is the first application of formal space–time point pattern analysis to locations of urban real estate transactions.

* Corresponding author.

E-mail addresses: lucia.paci@unicatt.it (L. Paci), asunbea@unizar.es (M.A. Beamonte), alan@stat.duke.edu (A.E. Gelfand), pigarga@unizar.es (P. Gargallo), salvador@unizar.es (M. Salvador).

http://dx.doi.org/10.1016/j.spasta.2017.06.007
2211-6753/

1. Introduction

Understanding the evolution of real estate markets is a subject of considerable interest both from a theoretical and applied perspective. Housing is an important part of the construction investment, the gross domestic product, and the wealth of a country. Additionally, the evolution of housing markets influences the lending policy of financial institutions (Thwaites and Wood, 2003) for instance, in countries that have a high level of debt involved in a homebuying (e.g., Ireland, Portugal, Spain).

Customarily, interest focuses on modeling of selling prices of properties at locations and times, the so-called hedonic framework of Rosen (1974), traditionally adopted for housing price determination. Hedonic models are employed using property-level, neighborhood-level and economic regressors, and have provided a framework for the use of spatial and spatio-temporal econometric techniques in order to determine the extent to which property prices are related across space (Basu and Thibodeau, 1998; Militino et al., 2004; Banerjee et al., 2004; Gelfand et al., 2007; Bivand et al., 2014, among others) and across space and time (Gelfand et al., 1998, 2003; Pace et al., 2000; Case et al., 2004; Beamonte et al., 2008, 2010).

A key feature ignored in hedonic modeling is the fact that the locations and time of dwelling transactions are random. That is, selling price is modeled conditionally on location and time; randomness is introduced only in the regression given features noted above at the location and time. Understanding the pattern of sales activity over space and time is a different but potentially useful problem. An analogy can be drawn with ecological modeling of species. We may be interested in features/traits of a given species but we are more interested in where species are and why. When we sample species we can go to fixed locations and see what is there. Alternatively, we can create an inventory of a domain and record the (random) locations where species are found. Investigating the pattern of property sales in space and time takes us to the latter setting. Moreover, the papers of Diggle et al. (2010) and Gelfand et al. (2012) on preferential sampling suggest that ignoring this aspect may lead to misleading inference and prediction with regard to the response, here, selling price.

As a result, the contribution of this paper is to focus on explaining the locations and times of property transactions. In particular, for a given market and a given period of years, we have a point pattern of observations over space and time. That is, the number of transactions is random and, given this number, the locations and times of the transactions are random. The goal here is to model such point patterns. The statistical analysis of spatial and spatio-temporal point patterns is a topic which has received much recent attention (Illian et al., 2008; Gelfand et al., 2010; Diggle, 2013; Banerjee et al., 2014) driven by applications to epidemiology, ecology, and the environment. In modeling such point patterns, we consider nonhomogeneous Poisson processes and log Gaussian Cox processes. We demonstrate that the latter are needed to better explain the process.

Nonhomogeneous Poisson processes and log Gaussian Cox processes are currently the workhorses of point pattern analysis. They are attractive in that they are *generative*. They generate a random number of points and then randomly locate these points. In our setting, they model an intensity surface over space and time using regression and, given this surface, they locate points independently over space and time. According to its height, the intensity surface clarifies where and when transactions are more or less likely to occur, providing an informative picture of the nature of market activity.

Under the nonhomogeneous Poisson process, the intensity surface is explained through fixed effects. Under the log Gaussian Cox process, random effects in space and time are introduced to enable suitable local adjustment to the surface provided by the regressors. This local flexibility provides better out-of-sample performance. The absence of these random effects in the nonhomogeneous Poisson process typically leads to over-fitting in-sample but poorer prediction out-of-sample. We offer model fitting within a hierarchical Bayes framework (Møller, 2012; Leininger and Gelfand, 2017). The log Gaussian Cox process models are demanding to fit over our space–time domain. We employ the nearest neighbor Gaussian process (NNGP; Datta et al. 2016a) which introduces convenient sparsity to facilitate such fitting. We consider model adequacy using empirical vs. nominal coverage of predictive intervals (Leininger and Gelfand, 2017) and model comparison using ranked probability scores (Gneiting and Raftery, 2007).

We assert that modeling at the point level is appropriate because that is the scale at which the transaction process operates. Alternative modeling such as aggregating observations to space by time