CJRS (Online)/ RCSR (en ligne) ISSN : 1925-2218
Vol. 33 (2): 57-72

IMPROVING ACCESS TO LAND PRICE DATA: A SPATIAL DECISION SUPPORT SYSTEM FOR CLEANSING LAND REGISTRY DATA

Jamie E. L. Spinney
School of Geography and Earth Sciences
McMaster University
1280 Main Street West
Hamilton, Ontario
Canada L8S 4L8
spinneyj@mcmaster.ca

Pavlos S. Kanaroglou
School of Geography and Earth Sciences
McMaster University
1280 Main Street West
Hamilton, Ontario
Canada L8S 4L8
pavlos@mcmaster.ca

Hugh A. Millward
Department of Geography
Saint Mary's University
923 Robie Street
Halifax, NS
Canada B3H 3C3
hugh.millward@smu.ca

Abstract.
Among the limited sources of land price data in Canada, the Land Registry office, or Registry of Deeds, maintains the population of real estate transactions within their jurisdiction. However, Land Registry data do not permit simple extraction of vacant land prices, they lack neighbourhood and land use information required for analysis of the urban land market, plus the reported sale prices are not verified as fair-market transactions. In response, a geo-referenced database was built within an object-oriented geographic information system (GIS) in order to illustrate a spatial decision support system (SDSS) framework, which was developed to fuse land use and neighbourhood information to Land Registry data, extract sales of vacant land, and verify accurate price information. An illustration of this SDSS framework, using empirical data for Hamilton, Ontario, indicates this SDSS framework provides an effective method for cleansing Land Registry data, thereby improving access to vacant land price data.

Key Words: Geographic information systems, decision support system, data cleansing, Land Registry data, vacant land.

JEL Codes: R1, R3.

Résumé. Améliorer l'accès aux données sur les prix fonciers: Un système de support de décision spatiale pour le nettoyage des données du registre foncier
Parmi les sources limitées de données sur les prix des terres au Canada, le Bureau d'Enregistrement Immobilier, ou d'enregistrement des titres, gère la masse de transactions immobilières relevant de leur juridiction. Toutefois, les données du Registre des terres ne permettent pas la simple extraction des prix des terrains vacants, elles n'ont pas d'informations requises sur les quartiers et sur l'utilisation des terres pour l'analyse du marché foncier urbain, ainsi que les prix de vente déclarés ne sont pas vérifiés comme étant des transactions équitables sur le marché. En réponse, une base de données géo-référencées a été construite au sein d'un système d'information géographique orienté objet (SIG) afin d'illustrer le cadre d'un système de support de décision spatiale (SSDS), qui a été développé pour amalgamer l'utilisation des terres et l'information de quartier aux données du Registre des terres, extraire les ventes de terrains vacants, et vérifier la précision des informations sur les prix. Une illustration de ce cadre SSDS, utilisant des données empiriques de Hamilton, en Ontario, indique ce cadre SSDS offre une méthode efficace pour le nettoyage de données du Registre des terres, ce qui améliore l'accès aux données sur les prix des terrains vacants.

Mots clés : Systèmes d'information géographique, système aide à la décision, le nettoyage des données, données du Registre des terres, terrains vacants.

Codes JEL : R1, R3.

Introduction

Land values represent the economic health of urban areas and regular surveillance of land market behaviour is essential for many private and public sector applications (Cheshire and Sheppard, 2005; Jones et al, 2005). Despite the many and varied potential applications of land price information for land use planning and real estate decision making, there is a dearth of research in Canada using vacant land price data. Contributing to this deficiency is a lack of accessible real estate transaction data, especially vacant land transactions (Clapp, 1990; Wyatt, 1995), and such data may not be in an easily usable form (Thrall, 1998). Furthermore, Feenan and Dixon (1992) suggest that these data accessibility challenges stem from a lack of motivation and a professional environment of secrecy. A simple solution would be to “improve access to, and the dissemination of, property data” (Wyatt, 1995: 69). However, based on our experience, and that of many others, the main barrier to land value research in Canada appears to be political rather than technical.

There are at least three sources of vacant land transaction price data in Canada. Among these sources, municipal and provincial assessment agencies have the most accurate datasets available, because they are, or are supposed to be, regularly visited by an inspector to ensure the accuracy of the database. However, these data may not include detailed inventories of vacant land and they are undeniably difficult to acquire. The latter is evidenced by the alleged “habit of secrecy” (Marin, 2006: 2) that is prevalent within Ontario’s Municipal Property Assessment Corporation (MPAC). Another potential source of vacant land transaction price data is private real estate firms that maintain real estate listings for residential and commercial properties. For example, the largest of these private firms in Canada is the Multiple Listing Service (MLS), which compiles large and detailed datasets of properties that are sold through their listing service. However, MLS data record asking price instead of sale price, are highly skewed towards the residential real estate market, rarely capture vacant land transactions, and are also difficult to acquire due to alleged privacy concerns. The final source of real estate data is Land Registry transaction data obtained from deed transfers, which includes the population of real estate transactions that occurred within the jurisdiction. Beyond access issues, the main problems with using Land Registry data are: they suffer from data quality issues such as data entry errors (e.g. $525,000 vs. $52,500); they contain “non-arm’s-length transactions in which reported prices are significantly below market levels because of a relationship between buyer and seller” (Pollakowski, 1995: 380); and, they lack land use (e.g. residential, commercial, vacant) and neighbourhood (e.g. median income, accessibility to parks) information.

Not surprisingly, since Land Registry data have recently been made accessible in the UK, they have been regularly used to study housing prices (Leishman and Watkins, 2002; Lim and Pavlou, 2007). Land Registry transaction data have also been widely used in the United States, partly due to data value-added resellers (DVARs), who mitigate the impediments to public access and usability (Thrall, 1998) by cleansing and adding contextual or value-added information to the Land Registry data. However, such DVARs are not widely available in Canada and there is currently no mechanism for public access or data cleansing, so Land Registry data are used less often in Canada[1]. Therefore, in order to mitigate the obstacles surrounding data quality, this research seeks to develop a mechanism to cleanse Land Registry data. Spatial Decision Support Systems (SDSS) provide interactive systems designed to assist decision making for a semi-structured spatial problem (Densham, 1991; Geertman and Stillwell, 2003). Similarly, GIS have a long history of being used to support spatial decision making, because decision makers are particularly attracted by GIS capabilities to combine numerous diverse datasets plus the ability to interactively display multiple data themes (Longley et al, 2005).

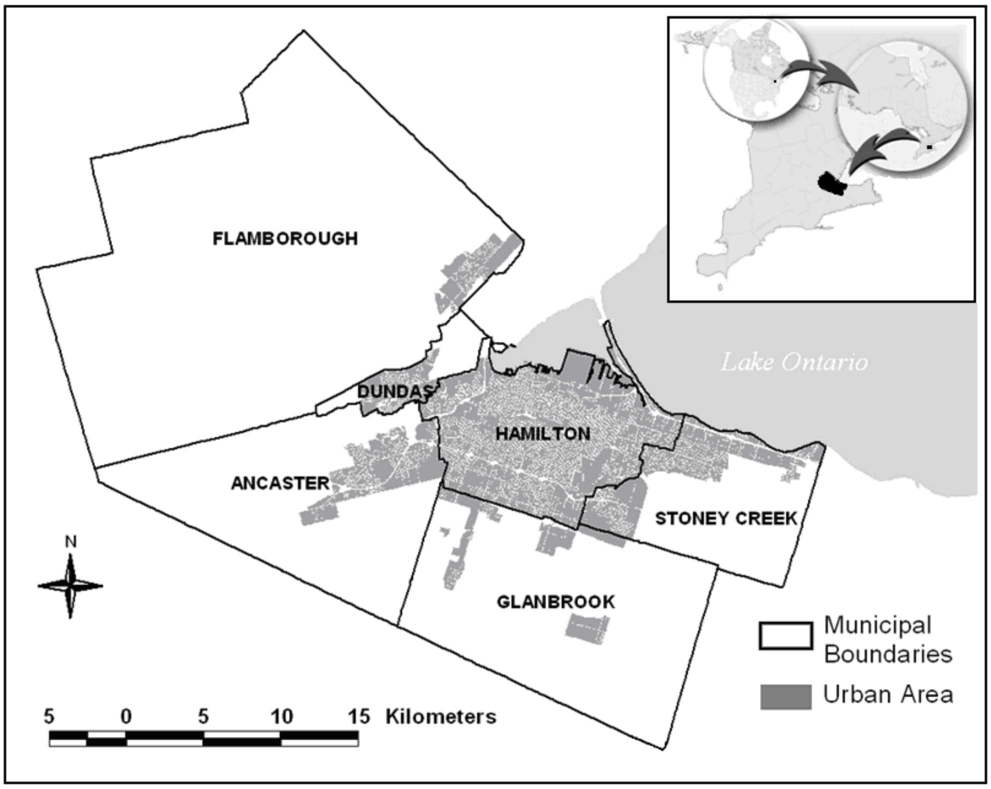
The objective of this research is to use GIS technology to develop an SDSS framework to fuse geospatial information, perform spatial analysis, and provide interactive visualization capabilities in order to cleanse Land Registry data. The purpose of the SDSS framework illustrated here is (i) to apply land use and neighbourhood information to Land Registry data, (ii) to extract vacant land sales from the population of real estate transactions, and (iii) to provide a mechanism to identify and remove erroneous prices. The SDSS framework uses GIS software to build a geo-referenced database by fusing Land Registry data with municipal tax rolls, municipal building permits, and high-resolution imagery. An illustration using Land Registry data for the City of Hamilton, Ontario indicates the SDSS framework was successful in applying land use and neighbourhood information, extracting vacant land sales, and verifying transaction prices, thereby improving access to vacant land price data. The concluding discussion examines the strengths and weaknesses of the SDSS framework and also discusses potential future developments. Descriptions of the digital geospatial data are presented following a brief introduction to the study area.

Study Area

The study area for this research is the amalgamated City of Hamilton, which is located at the south-western tip of Lake Ontario. Hamilton was amalgamated with five other municipalities in 2001, which include Ancaster, Dundas, Flamborough, Glanbrook, and Stoney Creek (see Figure 1). The amalgamated City of Hamilton's total population exceeded 490,000 people in 2001 (Statistics Canada, 2002), making it the ninth largest city in Canada.

[1] A notable recent exception is Provost et al (2006), who examined changes in land values by geomorphology and land use in the Haut-St-Laurent region of south-western Québec.

FIGURE 1 Location and Municipalities of the Amalgamated City of Hamilton



Source: Inset adapted from http://map.hamilton.ca

The municipality of Hamilton is completely urban and its traditional economic engine has been steel and heavy manufacturing industries located along the waterfront, but within the last decade the economy has been shifting towards the service sector. The municipalities surrounding Hamilton are historically agricultural centres, but are increasingly experiencing urban development due to the outward growth (sprawl) of Hamilton. Consequently, Hamilton may be described as exhibiting a typical pattern of urban development for a mid-sized industrial North American city.

Data Description

The City of Hamilton, along with every other municipal government in Canada has access to the digital data required to contextualise the local land market conditions surrounding each real estate transaction. The spatial-temporal digital data used to illustrate this SDSS framework are: (i) Land Registry transaction data; (ii) municipal tax rolls; (iii) municipal building permits; and, (iii) remotely sensed imagery.

Land Registry Transaction Data

Land Registry transaction data represent the population of real estate transactions, because they are based on deed transfers. Land Registry data were obtained from Teranet Inc. (www.teranet.ca), which manages the deed transfers recorded by Ontario's Land Registry Office. A total of 107,793 transactions of real property were acquired for this research and they represent the population of real property sales that occurred in Hamilton between January 1995 and mid-September 2004. These data include details of the transaction, such as the registration date and consideration value (i.e. sale price), and parcel details, such as area, perimeter, length, and ownership type.

Municipal Tax Rolls

The major own-source revenue for Canadian municipalities is the property tax. Annual property tax information is available from every municipal government across Canada, and is becoming increasingly accessible over the internet through distributed geographic information services. Municipal tax rolls for 1996 through 2003 were acquired from the City of Hamilton Finance Department and included three types of information. The first type of information relates to property identification, and included a tax roll number, a civic address, and the school board to which each property belongs. The second type of information refers to the assessed value of each property, and included both total assessed market value and assessed market value by property class. The third type of information concerns land use: there are different sub-classes within the broader land use categories of residential, commercial, industrial, farm, and managed forests.

Municipal Building Permits

A building permit is the basic administrative device used by municipal governments to enforce the laws that relate to the construction, demolition, alteration, addition (e.g. signs, swimming pool, fence), or renovation on a property. Monthly building permit data were acquired from the City of Hamilton's Planning Department spanning the period from January 1999 to September 2004, and were concatenated into a single database. Building permit data provide information about the location, type, and total costs associated with new construction, alterations, additions, and demolitions.

Remotely Sensed Imagery

Given the widespread availability of high-resolution imagery freely displayed for most Canadian cities on Google Maps, especially Street View, plus increasingly affordable pricing from an increasing number of vendors, most researchers and practitioners can afford access to terrestrial or aerial high-resolution imagery. High-resolution aerial imagery was obtained for Hamilton in the form of colour digital orthophotos, which were acquired in 1999 and again in 2002. The 1999 mosaic of colour digital orthophotos had a spatial resolution of 12.5 cm and a spatial coverage of only the built-up (i.e. urban) areas of Hamilton. The 2002 mosaic of colour digital orthophotos had a spatial resolution of 20 centimetres and a complete spatial coverage of the amalgamated city of Hamilton.

Methods

This research employs GIS techniques in order to develop an SDSS framework in order to (i) apply land use and neighbourhood information to Land Registry data, (ii) extract vacant land sales from the population of real estate transactions, and (iii) identify and remove erroneous transaction prices.

Spatial Decision Support System (SDSS) Framework

A geo-referenced database was built within an object-oriented geographic information system (GIS) in order to enable the required functions of the SDSS framework, which include data entry, data processing, data fusion, spatial analysis, and interactive visualization and query capabilities.

Data Entry: The SDSS framework begins with entry of geo-referenced digital data into a common software environment. The Land Registry transaction data, municipal tax rolls, municipal building permits, and remotely sensed imagery were each imported into the SDSS framework, which is currently built within an ArcGIS 9.3 software environment.

Data Processing: Once the data were entered, data processing was used to prepare the data for fusion with the other geo-referenced data. The first data processing operation involved the application of a common geographic projection. Processing the Land Registry data involved stratifying the real estate transactions by the type of ownership, which includes private (84%), condominium (15%), and freehold (< 1%). The remainder of this research is illustrated using only private real estate transactions (N=90,639) that occurred in the City of Hamilton between 1995 and 2004. Processing the municipal tax rolls involved dividing the assessed market value for each property class by the property's total assessed value, thus generating a series of new variables that reflect the percent of market value for each property class (e.g. percent residential, percent commercial, percent industrial), which is especially useful for highlighting mixed land use properties. Processing the municipal building permits involved converting the work type variable from a text variable (with many spelling mistakes) to a numeric categorical variable. After processing operations were complete, the next step in the SDSS framework is data fusion.

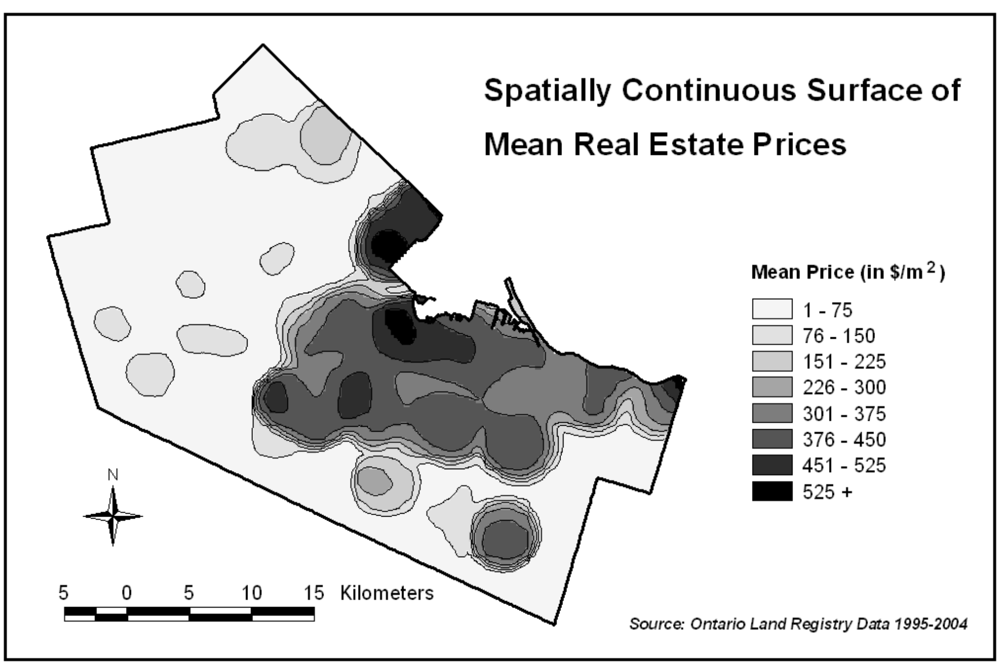
Data Fusion: Data fusion is defined as the process of combining datasets from multiple sources in order to infer information that would not otherwise be available. Data fusion is used here in order to spatially join the Land Registry data with annual tax rolls and building permit data, based on the corresponding location in both space (i.e. address) and time (i.e. month and year), in order to make decisions about land-use and local land market conditions for each transaction. It is important to note that a small percent of private real estate transactions (2.6%) had incomplete address information and did not join with the tax rolls. These transactions were consequently excluded from the following illustration of the

SDSS framework, leaving a total of 88,294 private real estate transactions.

Spatial Analysis: Spatial analysis of the fused datasets was used to generate a series of annual spatially continuous surfaces of mean sale prices per square metre for each property class, which, despite known diminishing returns to lot size (Colwell and Sirmans, 1978), is based on the assumption of a direct linear relationship between property values and lot size. A kernel interpolation technique was used to calculate a distance-weighted average sale price at a given location (using *x* and *y* coordinates) based on the prices of other properties that sold within the search radius (called bandwidth) of the kernel. There are two types of kernels; the first uses a fixed distance for the search radius, and the second uses a fixed number of nearest neighbours (Bailey and Gatrell, 1995). The latter type is known as an adaptive kernel and has the advantage of providing greater detail in areas with dense sales without compromising the estimates in areas with less dense sales. An adaptive kernel was chosen because the spatial distribution of real estate transactions had an obvious clustering in rapidly developing suburban areas and a more dispersed distribution in both rural and densely developed areas.

After temporarily removing repeat sales (keeping the most recent), an adaptive kernel was specified using 30 nearest neighbours and applied to all private real estate transaction prices that occurred in Hamilton between 1995 and 2004. The result was a spatially continuous surface of mean sale price per square metre (Figure 2), which clearly illustrates a polycentric distribution of transaction prices. The urbanised areas have higher transaction prices than the rural communities, with the exception of “prestige” suburban subdivisions.

FIGURE 2 Spatially Continuous Surface of Mean Residential Transaction Prices



The spatially continuous mean sale prices were subsequently fused with the transaction data in order to enable the computation of the spatial price filter ratio (SPFR). The SPFR is the quotient of the local mean price per square metre divided by the sale price of the property in price per square metre. A tax filter ratio (TFR) was also computed by dividing the sale price of each transaction by the assessed value of the property. Both the TFR and SPFR provided valuable contextual information, especially when combined with the remotely sensed imagery, for the interactive processes of vacant land extraction and verification of an arm's length transaction.

Interactive Visualisation and Query: The conventional solution for verifying land use information is to physically visit the site of each transaction and record the land use information. It has been suggested that:

> …it is in no way obsolete to go to the location of the study with the purpose of verifying that the structure of the problem and analysis and any conclusions make sense. This has at times been known as "muddy boots" and today more exotically referred to as "ground truthing" (Thrall, 1998: 52).

In response to the exorbitant costs associated with physically visiting each parcel of land in the study area and the inability to contact the representative actors in each real estate transaction, this paper uses an SDSS framework in order to virtually visit each site to (a) verify land-use information for each transaction and (b) select only transactions with sale prices that appear to reflect arm's length transactions.

Assigning Land Use and Neighbourhood Information

In addition to land use (i.e. property class) information assigned to the Land Registry data by fusing them with the municipal tax rolls, a wide variety of neighbourhood quality information can be used to contextualise the Land Registry data. For example, Statistics Canada's census data, available to the education community through its E-STAT website, can be fused with the Land Registry data to contextualise the social and economic conditions at several different geographic scales. When using census data or any other data that are spatially aggregated into zones, however, it is important to bear in mind that such data are subject to the modifiable areal unit problem (MAUP). MAUP is defined as "the imposition of artificial units of spatial reporting on continuous geographical phenomena [which] result[s] in the generation of artificial spatial patterns" (Heywood et al., 1998: 193). Additionally, there is a wide variety of other point, area, or spatially continuous data sources that can be imported into the SDSS framework to provide valuable contextual information.

Extracting Vacant Land Transactions

Several methods were developed in order to highlight, evaluate, and extract vacant land sales. A small number of commercial and industrial vacant land sales were extracted using property class information garnered from the tax rolls, but the property class information does not permit the extraction of vacant residential land, because all residential properties are, unfortunately, assigned the same tax code regardless of whether they are developed or not. The land registry data were used to extract vacant land sales by querying repeat sales for a dramatic increase in value between successive sales, which indicates there has been development during the intervening period. The land registry data were also used to extract vacant land sales by querying SPFR values less than 0.4, based on the fact that the cost of housing construction is typically at least three times the cost of serviced land. The municipal building permit data were used to extract vacant land sales by querying for demolitions, because demolitions provide the best evidence of vacant land prices in completely developed neighbourhoods. However, it is important to note that the cost of the demolition needs to be added to the transaction price in order to accurately reflect the true price paid for *vacant* land.

All extracted vacant land sales were verified using the colour digital orthophotos to provide visual clues whether to support or contest that the parcel of land was indeed vacant at the time of sale. For example, properties that were developed, demolished, or remained vacant between 1999 and 2002 were quickly and easily identified. For the purposes of this analysis, *developed* properties were defined as those with a building or structure that is not slated for demolition. The aerial imagery was not as effective for transactions that occurred either before 1999 or after 2002, but there was sufficient contextual information afforded by the assessment rolls, approximate age of nearby construction, and nearby sale prices for similar properties to verify that each extracted transaction was indeed vacant at the time of sale. After all vacant land sales were extracted from the population of real estate transactions they were subsequently assigned neighbourhood information.

Verifying Arm's Length Transactions

The conditions of a fair price (i.e. arm's length) transaction are met when a willing buyer and seller make informed decisions under the conditions of a fair sale. The best method of verifying that transaction prices are representative of a fair market conditions is to contact the representative actors. Unfortunately, due to privacy concerns, the names and contact information for the actors in the transaction are suppressed due to privacy concerns. Therefore, an alternative approach was developed that considers the transaction price, the transaction date, the assessed value, and land use information in order to decide whether each sale price is representative of an arm's length transaction. It is important to bear in mind that the price people pay for property is occasionally ill-advised (Skaburskis, 2002), which results in market prices that are well above or below market value. The distinction between market price and market value is an important one. Market *price* is the amount actually paid in a particular transaction, while market *value* is a hypothetical or estimated sale price that would result from careful consideration by the buyer and seller of all data, with primary reliance on those data that reflect the actions of responsible, prudent buyers and sellers under conditions of a fair sale (UBC Real Etate Divison, 2001).

Market prices are reflected by the Land Registry transaction data, while market values are reflected by the total assessed values in the municipal tax rolls. Consequently, a SPFR value near 1.0 provides some support for an arm's length transaction, because it is based on an average of surrounding market prices, but a TFR value reasonably close to 1.0, meaning that the sale price was close to the property's assessed value, provides compelling evidence of an arm's length transaction. The interactive visualization and query capabilities of the SDSS framework also permit detailed inspection of the property in relation to surrounding properties. For example, repeat sales were used to provide contextual information about the *reasonableness* of each successive sale price, especially given any alterations or additions listed in the building permit data. The final decision of whether each sale price was representative of an arm's length transaction was based on an evaluation of the contextual information available.

Results and Discussion

The results of the SDSS framework are presented for extracting vacant land transactions and validating land use information, assigning neighbourhood information, and verifying arm's length transactions.

Extracting Vacant Land Transactions

Using the SPFR and from the building permit data, coupled with colour digital orthophotos for two snapshots in time, the SDSS framework was successful in the identification and validation of 2,524 vacant land sales. This represents 2.9% of all private real estate transactions that occurred in Hamilton between 1995 and 2004. The distribution of both the number and total area (in hectares) for extracted and validated vacant land sales is illustrated in Table 1 by year of sale and land use category.

Results in Table 1 indicate that the number of residential sales appears to be steadily increasing between 1995 and 2002. Only the first four months of 2004 are represented by the transaction data, which explains that year's remarkably lower number of sales. The distributions for other land uses, despite being overshadowed by vacant land for residential use, display a relatively stable number of transactions from one year to the next. Most cities have about 35% of their total area in residential uses (streets excluded), and about 45% of their developed area is residential (Hartshorn, 1992: 220-224). Since residential parcels tend to be much smaller than those for other land uses, however, about three-quarters of all urban lots are typically residential. This helps to explain why the large numbers of vacant lots which become residential (79.4%) only comprise a small percent (7.8%) of the total area. It is important to note that the total area of vacant land in each of the other land use categories shows a very different distribution. In contrast, the small group of farm and managed forest (i.e. resource) properties (5.6%) has a very large total area (76.1%), and those in the industrial and commercial categories are also typically large.

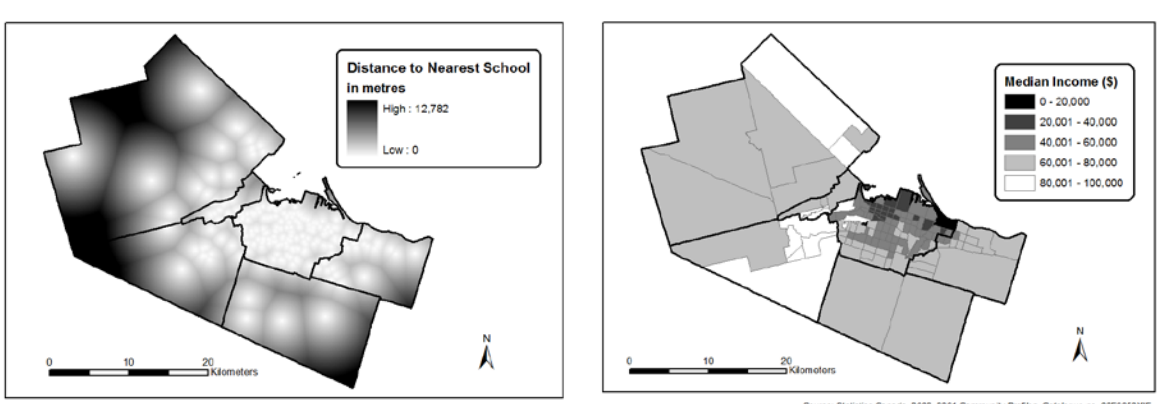
TABLE 1 Temporal Distribution of Vacant Land Sales by Land Use Category

	Residential		**Commercial**		**Industrial**		**Resource**	
Year	***N***	***Area (ha)***	***N***	***Area (ha)***	***N***	***Area (ha)***	***N***	***Area (ha)***
1995	67	10.8	14	10.2	4	2.7	7	168.4
1996	88	15.0	20	17.5	2	0.6	12	186.5
1997	177	19.1	22	81.1	14	23.7	12	390.1
1998	209	32.6	35	37.8	10	8.6	17	402.9
1999	269	29.5	36	10.0	14	13.0	11	225.1
2000	230	44.0	36	37.1	15	11.0	20	427.0
2001	269	23.7	25	6.55	21	34.4	13	327.2
2002	381	49.6	30	13.6	11	5.4	6	164.8
2003	235	40.5	31	17.6	16	11.0	29	264.3
2004	78	20.5	20	7.2	4	2.9	14	234.8
Total	2003	285.3	269	238.7	111	113.2	141	2791.3
Percent	**79.4**	**7.8**	**10.6**	**6.5**	**4.4**	**3.1**	**5.6**	**76.1**

Assigning Neighbourhood Information

Neighbourhood quality information can be fused with Land Registry data in order to provide valuable contextual information for subsequent analysis and modelling efforts. There is a variety of potential variables that can be used to implicitly value neighbourhood attributes (for a review see Cheshire and Sheppard, 1995). Despite the inherent difficulties of quantifying the value of neighbourhood amenities, specific benefits are derived from developing the SDSS framework within a GIS software environment, because of its ability to create location - and distance-related variables "that would be otherwise difficult or time-consuming to create" (Rodriguez et al, 1995: 170). For example, distance to school in Figure 3 is measured using straight-line distances, but a GIS can also compute network distances, network travel times, and even network travel costs could be used to measure the relative and absolute distance to any service.

FIGURE 3 Examples of Neighbourhood Information



Source: Statistics Canada. 2002. 2001 Community Profiles. Catalogue no. 93F0053XIE

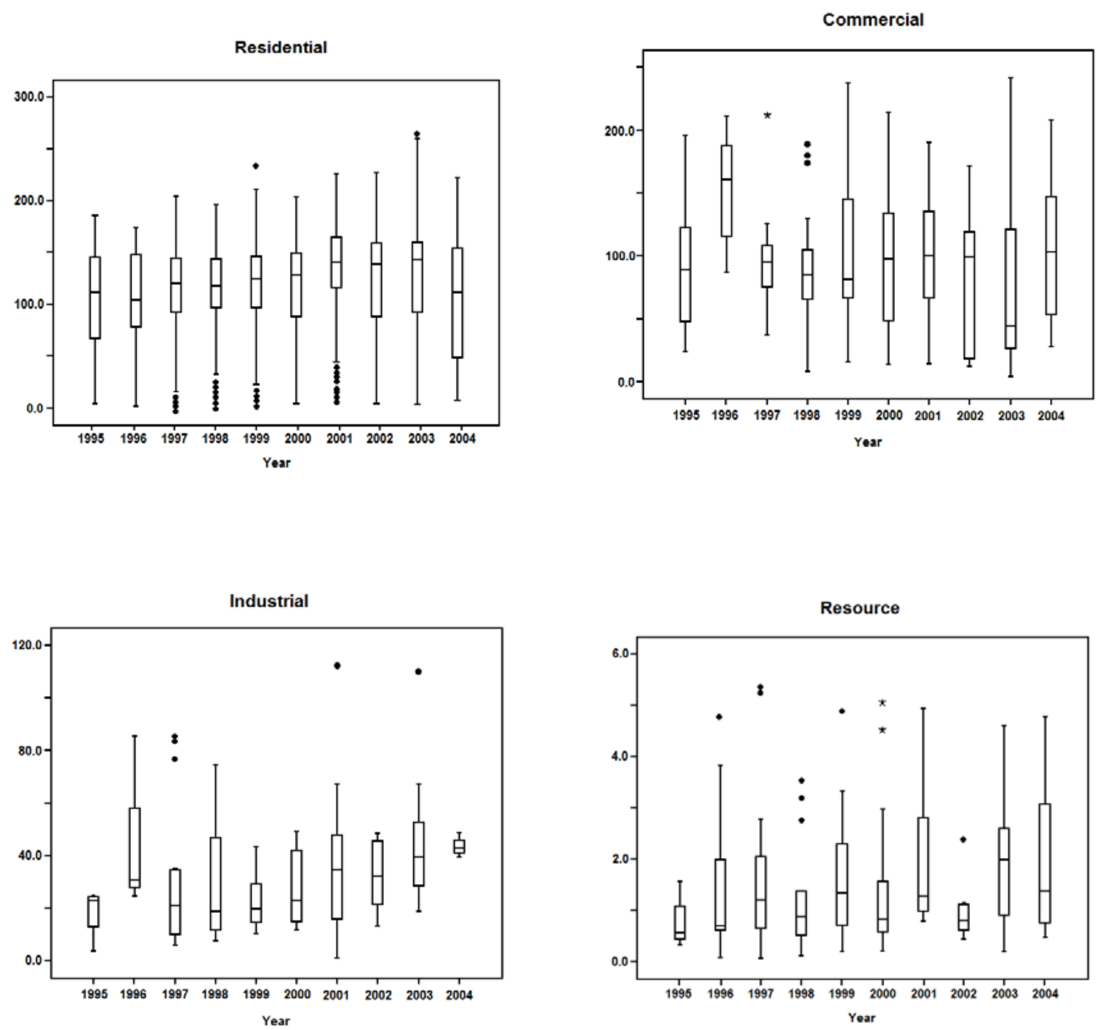
Distance to a school, although often not as important as the quality of the school, acts as a proxy for residential density and approximates the distribution of mean real estate prices in Figure 2. Mean real estate prices in Figure 2 indicate a polycentric pattern of decreasing prices with increasing distance from the city centre. Conversely, median family income exhibits a general increase with distance from the city center. The paradox "that has intrigued students of American cities since the turn of the century"… "is, then, the well-to-do live on cheap land while the poor live on expensive land" (Alonso, 1964: 227). These and other research questions can be explored by adding contextual information to the Land Registry data, but only after validation of sale price information.

Verifying Arm's Length Transactions

The objective of verifying that each sale price represents an arms' length transaction is to eliminate erroneous, and often extreme, values. Based on location and land use it was possible to make a well-informed decision whether or not the sale price appears representative of an arm's length transaction. In fact, the SDSS framework was successful in identifying and removing many erroneous prices, of which many appeared to be typographical but were not altered. The distributions of sale prices verified as arms' length transactions are illustrated in Figure 4 using box plot (i.e. box-and-whisker) diagrams for the main land use categories. These diagrams graphically represent the distribution of the data and are especially useful for identifying outliers and skew. The top and bottom of the box generally correspond to the upper and lower quartile, while the median is represented by a thick line that dissects the box. The vertical lines, or whiskers, are capped with horizontal lines, called adjacent values, and encompass about 99% of the distribution between them. Beyond the adjacent values are outliers, represented by a circle, and extreme outliers, represented by an asterisk.

The mean vacant residential and industrial land prices per square metre in Figure 4 indicate a general trend of increasing land prices over time. The distributions in Figure 4 also suggest that sale prices for vacant land in Hamilton appear particularly "noisy", especially for residential land. However, it is important to consider that vacant land prices vary considerably based on location within the study area (e.g. urban, suburban, and rural). For example, there is a general pattern of decreasing land values with increasing distance from the city centre, which helps explain the apparently large standard deviations indicated in Figure 4. Furthermore, there are small-area effects on land prices as well, such as "prestige" subdivisions, where land prices can far exceed land prices in relatively close proximity, but in a different subdivision. Despite the apparent "noisy" data, it is important to bear in mind that the end-user is likely to apply their own removal of outliers.

FIGURE 4 Distributions of Vacant Land Prices by Land Use Category



Conclusions

Appreciation and depreciation rates for vacant land prices reflect changes in the economic health (i.e. demand) of urban neighbourhoods. Such information is required for land use zoning, eminent domain, and fiscal planning, and for real estate investment and lending decisions. The study of changes in land prices is severely hampered by a lack of access to transaction price data for vacant land. This paper has utilised data management, spatial analysis, and interactive visualisation capabilities afforded by GIS in order to develop an SDSS framework that was used to (i) apply land use and neighbourhood information, (ii) extract vacant land sales, and (iii) identify and remove erroneous prices. The SDSS framework is straightforward and easy to implement for any jurisdiction using most commercial GIS software. The illustration in Hamilton provides convincing evidence that the SDSS framework was successful in effectively cleansing Land Registry data. Until DVARs become more widespread in Canada, the SDSS framework provides a mechanism for improving access to vacant land price data.

While vacant land prices are essential for understanding and monitoring land market behaviour, there is also considerable interest in the most active sector of the real estate market – the residential housing market. While the sale prices can be effectively verified, and the required land use and neighbourhood information can be applied using the data and methods described herein, most residential housing price research applications, with the possible exception of repeat-sales house price index construction, require structural information, such as number of bedrooms and bathrooms. Detailed inventories of structural information are maintained by the provincial property assessment authority, but in the absence these inventories there are alternate approaches that can be used, with varying success, to apply structural information to the Land Registry transaction data. One approach is to use building permit data, not because they provide detailed structural information, but because they do provide geo-referenced time-stamped total price information for new construction. That is, total price information can be used as a proxy for structural attributes by representing the total cost of the structure as opposed to seeking the marginal prices for each attribute. This approach should work well for structures of any size, regardless of their use.

An alternative approach is to infer exterior structural attributes, such as the type, size, and approximate age, from the digital imagery; an approach that can be applied to all land uses. For example, techniques exist for the automated extraction of building footprints using high resolution imagery (e.g. Mayer, 1999; Lee et al, 2003). With keen interpretation, the imagery can also be used to gather subjective structural information, such as the "quality" of the structure and surrounding properties. Structural *quantity* (e.g. square footage, number of bedrooms, number of bathrooms) has been well researched, but research into the impact of structural *quality* has been relatively ignored. Seminal work by Kain and Quigley (1970) used structural quality attributes (e.g. condition of exterior structure, windows, and driveway) to study housing prices and found and they explained as much as structural quantities (e.g. number of rooms, number of bathroom). This approach should work well for single-family residential properties, but not as well for multi-tenant residential, commercial, and industrial properties.

References

Alonso, W. 1964. "The Historic and the Structural Theories of Urban Form: Their Implications for Urban Renewal". *Land Economics* 40 (2): 227-231.

Bailey, T. C. and A. C. Gatrell. 1995. *Interactive Spatial Data Analysis*. Essex, England: Longman Group Limited.

Cheshire, P. C. and S. C. Sheppard. 1995. "On the Price of Land and the Value of Amenities". *Economica* 62: 247-267.

Cheshire, P. C. and S.C. Sheppard. 2005. "The Introduction of Price Signals into Land Use Planning Decision-Making: A Proposal". *Urban Studies* 42 (4): 647-663.

Colwell, P. F. and C. F. Sirmans. 1978. "Area, Time, Centrality and the Value of Urban Land". *Land Economics* 54: 514-519.

Densham, P. J. 1991. "Spatial Decision Support Systems". In: D.J. Maguire, M. F. Goodchild and D.W. Rhind (eds.), *Geographical Information Systems: Principles and Applications*, Essex, England: Longmon, Harlow, Vol.1., 403-412.

Feenan R. and T. Dixon. 1992. *Information and Technology Applications in Commercial Property*. London: Macmillan.

Gatzlaff, D. H. and D. C. Ling. 1994. "Measuring Changes in Local House Prices: An Empirical Investigation of Alternative Methodologies". *Journal of Urban Economics* 35: 221-244.

Geertman, S. and J. Stillwell (eds.). 2003. *Planning Support Systems in Practice*. Heidelberg: Springer.

Hartshorn, T. 1992. *Interpreting the City: an Urban Geography, 2nd Ed*. New York: Wiley.

Heywood, I., S. Cornelius and S. Carver. 1998. *An introduction to Geographical Information Systems*. New York: Addison Wesley Longman.

Jones, C., C. Leishman and C. Watkins. 2005. "Housing Market Processes, Urban Housing Submarkets and Planning Policy". *Town Planning Review* 76 (2): 215-233.

Kain, J. F. and J. M. Quigley. 1970. "Evaluating the Quality of the Residential Environment". *Environment and Planning A* 2 (1): 23-32.

Lee, D. S., J. Shan and J.S. Bethel. 2003. "Class-Guided Building Extraction from Ikonos Imagery". *Photogrammetric Engineering and Remote Sensing* 69(2): 143-150.

Leishman, C. and C. Watkins. 2002. "Estimating Local Repeat Sales House Price Indices for British Cities". *Journal of Property Investment and Finance* 20(1): 36-58.

Lim, S. and M. Pavlou. 2007. "An Improved National House Price Index Using Land Registry Data". *RICS Research Paper Series* 7 (11): 10-14.

Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind. 2005. *Geographic Information Systems and Science*. Chichester: Wiley.

Marin, A. 2006. *Getting it Right: Investigation into the Transparency of the Property Assessment Process and the Integrity and Efficiency of Decision-Making at the Municipal Property Assessment Corporation*. Ombudsman Report, Ombudsman of Ontario. Available online at: http://www.ombudsman.on.ca/media/3283/getting_it_right_20060328.pdf (Accessed January 1, 2011).

Mayer, H. 1999. "Automatic Object Extraction from Aerial Imagery - A Survey Focusing on Buildings". *Computer Vision and Image Understanding* 74 (2): 138-149.

Pollakowski, H. O. 1995. "Data Sources for Measuring House Price Changes". *Journal of Housing Research* 6 (3): 377-387.

Provost, E., A. Ménard, A. Frihida, G. Domon and D. J. Marceau. 2006. "Fluctuations in Land Values in a Rural Municipality in Southern Québec, Canada". *Canadian Geographer* 50 (4): 450-464.

Rodriguez, M., C. F. Sirmans and A. P. Marks. 1995. "Using GIS to Improve Real Estate Analysis". *Journal of Real Estate Research* 10 (2): 163-174.

Skaburskis, A. 2002. "Generational Differences and Future Housing Markets". *Canadian Journal of Regional Science* XXV(3): 353-376

Statistics Canada. 2002. *2001 Community Profiles*. Released June 27, 2002. Last modified: 2005-11-30. Statistics Canada Catalogue no. 93F0053XIE.

Thrall, G. I. 1998. “GIS Applications in Real Estate and Related Industries”. *Journal of Housing Research* 9 (1): 33-59.

UBC Real Estate Division. 2001. Mass Appraisal: The Three Approaches to Value. Vancouver, BC: UBC Real Estate Division.

Wyatt, P. 1995. “Using a GIS for Property Valuation”. *Journal of Property Valuation & Investment* 14 (1): 67-79.