Measurement and Monitoring of Urban Sprawl in a Rapidly Growing Region Using Entropy

Anthony Gar-On Yeh and Xia Li

Abstract
Rapid urban development and dramatic change of landscape have been recently witnessed in some developing countries as a result of rapid economic development. The measurement and monitoring of land-use changes in these areas are crucial to government officials and planners who urgently need updated information for planning and management purposes. This paper examines the use of entropy in the measurement and monitoring of urban sprawl by the integration of remote sensing and GIS. The advantages of the entropy method are its simplicity and easy integration with GIS. The measurement of entropy is devised based on two locational factors—distances from town centers and roads—to capture and reveal spatial patterns of urban sprawl. The entropy space can be conveniently used to differentiate various kinds of urban growth patterns. The application of the method in the Pearl River Delta, one of the fastest growing regions in China, has demonstrated that it is very useful and effective for the monitoring of urban sprawl. It provides a useful tool for the quantitative measurement that is much needed for rapidly growing regions in identifying the spatial variations and temporal changes of urban sprawl patterns.

Introduction
Urban sprawl is the scattering of new development on isolated tracts, separated from other areas by vacant land (Ottensmann, 1977). It is also often referred to as leapfrog development (Gordon and Richardson, 1997). Urban sprawl has been criticized for inefficient use of land resources and energy and large-scale encroachment on agricultural land. There are many problems related to fragmented conversion of agricultural land into urban use. Many advantages have been associated with a larger patch of land use, such as the reduction of environmental cost and development cost (Bilton, 1994). It is found that fragmentation of land use is also harmful to biological conservation. A larger area usually contains a greater diversity of habitats because it provides greater spatial and temporal variation in resources (McArthur and Wilson, 1967; O'Connor et al., 1990). There have been a lot of debates on how to confine urban sprawl and conserve agricultural land resources (Bryant et al., 1982; Ewing, 1997; Daniels, 1997). However, there is a general lack of good indicators to measure urban sprawl and assess their impacts.

This paper examines the use of Shannon's entropy to measure urban sprawl with the integration of remote sensing and GIS. Remote sensing and GIS have been successfully applied in land-use inventory, land-use change monitoring, and assessment of environmental impacts of land development (Jensen et al., 1993; Jensen et al., 1995; Yeh and Li, 1997; Li, 1998). Remote sensing has great potentials in studying urban environments and urban/suburban landscape when high spatial resolution imagery is available (Jensen and Cowen, 1999). An important part of urban monitoring is to obtain information about the geometric elements of urban landscape. Studies have shown that entropy is a good statistic for measuring the spatial distribution of various geographic phenomena (Butty, 1972; Thomas, 1981). In this study, Dongguan, a city in the Pearl River Delta which is one of the fastest growing regions in China, is used to show how entropy can be conveniently calculated using remote sensing and GIS for the measurement and monitoring of urban sprawl.

Methodology
Remote sensing data are capable of detecting and measuring a variety of elements relating to the morphology of cities, such as the amount, shape, density, textural form, and spread of urban areas (Webster, 1995; Mosev et al., 1995). Remote sensing data are especially important in areas of rapid land-use changes where the updating of information is tedious and time-consuming. The monitoring of urban development is mainly to determine the type, amount, and location of land conversion. There are many studies on the use of remote sensing to monitor land-use changes and urban development (Howarth, 1986; Fung and LeDrew, 1987; Martin, 1989; Eastman and Fulk, 1993; Jensen et al., 1993; Jensen et al., 1995; Li and Yeh, 1998). Various techniques have been developed to improve change-detection accuracy, including image differencing (Toll et al., 1980), image ratioing (Nelson, 1983), post-classification comparison (Howarth and Wickware, 1981), the masking method (Pilon et al., 1998), and principal component analysis (Fung and LeDrew, 1987; Li and Yeh, 1998). A conversion matrix which can be generated from change detection techniques is frequently used to indicate the details of land-use conversion. The encroachment of urban areas on other land-use types can be clearly shown in the conversion matrix.

Morphological measurement of urban development is important for land-use planning and the study of urban development (Webster, 1995). Urban planners and researchers are concerned with the change in shape, size, and configuration of the built environment. The measurement of urban form can provide a more systematic analysis of the relationships between urban form and process. Urban economics, transport, and social structure are predicted in spatial terms, and, thus, the effects of such a theory are often articulated through geometric notions involving the shape of urban land use and the manner in which it

Centre of Urban Planning and Environmental Management, The University of Hong Kong, Pokfulam Road, Hong Kong (hdxyugoy@hkucc.hku.hk).
X. Li is also with the Guangzhou Institute of Geography, Guangzhou, P.R. China (lixia@graduate.hku.hk).
spreads (Mesev et al., 1995). Banister et al. (1997) found that there are significant relationships between energy use in transport and physical characteristics of a city, such as density, size, and amount of open space. Batty and Kim (1992) have presented an interesting approach in the measurement of self-similarity using fractal analysis. Fractal geometry can provide a more in-depth insight into density functions than has been available hitherto. It provides ways in which the form of development can be linked to its spread and extent (Mesev et al., 1995). Shape is also a very important element in considering the optimal location of development. It can be used to find the best location for a particular land use using GIS (Brookes, 1997).

The main focus of this study is on the measurement and monitoring of urban sprawl using remote sensing and GIS. Although there have been many studies on the measurement of urban form (Batty and Longley, 1994; Webster, 1995), they have limitations in capturing the characteristics of urban sprawl. In this paper, entropy is used to measure the extent of urban sprawl with the integration of remote sensing and GIS. The measurement is directly carried out within a GIS to facilitate access to its spatial database.

Shannon's entropy ($H_n$) can be used to measure the degree of spatial concentration or dispersion of a geographical variable ($x$) among $n$ zones (Theil, 1967; Thomas, 1981). It is calculated by

$$H_n = \sum_{i=1}^{n} p_i \log(1/p_i)$$

where $p_i$ is the probability or proportion of a phenomenon (variable) occurring in the $i$th zone ($p_i = x_i/\sum x_i$). $x_i$ is the observed value of the phenomenon in the $i$th zone, and $n$ is the total number of zones. The value of entropy ranges from zero to $\log(n)$. If the distribution is maximally concentrated in one zone, the lowest value, zero, will be obtained. Conversely, an evenly dispersed distribution among the zones will give a maximum value of $\log(n)$. Relative entropy can be used to scale the entropy value into a value that ranges from 0 to 1. Relative entropy $H'_n$ is (Thomas, 1981)

$$H'_n = \sum_{i=1}^{n} p_i \log(1/p_i)/\log(n).$$

Entropy can be used to indicate the degree of urban sprawl by examining whether land development in a city or town is dispersed or compact. If it has a large value, then urban sprawl has occurred. The buffer function of a GIS can be used to define buffers or zones from city/town centers or roads, and the density of land development in each of these buffer zones can be used to calculate the entropy.

Because entropy can be used to measure the distribution of a geographical phenomenon, the measurement of the difference of entropy between time $t$ and $t + 1$ can be used to indicate the magnitude of change of urban sprawl: i.e.,

$$\Delta H_n = H_n(t + 1) - H_n(t).$$

The change in entropy can be used to identify whether land development is toward a more dispersed (sprawl) or compact pattern. The following section will discuss how to use the entropy method to measure rapid urban sprawl in one of the fastest growing regions in the Pearl River Delta of China with the integration of remote sensing and GIS.

**Implementation and Discussions**

The study area includes the whole of Dongguan City which is one of the fastest growing regions in the Pearl River Delta of China. It is located north of Hong Kong and Shenzhen and south of Guangzhou on the east bank of the Pearl River Delta. The region has a city proper and 29 towns with a total area of 2,465 km². Before economic reform in 1978, the major land use in the study area was agriculture, especially the growth of paddies. This has undergone rapid changes since economic reform in 1978 (Yeh and Li, 1999). The average annual growth of the gross domestic product (GDP) was 23.3 percent from 1986 through 1996. Recent land development in the Pearl River Delta has led to a substantial change of landscape in both the urban and rural areas. Remote sensing imagery has revealed intensive land-use changes, such as diversified agricultural cultivation, massive urban development, and encroachment of urban land development on agricultural land (Yeh and Li, 1997). In particular, severe urban sprawl has appeared in the '90s as the result of a property boom and land speculation. The urban sprawl is characterized by (1) massive dispersed land development along roads, (2) leapfrog and isolated land development among agricultural land, (3) fast growth around towns instead of around the city proper, (4) the proliferation of idle land, and (5) the unusual high rate of land conversion.

Landsat Thematic Mapper (TM) images dated 10 December 1988, 13 October 1990, and 22 November 1993 were used to estimate the amount of urban expansion and to measure and compare the spatial patterns of urban sprawl during these periods. Landsat Thematic Mapper is one of the currently available remote sensing systems that can provide some of the urban/suburban infrastructure and social-economic information when the required spatial resolution is less than 4 by 4 m and the temporal resolution is between 1 and 55 days (Jensen and Cowen, 1999). TM images have a resolution of 30 m, which is enough to provide information on urban expansion. Principal component analysis of the stacked multi-temporal images technique was used to obtain the land-use classes and land-use changes (Li and Yeh, 1998). It was found that the area for urban land development rapidly expanded from 18,351.4 ha. in 1988 to 19,604.4 ha. in 1990 and to 39,636.4 ha. in 1993. The annual growth rate of urban land use was only 3.4 percent from 1988 to 1990, but it increased tremendously to 34 percent from 1990 to 1993. Excessive use of land resources has been identified as urban land consumption per capita increased substantially from 128.4 m² to 295.8 m² from 1988 to 1993, which were both much higher than the national standard of 100 m². Urban expansion had resulted in the loss of 13.2 percent of its total agricultural land. The loss of prime agricultural land in the course of urbanization has profound impacts on sustainable development in the region.

It can be seen from Figure 1 that land development usually occurs around existing built-up areas and along roads. In order to examine the influences of these locational factors on the spatial pattern of land development, the buffer function of a GIS was used to create buffer zones around town centers and roads in order to calculate the density of land development in each buffer zone (Figure 2). The buffer zones have a width of 250 m, which is about nine pixels in TM images. Too wide a zone will cause the loss of information due to aggregation. Forty-eight buffer zones around each town center and 24 buffer zones around each road were used in the study to ensure consistency of comparison between different towns and different dates of images. Most of the urban sprawl will fall within the range of these zones (Figure 3). The overlay of the urban land-use images on the two buffer images was carried out to find the densities of land development, which is the amount of land development divided by the land area in each buffer zone. The distribution of land development densities over the buffer zones was obtained using the summary function of ERDAS IMAGINE 6.3. The summary function will overlay the buffer zone layer on top of the urban land development layer and find the number (or hectares) and percentage of urban land development pixels for each buffer zone.
Figure 1. Spatial pattern of land development in Dongguan from 1988 to 1993.

Distance decay functions of land development can be observed in the scatterplots of the average density of land development in the buffer zones of the city proper and 29 towns of Dongguan (Figure 3). The density of land development declines rapidly as the distance from roads increases. The pattern of land development away from town centers, however, is slightly different from that away from roads. It can be seen from Figure 3 that the density of land development sharply increases to a peak in a short distance and then declines away from the town center. The relationships between average densities of land development and distances from town centers and roads can be summarized by the following regression equations:

(a) Distance from Town Centers

\[
\overline{DEN}(r) = \begin{cases} 
0.069807 + 0.055268r & (r \leq 1.2000m) \\
0.651583e^{-0.568825} & (r > 1.2000m)
\end{cases}
\]  

(b) Distance from Roads

\[
\overline{DEN}(r) = 0.172950e^{-1.41977r}
\]  

where \( r \) is the distance in meters away from the town centers or roads and \( \overline{DEN}(r) \) is the average density of land development at distance \( r \). It is found that, on average, the density of land development increases steadily from the town centers to around 1.2 km and then declines gradually with increasing distance from the town centers. The decline in the density of land development was much faster from roads than from town centers.

Relative entropy of the two types of buffer zones for each town is used to measure the degree of urban sprawl and monitor its change with time. It is calculated by

\[
H_n^r = \sum_{i=1}^{n}蒲DEN_i\log(\frac{蒲DEN_i}{DEN})/\log(n)
\]

where \( PDEN_i = \overline{DEN}/\sum_i^{n} DEN_i \) and \( DEN_i \) (the observed variable) is the density of land development. Two types of relative entropy, \( H^r_i \) and \( H^r \), which are based on the town buffers and road buffers, respectively, were calculated for each town in 1988, 1990, and 1993. The results show that there is substantial variation in the patterns of urban sprawl among the towns of the study area. In general, urban sprawl is quite obvious in the whole region. The average entropy of urban sprawl from town centers was 0.767 (the maximum is 1) and from roads was 0.806 in 1993. Some towns witnessed quite unusual sprawl of land development, as can be seen from their entropy values. In 1993, the top five towns with the highest entropy from town centers were Fenggang (0.915), the city proper (0.881), Tangsha (0.867), Houjie (0.862), and Qiaotou (0.831). The top five towns with the highest entropy from roads were Fenggang (0.941), Shatian (0.936), Shilong (0.905), Dalingshan (0.900), and Changan
1.0 - 0.888.

Three typical urban sprawl patterns can be identified in the towns of Dongguan from the analysis of entropy from town centers (Plate 1). The first type is *Concentrated Low Development* as represented by Hongmei which had only very limited land development from 1988 to 1993. The second type is *Dispersed Medium Development* as represented by Dalang which exhibited some dispersal away from the town center. The third type is *Highly Dispersed High Development* as represented by Tangsha. It had an upward increase in the density of land development and dispersal of urban development away from the town center. These three types of urban sprawl patterns can be reflected from the entropy. In the first type of urban sprawl, *Concentrated Low Development*, most of the land development is near the town center and the entropy is relatively small. Areas farther away from the town center were not so favorable for land conversion, and most of the land development was carried out only within distances very close to the town center. There is more spread of land development in the second type, *Dispersed Medium Development*. The entropy is higher than that of the first type. For the third type of urban sprawl, *Highly Dispersed High Development*, land development spreads over the urban fringe and to the surrounding rural areas, and the entropy is the highest among the three types of urban sprawl.

The spatial relationship between urban sprawl and the two distance variables can be more clearly identified by using the two-dimensional entropy space (Figure 4). Relative entropy of the town center buffers ($H_T$) and road buffers ($H_R$) of the density of land development in 1993 was calculated. Urban sprawl from town centers will produce a higher value of $H_T$, whereas urban sprawl from roads will produce a higher value of $H_R$. It is found that different urban growth patterns can be identified from these two entropy spaces. It is interesting to note that the towns of Dongguan have quite distinctive growth patterns according to their entropy values. This can help government officials and planners to identify the towns that have irrational development patterns. In Figure 4, the towns that fall within region D have quite dispersed patterns because their entropy is high for the two distance variables. Their highly dispersed patterns of urban sprawl should be of major concern to the city government.

The temporal change of spatial patterns of urban development can be easily measured from the change of entropy using Equation 3 (Figure 5). The increase in the value of entropy indicates that there is an increase in urban sprawl, and develop-
Plate 1. Three typical urban sprawl patterns. (a) Concentrated low development type. (b) Dispersed medium development type. (c) High dispersed high development type.
Figure 5. The increase of relative entropy and urban sprawl patterns (a) from 1988 to 1990 and (b) from 1990 to 1993.

Figure 5. The increase of relative entropy and urban sprawl patterns (a) from 1988 to 1990 and (b) from 1990 to 1993.

Entropy tends to be more dispersed. It was found that the average increase in entropy from town centers was only 2.7 percent for the whole region from 1988 to 1990. This means that there was only a slight trend of urban sprawl in this period. However, the average entropy increased by 8.9 percent from 1990 to 1993. This is the result of rapid land development and a property boom in southern China from 1992 to 1993 as influenced by the property boom in Hong Kong (Yeh and Li, 1999). Figures 5a and 5b show the variation in the increases of entropy from town centers from 1988 to 1990 and from 1990 to 1993, respectively. There were some towns with negative figures during these periods, indicating relatively more concentrated development rather than dispersed development. These towns were in the minority among the 29 towns of Dongguan. There were 11 towns from 1988 to 1990 and only four (including the city proper) from 1990 to 1993 that had negative increases. This indicates that the majority of development is towards dispersed development rather than concentrated development.

Zhangmutou, a town which is deeply involved in the development of property markets for Hong Kong's residents, had an increase of entropy as high as 24.1 percent from 1990 to 1993, as compared with an average of 8.9 percent for the whole region. Two other towns, Tangsha and Qingxi which are very close to Hong Kong, were also among those with a high increase of entropy with rates of 18.1 percent and 24.9 percent, respectively. Dalingshan and Dalang, which are located in the hilly areas, however, also witnessed a high increase of entropy from 1990 to 1993. The development sites in these two towns are scattered in the satellite images. The four towns that had a negative increase of entropy from 1990 to 1993 were Wangniudun, Daqiao, Hongmei, and Xianwan. This means that land development in these towns was less dispersed during this period.

Entropy is a more robust spatial statistic. Many studies have shown that spatial dispersal statistics, such as the Gini coefficient, are often dependent on the size, shape, and number of regions used in calculating the statistics, and the results can change substantially with different levels of areal aggregation (Smith, 1975). This is a manifestation of the scale problem or modifiable areal unit problem which may exert unspecified influence on the results of spatial analysis (Openshaw, 1991). Thomas (1981) indicates that relative entropy is better than traditional spatial dispersal statistics because its value is invariant with the value of n, the number of regions. However, relative entropy is still, to some extent, sensitive to the variations in the shapes and sizes of the regions used for collecting the observed proportions. For example, if there are two scales of analysis for the dispersion of population in a country, such as regions and sub-regions, different entropy values will be obtained if the data are collected based on regions instead of sub-regions. The entropy decomposition theorem can be used to identify different components of the entropy that are related to different zone sizes in collecting the data (Batty, 1976; Thomas, 1981). It can alleviate the problem for comparing the results between different zone sizes because the influence of scaling can be exactly measured. If the larger zones (regions) are desegregated into smaller zones (sub-regions), the resulting entropy can be accurately estimated by using the following equation (Theil, 1972; Thomas, 1981):

\[ H_n = \sum_{j=1}^{m} p_j \log(1/p_j) + \sum_{j=1}^{m} \left( \frac{n_j}{p_j} \right) \log \left( \frac{p_j}{p_j} \right) \]  

where \( j \) is the \( j \)th zone at the region scale, \( m \) is the total number of zones at the region scale, \( p_j \) is the proportion at the \( j \)th zone at the region scale, \( p_j/n_j \) is the proportion at the \( j \)th sub-region within the \( j \)th region, and \( n_j \) is the total number of zones of sub-region at the \( j \)th region. The equation can be written in a simpler form: i.e.,

\[ H_n = H_m + H_{m/n} \]  

Entropy is decomposed into two terms with the first \( (H_m) \) measuring variations between regions and the second \( (H_{m/n}) \) measuring variations within regions. The equation indicates that the increase in the number of zones (with a smaller zone size) will cause the increase in the entropy value because of the gain of information within smaller sub-regions.

To test the robustness of the entropy methodology, we used four zone sizes, namely 6, 12, 24, and 48 buffer zones, from town centers to calculate relative entropy in 1993. A smaller
number of zones is obtained by using a larger buffer distance and vice versa. The calculation is based on the density of land development that is obtained by overlaying these buffer zones on the 1993 urban land-use image. Figure 6 shows the results of the test in which the relative entropy values of the three towns (the three typical urban sprawl patterns in Plate 1) are plotted against the number of zones used in the test. It can be seen that relative entropy from town centers increases as the number of zones (smaller zone sizes) increases. However, the lines are almost parallel with each other with a similar slope. Despite the fact that the relative entropy of a town will increase as the zone size gets smaller, the relative differences in entropy among the towns will not change if the same zone size is used to calculate the entropy. This shows that the entropy method is quite robust and can well handle the scaling problem if the same zone size is used.

**Conclusion**

Much in the literature has discussed the negative effects of the proliferation of urban sprawl. However, there is a general lack of research on the development of methods and indicators to quantify and study urban sprawl. There is also an urgent need to provide government officials and planners a scientific method to measure and monitor urban sprawl, especially in rapidly growing cities in developing countries where urban sprawl has caused many environmental problems. This paper has demonstrated how entropy can be used to measure, monitor, and identify spatio-temporal patterns of urban sprawl by the integration of remote sensing and GIS. TM images are able to provide the desired information for monitoring urban sprawl patterns when the study area is large and land use changes quickly.

The entropy method can be easily implemented within a GIS to facilitate the measurement of urban sprawl away from urban centers and away from roads. The study shows that entropy is a good indicator for identifying the spatial problems of land development. It is capable of identifying which town has a better spatial efficiency for land development in terms of compactness. Furthermore, the two-dimensional entropy space can be used to differentiate various growth patterns more clearly. The method can provide useful information for government officials and planners to monitor the land development process and to identify land-use problems.

The scaling problem or modifiable areal unit problem exists in spatial indicators. The values of entropy are still influenced by the sizes of zones. However, as can be seen from the study, although entropy of a town will increase as the zone size gets smaller, the relative differences in entropy among the towns will not change if the same zone size is used to calculate the entropy for different towns or time periods. Entropy is a quite robust method for comparing urban sprawl patterns.

The application of the method reveals that the study area has experienced severe urban sprawl in the '90s with the lack of proper development control and management. The spatio-temporal pattern of urban sprawl is identified by calculating the entropy from multi-temporal TM satellite images. Some towns in the study area have an unusually high degree of urban sprawl due to land speculation. Government officials and planners can use the method and results of the study to control and monitor such development patterns for protecting the environment and making better use of land resources at the urban fringe areas. Entropy is a good indicator for the measurement and monitoring of land development and the identification of urban sprawl patterns in the region. The method is also applicable to other fast growing regions in the world that are facing similar urban sprawl problems.

**Acknowledgments**

This study is supported by the funding from the Croucher Foundation, Hong Kong. The authors are most grateful to the valuable comments from the four anonymous referees in helping us to improve the paper.

**References**


