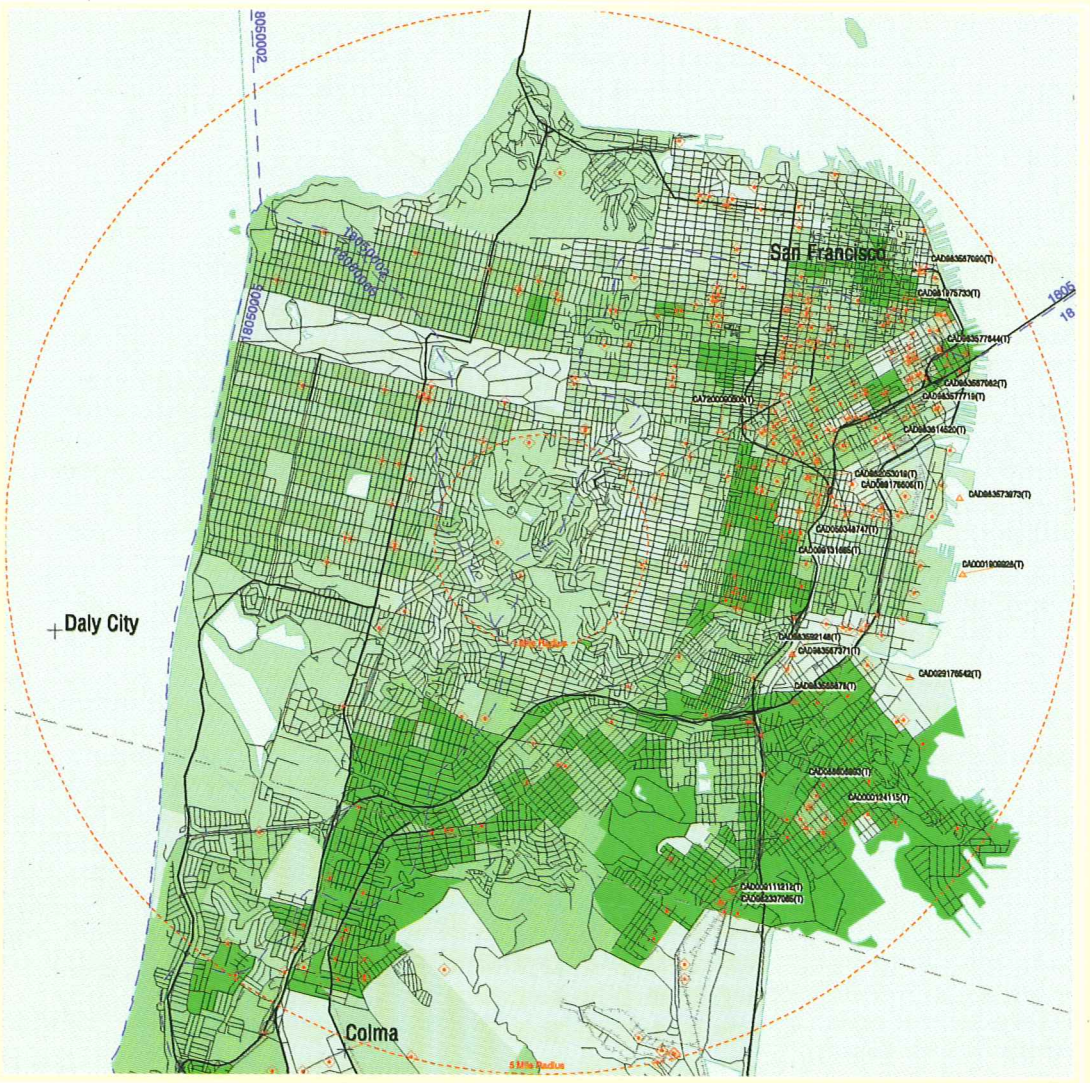
VOLUME 3 NUMBER 1-2 (ISSN 1082-4006) DECEMBER 1997

GEOGRAPHIC INFORMATION SCIENCES

地 理 信 息 科 学

第三卷 第一至二期 一九九七年十二月



A Journal of The Association of Chinese Professionals in Geographic Information Systems (Abroad)

国家科委副主任徐冠华院士在“纪念中国海外地理信息系统协会创会五周年学术研讨会”上的讲话

（根据录音整理）

各位来宾：

首先我代表国家科委对中国海外地理信息系统协会成立五周年表示热烈的祝贺，为海外学子与国内地理信息系统工作者的密切合作，发展中国地理信息系统事业所做出的努力表示衷心地感谢。

刚才陈宜瑜副院长做了很重要的讲话，其他同志也做了很好的发言，听到这些发言深受启发，同时也有了一个问题，就是我讲什么好？大家讲的都很全了，我只好讲一些也许与CPGIS不是非常相关，但我认为还是需要讲的话。第一个问题是：如何进一步发挥海外学人的作用问题。大家都承认，当前世界的竞争主要是经济竞争，归根结蒂是国家技术创新能力的竞争。人才在科学技术竞争中占有重要地位，特别是科学研究工作，是创造性的智力活动，人才决定科技的前途。我经常谈这样一个观点，一流的人才创建一流的研究所；二流的人才只能创建二流的研究所，而不管你研究所有多好的设备，有多少资金和经费支持。从这个意义上来讲，科学技术的竞争，最终是人才的竞争，所以人才在国家的经济、科技发展中具有突出的重要性。这是人才的重要。另外还想讲人才难得：一个国家的人才结构象是一个巨大的金字塔，接受普及教育的，有几亿人，构成了巨大的金字塔塔基，是大量财政、人力投入的结果。各种因素，包括经济、社会、个人因素，决定金字塔越往上越小，受过高等教育、研究生教育是很少一部分人，这些尖子人才的诞生，是大量普及教育的结果，即大量人力、财力投入的结果。因此很难用培养一人才花了多少钱来衡量，我总认为出国交多少个人培养费就解决问题了？实际上每一个人才都是众多人大量投入的结果，在这个基础上才能产生出尖子，这些人对国家的发展有突出的重要性。我不太同意经常谈的“三个臭皮匠，顶个诸葛亮”的说法，数量和质量不能相比，一个就顶一个，这个人的作用是多少人也顶不了的，就是这一个人的作用。因此人才是很难得的。当前有很多急需人才流失到国外，我深感痛心，我不是说我们应该采取行政的手段把关卡死，不要出国，这是非常愚蠢的做法，和我国的开放政策不相适应，不利于科技发展，也不能为大家所接受。关键就在于我们要采取相应政策，一方面千方百计地留住人才，另一方面要充分发挥已在海外人才的作用。应当把海外人才作为我国人才队伍的一个重要的组成部分。这不仅仅因为这些同志有着很深厚的爱国主义的传统，而且在国家的科技发展当中，发挥着重要的作用，特别是在传播先进国家的科学技术知识、经验方面发挥着不可代替的作用。因此应当充分肯定在国外工作的同志同样是我国科技工作的一支重要力量，要充分发挥他们的作用。

第二个问题我想谈一谈从五年的CPGIS的活动当中得到的启发。回忆起一九九二年在布法罗我们召开了CPGIS的第一次成立大会，当时参加的人有几十人，场面也很热闹，但是我并没有对CPGIS能发挥多大作用有充分的思想准备。五年后，实践证明CPGIS对于中国和海外在GIS方面的交流、促进中国GIS发展确实发挥了重要的作用，有几点体会很深：第一，是海外学人的高度的爱国主义精神。这种精神是中华民族延续几千年不衰的力量所在，是中华民族在危难、困难时表现出高度凝聚力的力量源泉。这些海外学人身在海外，心系祖国，很多事情令人感动。CPGIS的几任会长林珲、宫鹏、李斌，他们不断给我写一些信，就中国GIS的发展、GIS产业的前景提出了很多重要的意见，他们虽然身在海外想到的是祖国的建设、祖国的科技发展。要出版CPGIS科技刊物，除了少量的会费以外，基本上没有经费来源。只好千方百计筹集经费来做，要没有点精神力量是不可能的，从中确实体现了爱国主义精神。第二，从CPGIS的活动中体现了科学家的开拓精神，这不仅在科技领域，而且也在CPGIS活动组织方面，也是有创造性的。五年当中，CPGIS采取了多种形式促进国内与海外的学术交流，包括组织学术讨论会、组织学术培训班、参加国内GIS软件测评和其他工作；另外还探索组织合作研究机构等。我想这些探索对于如何发挥海外学人在国家的科技、经济建设中的作用有重要的参考价值，而且通过进一步的实践，会有更多的新鲜经验，发挥更大的作用。第三，在CPGIS工作中体现的求实精神。所做的事都是团结促进海内外GIS的系统和产业化发展，不是做表面文章、喊口号，这和有些海外的协会，选个会长、发个公告，然后名存实亡，大不一样，这也非常可贵。今后应该继续发扬这种精神，在进入下一个世纪的时候，CPGIS会为加深国内外GIS的相互了解，加强中国GIS研究和产业的发展做出更大的贡献。第三，鉴于CPGIS在过去五年中所取得的成绩，我想借这个机会呼吁国家有关部门应当采取实际的措施来扶持CPGIS的活动，包括组建联合研究机构，包括准备开展的项目，促使海外科学家能够每年有几个月的时间，或者采取其他形式参与国内的科研工作，为促进中外GIS的交流和发展中国的GIS产业做出新的贡献。也为发挥海外学人在科技和经济建设中的作用积累更多的经验。

（一九九七年六月）

Future Research on Application of GPS/GIS/RS for Farmcrops Temporal Arrangement

Deren Li*, Zequn Guan* and Xiufeng He†

*Institute of Remote Sensing and Information Engineering
Wuhan Technical University of Surveying and Mapping, Wuhan, 430070, China
†Department of Automatic Control
Nanjing University of Aeronautics and Astronautics, Nanjing, China

Abstract

Currently, farmcrops temporal arrangement is constrained by the generally inadequate treatment of spatial simulation in terms of socioeconomic and ecological information, resulting in artificially deflected planning. For farming in the future, positional information is of particular importance. In this paper, a description is given of the farmcrops temporal arrangement method, to reduce errors and time in precision farming, based on GPS/GIS/RS and planning techniques currently being used in many institutions together with respective approaches, usability, and trends. The system, being presented here, appears to be particularly suited to data processing and data analysis incorporating image segmentation with location sensing, which will help to combine the advantage of small field sample locations with large-scale, cost-efficient image processing methods.

摘 要

目前，由于依据社会经济和生态信息来正确模拟农作物在特定空间上的生长过程还受到定位信息不足等限制，因而容易导致对农作物的时间安排受人为因素影响而产生偏差。这表明，农业的进一步发展，越来越需要定位信息。本文基于GPS/GIS/RS技术及新近在其它方面已被采用的规划技术，从减少精细农业的失误和缩短耕作时间出发，提出了一种新的农作物时间安排方法以及相关的系统。该系统的目标是在数据处理和数据分析中综合运用定位传感和图像分割技术，以便能将大范围而有价值的图象处理方法与小范围的精细定位采样结合起来。

I. INTRODUCTION

Since China embarked on economic reform and opening up to the outside world, her agriculture has seen relatively rapid development, and it has since produced abundant farm products to meet general social demands. However, from a long-term point of view, China's agriculture is not only confronted with multiple pressures such as growth of population but greater demands on farm products by accelerated industrialization . The prospects brook no over-optimism.

How shall farm products and increase of incomes of peasants be effectively supplied? One of the valid ways is to develop agriculture from a weak-quality industry to a highly effective industry, and it is the key for us to effect the process through the introducing of scientific methods and management , one of whose nature is to handle the precision information of farming.

Precision farming problems are inherently linked to planing problems. To create better Planning Support System(PSS) that will begin to address these complex issues, the changes of farmcrops and the relationships between them and geo-referenced information must be positioned, recognized and understood. One approach to meeting these needs is the integration of remote sensing (RS) image processing, Geographical Information System (GIS), farmcrops inventory, and Global Positioning System (GPS) for an operational farmcrops temporal arrangement.

Lachapelle et al (1994) outlined the precision farming objective of optimizing field potential based on information known about the field. Hermann et al (1995) pointed out that the major application of positioning with GPS/DGPS is to be seen in the area of local information and documentation. Chen et al (1992) developed a portable information manage-

1082-4006/97/0301~2-1$3.00

ment system which integrated the remote sensing (RS), GIS and GPS with basic interface screen overview.

Although the integration of RS, GIS and GPS has been extensively studied in the past, little has been done in developing some adaptive algorithms for incorporating image segmentation with location sensing, and farmcrops temporal arrangement.

The PSS discussed in this paper is intended to operate in a WINDOWS environment.. The PSS graphically integrates digital images and map information with observed positions. In general, the PSS automatically plots a position onto a digital image or map. The real value of the PSS will be realized when data analysis tools permit user to "see" further into the growing digital database of spatial information that is possible using traditional cartographic renditions of the data.

New analytical tools that extend our vision into location sensing and extract the useful information needed to practise farmcrops temporal arrangement are described in this paper.

II. BASIC FUNCTIONALITIES

An application system is proposed for the farmcrops temporal arrangement project. The system under consideration consists of five major functional blocks shown in Figure 1. A standard Microsoft WINDOWS graphical user interface is implemented for PSS system equipped with mouse or other interface tools. A brief discussion will be provided for each function in the following paragraph.

1. Data processing. All the data process of the observation and images until one gets the final farmcrops planning can be arranged in the following steps:

1). To input the multi-temporal images data from the user.

2). To determine if there are significant changes, either increases or decreases, occurring in the images.

3). To classify the images.

4). To document the features which could characterize the change areas of farmcrops, such as size, time, speed, domain and so on.

2. GPS. A GPS unit concerns:

1). To feed field measurement information into

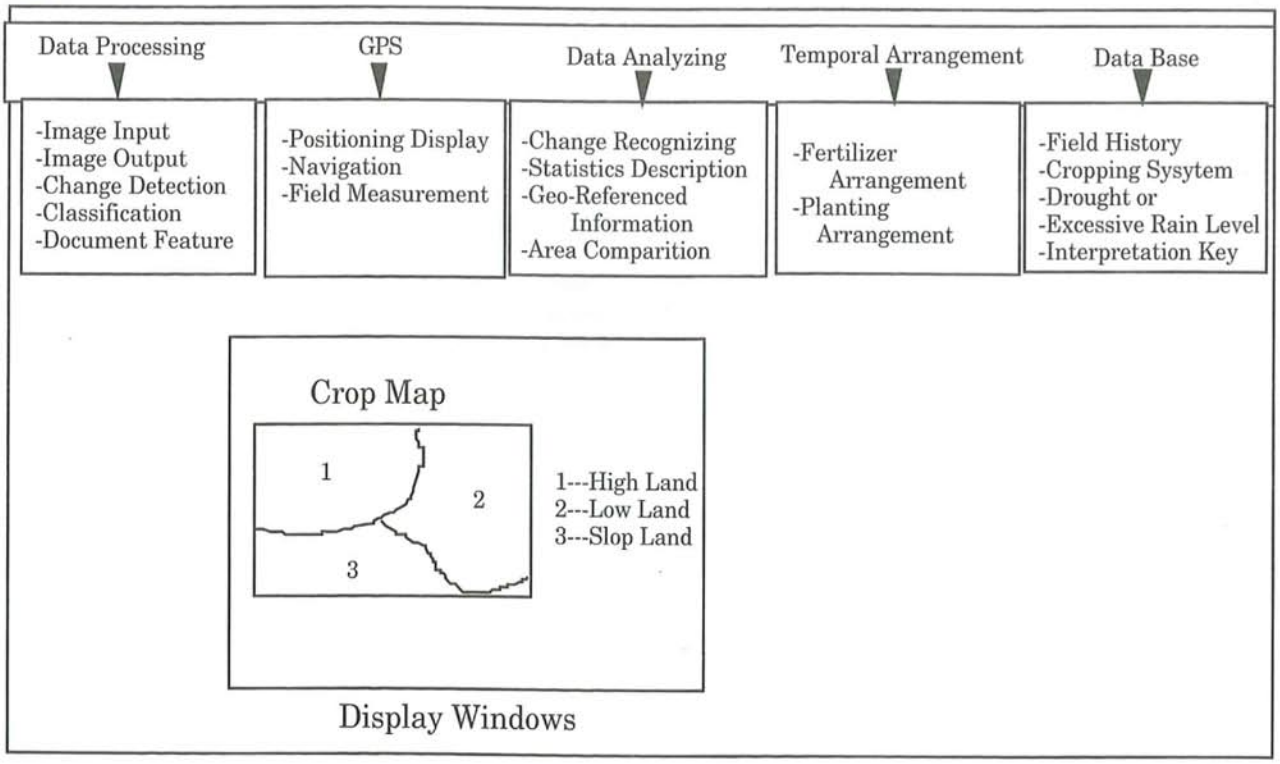


Figure 1. Basic Interface Screen Overview and Display Windows

the system.

2). To feed position information into the system.

3). To feed navigation information into the system.

3. Data analyzing. The analysis subsystem will be in charge of following several items:

1). To determine possible correlation between the features and interpretation of significance to farmcrops temporal arrangement

2). To describe the situation for a service area around each of pointed locations based on statistics

3). To compare the areal extent of the farmcrops defined in each set of satellite data

4. Temporal arrangement. The study which is proposed has two important appliances:

1). To find the appropriate procedure for monitoring and detecting farmcrops change

2). Planning techniques are used to arrange the rotation of crops and the adequate time of fertilization.

5. Database. Field history, cropping system, level of drought or excessive rain, interpretation key and so on are maintained, managed and created for the farmcrops temporal arrangement.

III. PLOTTING A POSITION ONTO A SEGMENTATION IMAGE

Image segmentation is an important step in the analysis of digital images. Usually, it is employed after image enhancement and before object recognition. In this section, the problem of how to transform a primary image into a segmentation image, and homogeneous region image is described. At the same time, we will apply an adaptive Voronoi tessellation to the image.

At first, the primary image is segmented by the judgment of intraparallelism which involves the attributes stemming from the gray level, such as intensity, hue, saturation etc. To each pixel of the segment that maybe contains multiple objects, if the attributes are parallel, then give it a corresponding label, otherwise give it a question mark label which is a control label used to indicate where the segments must be continuously subdivided. As a result, an image composed of these labels is obtained. It is called the segmentation image.

The pixel intensities of neighboring pixels from the same object can be assumed to be incompletely correlated. For many objects more than one segment may be obtained. Thus these segments must be conbined into homogeneous regions on the basis of the knowledge of objects that determine collections of segments, which form "natural" components of the scenes.

Figure 2 illustrates the elements of segmentation image and homogeneous region image, and their transformation. The attribute and uniformity of segments, which are mainly based on statistical properties and sensor dependent; while the attribute and homogeneity of homogeneous regions mainly refer to the knowledge of objects and themes, which in-

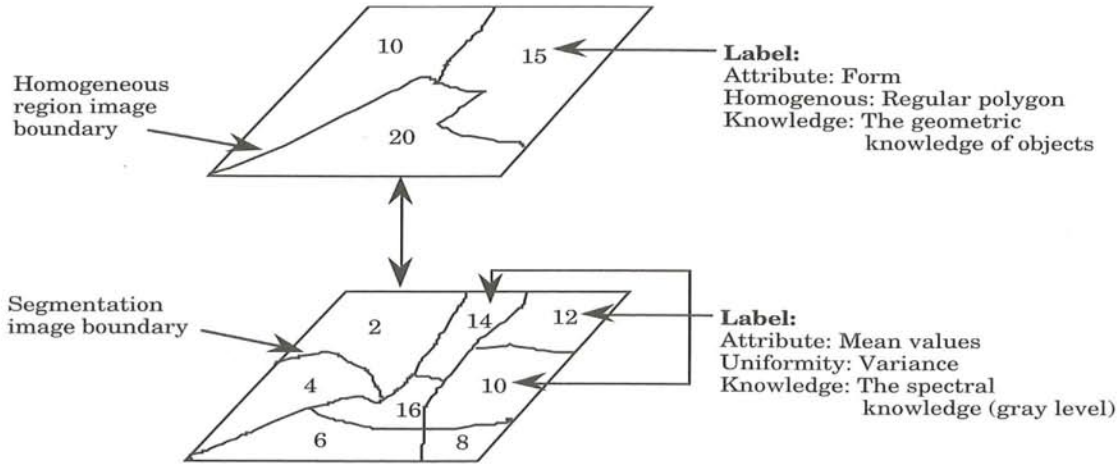


Figure 2. The elements of segmentation image and homogeneous region, and their transformation.

clude the geometric characteristics. For example, segments 10, 12, 14 in Fig.2 indicate three rice lands with different water depths. Because all of them have the same attribute, regular polygon, they should be merged into a single homogeneous region by using form attribute.

Pyramidal techniques have been shown to be efficient methods for image segmentation. On this already irregular structure, an irregular pyramid will be build. The advantage of this approach is that the Voronoi tessellation and the Delaunay graph offer a reduced description of the image and the neighbour relation between regions, which is already adapted to the image content(Etienne 1996).

Voronoi diagrams built from a distribution of discrete seeds are also of interest in the content of geometrical partitions. Given a distance function, each Voronoi polygon is associated with a seed and is defined as the set of points which are closer to this seed than from any of the other seeds (see Figure 3).

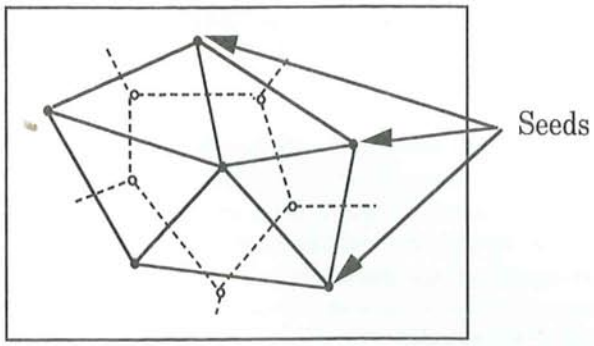


Figure 3. A Voronoi diagram and the associated Delaunay tessellation. The seeds are contained in the interiors of the polygons.

Each seed is connected with field sample locations using GPS/DGPS so that a consistent reference system could be used. The number of sample points is approximately that of the seeds.

Geo-referenced information means information available at a respective seed position. Thus, if positioning is available at the seed, image segments may be accessed. A number of intermediate stages are conceivable, starting with a display for the user only and leading to totally automated processes.

Image segmentation proceeds in two steps. The first step is the splitting: polygons are added in the support of the seeds and image until convergence. This step involves a dynamic management of the Voronoi diagrams. The second step is the merge step: some polygons and seeds are deleted. If a polygon P is nonhomogeneous, a seed should be added in the middle of each Voronoi edge $e = P \cap Q$ where the polygon Q is neigbbour of P, otherwise, we do nothing.

IV. PLANNING TECHNIQUES FOR FARMCROPS TEMPORAL ARRANGEMENT

Economic representation of data with all their interrelationships is one of the most central problem in information sciences. In thinking, and in the subconscious information processing, there is a general tendency to compress information by forming reduced representations of the most relevant facts, without loss of knowledge about their interrelationships. The purpose of intelligent information processing seems in general to be creation of simplified images of the observable world at various levels of abstraction, in relation to a particular subset of received data.

Temporal arrangement of crops is a vivid example of this kind of problem hence the name "temporal arrangement problem". The term "temporal" strongly evokes the idea of a one-dimensional planning space but the temporal arrangement problem may also refer to temporal-spatial planning in a four-dimensional workspace. However the section of this paper will focus on the one-dimensional case.

Consider the map shown in Figure 4. It consists of three landforms: high-land, slope-land and low-land, where one is a neighbour of the others which are planted or are not.

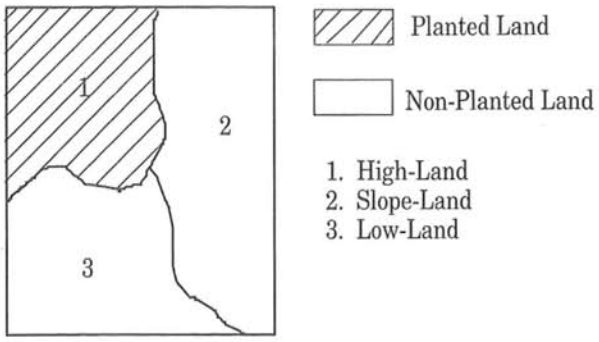


Figure 4. A map of cropping system

The study area is chosen in lower reaches alluvial plain of Yellow river. This study area selection is made on the basis of the geographic coverage of the socioeconomic data that corresponds to the mainly farming area in China. In addition, area bounding the Yellow river and suffering from severe drought or excessive rain is the major factors considered in choosing the study area. The relationships between landform and drought or excessive rain are demonstrated in Table 1. The farmcrops in this area may be classified five types:

(1) Spring-sown crops April-August.
(2) Spring-Summer-sown crops May-Septem ber.
(3) Summer-sown crops June-October.
(4) Summer-Autumn crops July-November.
(5) Autumn-sown crops October- June.

To make a farmcrops temporal arrangement, we give the constraints as follows:

(a) Balance spring-sown crops with autumn-sown crops.
(b) The farmcrops temporal arrangement of high-land is prior to slope-land, in turn slope-land is prior to low-land.
(c) Adjust farmcrops temporal arrangement cor responding to the level of drought or excessive rain.

The planning techniques used for farmcrops temporal arrangement by this system have two cases:

(1) When the middle-long-term weather forecasting is not considered, the farmcrops temporal arrangement is a combination of Plot 1 + Plot 2 + Plot 3 temporal arrangement. In terms of constraints (a), (b),(c) above, we may get the arrangements shown in Figure 5.
(2) According to middle-long-term weather forecasting, the farmcrops temporal arrangement appears to be a more complicated situation to represent the relationships between the plots. An algorithm for farmcrops temporal arrangement considering middle-long-term weather forecasting is presented. According to constraints (a), (b), (c) above, we may get the arrangement shown in Figure 6.

V. CONCLUSIONS

The use of PSS is expected to increase in the future, but the development of both the hardware and the software will have to be based on rational considerations and projects. To enhance the interactivity, efficiency, and computational power in a WINDOWS environment as much as possible, suitable system may be developed as, for instance, the framework presented in Section 2. A novel data processing and data analysis algorithm that attempts to build Voronoi diagrams from a distribution of discrete seeds, which are connected with field sample location using GPS/DGPS, is proposed. The incorporation of image segmentation and location sensing helped to combine the advantages of small field sample locations with large-scale, cost-efficient image processing methods is pointed out. At the same time, a farmcrops temporal arrangement method in terms of the integration of GPS/GIS/RS and planning techniques is proposed. It will be useful to reduce errors and time in precision farming.

REFERENCES

[1] Chen, C. F. et al, 1992, Integrating remote sensing, GIS and GPS in PEN computing environment,

Table 1. The relationships between landform, level of drought or excessive rain and rotation of crops

Landform	Level of drought or excessive rain	Rotation of crops	Field yield
High-land	Other	Twice / a year	High
	Very serious	Thrice / two years	Middle
Slope-land	Less serious	Twice / a year	High
	Serious	Thrice / two years	Middle
	Very serious	Once / a year	Low
Low-land	Less serious	Thrice / two years	Middle
	Serious		
	Very serious	Once / a year	Low

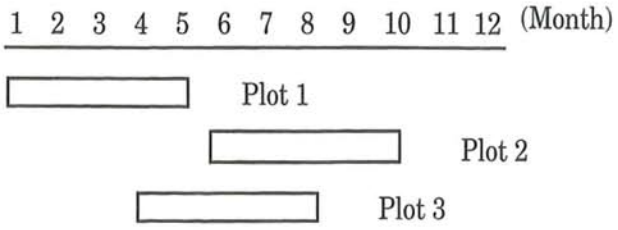


Figure 5. Farmcrops temporal arrangement ignoring weather forecasting

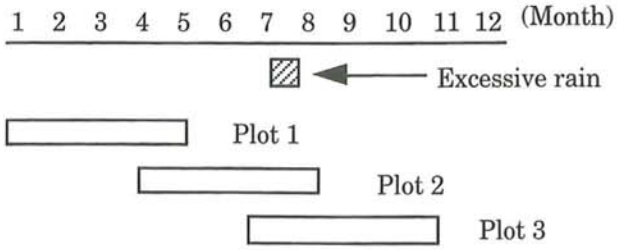


Figure 6. Farmcrops temporal arrangement considering weather forecasting

ASPRS/ACSM/RT 92, pp. 309-314.

[2] Etienne, B. et al, 1996, Vornoi pyramids controlled by Hopfield neural networks, *Computer Vision and Image Understanding*, Vol. 63, No. 3, pp. 462-475.

[3] Hermann, A. et al, 1995, GPS and DGPS as a challenge for environmentally-friendly agriculture, *The Journal of Navigation*, Vol. 48, No. 2, pp. 269-278.

[4] Lachapelle, G. et al, 1994, GPS system integration and field approaches in precision farming, *Navigation, Journal of the Institute of Navigation*, Vol. 41, No. 3, pp. 323-335.

CPGIS 于 1997 年 6 月 2 日在北京举行了创会五周年庆祝活动。在京的 60 多位 GIS 专家、领导参加了这一活动。会后，由 CPGIS 会员举办了七场专题讲座。图为 CPGIS 五任会长，从左到右：宫鹏、李斌、丁跃民、柳林、林珲。又见本刊 19 页和 50 页。

On June 2, 1997, CPGIS held its 5th Anniversary Celebration Meeting in Beijing. All the Presidents of CPGIS attended the meeting. From left to right, they are Peng Gong, Bin Li, Yuemin Ding, Lin Liu, and Hui Lin.

Linear Feature Modeling with Curve Fitting: Parametric Polynomial Techniques

Xiaoming Zheng and Peng Gong

Center for Assessment and Monitoring of Forest and Environmental Resources
Department of Environmental Science, Policy and Management,
University of California, Berkeley, CA 94720-3310, USA

Abstract
A decomposition model is described to model linear features sampled by manual digitization or field survey. The model consists of three components, original data, systematic pattern, and random error. Least squares and moving least squares techniques are introduced for polynomial curve fitting. Polynomial functions are proposed to represent linear features. The position deviation between sampled points and the polynomial function is used as an approximation of the random error. Experimental results are presented to show the effectiveness of the decomposition model. Potential applications of the model have been discussed including estimation of errors associated with points sampled along linear features, digital representation and mapping of linear features.

摘 要
本文介绍对手工数字化或野外测量的线性特征进行模拟的方法。本模拟有三个组成部分: 模拟出的原始数据、系统模式和随机误差。我们的模型使用多项式参数曲线拟合技术配以最小二乘法和滑动最小二乘法技术进行参数估算。所得的多项式参数函数可用来模拟曲线。由多项式函数和数字化点之间的偏差用来近似估算随机误差。试验结果表明我们的模型有一定功效。该模型在数据误差估算、线性特征的模式、数字表达和线性特征制图等有应用潜力。

I. INTRODUCTION

In a vector-based GIS, digital representation of curve features is done through the use of a series of point coordinates sampled along the curves. Manually digitizing paper maps is a predominant method of point sampling. This has been recognized as a significant source of error of spatial data (Chrisman, 1982). Perhaps a more precise method is to use global positioning systems (GPS) units in the field. The high precision of GPS receivers, however, does not offer much help for curve features because such features are approximated by and handled as a series of straight line segments joining the consecutively sampled points. The facts that use of discrete points to represent curve features is prone to errors and that such errors are not quantified are fundamental problems, some yet unsolved tasks, in vector-based GIS (Brunsdon and Openshaw, 1993).

To exactly calculate position errors of discrete points, we need to compare the sampled points with their true position along a curve. In most cases, however, the true curve is not known. We must first approximate the true curve from sampled points and then derive position uncertainties by comparing the sampled points with the estimated curve. To do so, we must (1) develop a mathematical model that can approximate and represent the true curve features in a map or in reality based on the sampled points, and (2) provide a procedure to estimate the errors or uncertainties associated with the curve model. Splines function used for curve fitting and interpolation is not suitable for those purposes because it forces the fitted curve to go through the sampled points. In time series analysis, some prediction models such as the autoregressive model, the Autoregressive Integrated Moving Average (ARIMA) model and adaptive filtering based on Wiener-Levinson and Kalman Filter theories may be useful, but they use data observed in the past to predict the future behavior of the modeled phenomenon (Janacek and Swift, 1993, Graupe, 1984). In spatial data modeling, it is desirable to use both the "past" and "future" points. Some shape

1082-4006/97/0301~2-7$3.00

analysis methods (e.g., Lin and Hwang, 1987; Gunther and Wong, 1990; Grogan et al., 1992) including strip tree, curve fitting method using Bezier curves, arc tree, and Fourier descriptors may only be useful for curve representation not for estimation of errors or uncertainties.

Polynomials can be used for the two purposes mentioned above, particularly for intuitively smooth curves that are continuously differentiable. In reality, not all linear features have this mathematical property. Many linear features resulting from human activities may not be continuously differentiable. Examples are roads, utility lines, cadastral and administration lines. On the other hand, most linear features delineating natural phenomena can be considered as continuously differentiable. These include contours, streams, and natural resource boundaries (e.g., soil, climate, vegetation, wetland, etc.). Because of the increased amount of human abstraction realized by map generalization, on smaller scale maps we observe a larger proportion of differentiable curves. It is possible to store polynomial coefficients and use polynomial functions to represent curve features particularly if lower order polynomial functions can fit curves with sufficiently high accuracy. It may require less space to store polynomial coefficients than to store sampled point coordinates. In addition, it is effective to use polynomial coefficients to represent curve shapes. Curve shape analysis may be made based on polynomial functions for subsequent curve generalization, curve matching for object registration or recognition. Thus, curve representation with polynomial coefficients may have some advantages in data storage and curve shape analysis over the traditional curve representation method involving consecutive straight lines.

The objective of this paper is to develop a decomposition model for curve fitting by employing polynomial functions. The model consists of three components, original data, systematic error, and random error. It is used to simulate differentiable curve features from sampled points and to approximate sampling errors or uncertainties. Without loss of generality, we concentrate on the development of the model and its application to digitized curve features. Sampled points through GPS units can be processed in the same manner. In the next section, we introduce a framework for spatial data modeling particularly curve modeling based on the decomposition model. In section 3, we introduce an epsilon band model for the estimation of errors or uncertainties of sampled points that constitute a stationary random data series. For estimation of polynomial coefficients, we describe least squares and moving least squares methods in sections 4 and 5, respectively. The two methods are used to implement the curve models. Some experiment results with digitized data from simulated curves are presented in section 6 followed by some conclusions.

II. SPATIAL DATA MODELING

There are two kinds of natural or social phenomenon that can be described with a mathematical model. One is deterministic physical process or signal, which is entirely known and can be represented exactly with a mathematical function. The other is random event, which can only be described using a stochastic model based on random samples.

Spatial data digitization is a stochastic process (Keefer et al, 1988). The random error is introduced during the generation, analysis and processing of the digitized spatial data. To model a digitized line, its uncertainty and random error, three steps are needed. These are model selection, model estimation, and model evaluation.

Model Selection

A spatial series model describing a curve or a linear feature should be capable of (1) representing the original data with a deterministic mathematical function that can simulate or account for the sampled data series, and (2) estimating the random error distribution for evaluating the accuracy of the fitting, interpolation and prediction of the linear feature. Selecting an appropriate model is one of the most crucial steps in spatial data modeling. Model selection criteria depend on the objective of data simulation. To describe the behavior of a physical phenomenon, we may derive a model based on physical laws so that we can precisely represent or predict the value of a physical parameter in a given time or space. To estimate the trend of a random phenomenon affected by many unknown factors, stochastic process models can be chosen. It is helpful to plot the data first for understanding the type, pattern and trend of the random data sets. The knowledge on the physical process of data acquisition is also useful in mathematical model selection.

For a discrete series represented by digitized points $\{P_t, t = 1, 2, \ldots, n\}$, the series can be expressed with a general stochastic decomposition model

$$P_t = F_t + E_t + R_t \quad (1)$$

where

- P_t is the digitized point;
- t is a number index for a particular point;
- F_t is the component that represents the original, undistorted part of the data series;
- E_t contains the systematic pattern or systematic error that can be removed if known.
- R_t is the random component, which can only be estimated using some a priori knowledge about its distribution.

A digitized-point series consists of a sequence of X and Y coordinates, which can generally be represented as:

$$\{P_t = (X_t, Y_t), \quad t = 1, 2, \cdots, n\} \quad (2)$$

Because sample points are a discrete series with unequal intervals, it is more practical to write the point set in a parametric form

$$\begin{cases} P_t = (X_t, Y_t) \\ X_t = X(s_t) \\ Y_t = Y(s_t) \\ t = 1, 2, \cdots, n \end{cases} \quad (3)$$

where s_t is a distance parameter between the origin and point t. Both X_t and Y_t are the functions of parameter s_t. X_t and Y_t can be fitted by using the same mathematical form with different coefficients. We only focus on the discussion of data modeling with the series of $\{X_t, \quad t = 1, 2, \cdots, n\}$ in this paper. The decomposition model can be written as

$$X_t = f_t + e_t + r_t \quad (4)$$

The component of e_t largely depends on the physical process of data acquisition and the digitizer used. It is difficult to use a mathematical expression to describe the systematic pattern without a complete knowledge of a specific data series to be modeled. One of the common systematic errors is linear shift. To remove the systematic effect of a linear shift, a linear parametric function can be used to rectify the error. Some systematic patterns or errors may be modeled separately through visual analysis of the digitized data. Visualizing the sample points may allow systematic errors to be detected and corrected through manual editing. To simplify the discussion, we assume that there is no systematic pattern and error in a digitized data series, that is, the component of e_t is zero. We have

$$X_t = f_t + r_t \quad (5)$$

Generally, any data series can be decomposed into a deterministic and a random part and can be represented by equation (5) (Janacek and Swift, 1993). Model selection includes the determination of the mathematical expression for the parametric equation f_t and the definition of the distribution of the random component r_t.

Model Estimation

A mathematical model for describing a random event contains some unknown parameters, which should be estimated with the available sample data. Least squares is an important statistical technique for estimating model parameters based on some specified standard and criteria.

Suppose a model for a spatial data series takes the form

$$X_t = f_t + r_t = f(t, \theta) + r_t \quad (6)$$

where θ is the parameter vector and $\theta = (\theta_1, \theta_2, ..., \theta_k)^T$.

Let $\hat{X}_t = f(t, \theta)$ be an estimation of the original data set X_t. Then the deviation of the estimation for the t-th data point is

$$r_t = X_t - \hat{X}_t, \qquad t = 1, 2, ..., n \quad (7)$$

The sum of squared deviations is

$$R = \sum_{t=1}^{n} r_t^2 = \sum_{t=1}^{n} (X_t - \hat{X}_t)^2 \quad (8)$$

The criterion for calculating parameter θ_i is that the parameter can minimize the sum of the squares. Let

$$\frac{\partial R}{\partial \theta_i} = 0, \qquad i = 1, 2, ..., k \quad (9)$$

Since $\hat{X}_t$ is a linear function of the parameter vector θ, we can draw a set of k linear equations from (9). Solving the linear equation set, we can obtain the parameters, θ_i (i=1, 2, ..., k; k≤n), which minimize the sum of squared deviations in (8).

Model Evaluation

When a model is estimated based on the available data sets, it is necessary to diagnostically check the goodness of fit between the estimated and the digitized data. We need to ascertain if the model is appropriate for the data set, and evaluate the estimated characteristic parameters.

III. STATIONARY RANDOM DISTRIBUTION AND EPSILON MODEL

Suppose there is a random data series $\{X_t | t = 1, 2, ...\}$. If its statistical characteristics do not change with variable t, that is, the characteristics is independent of the origin of variable t, we call the random data series stationary (Janacek and Swift, 1993).

A stationary data series have a constant mean

$$\mu_t = E[X_t] = \mu \tag{10}$$

and, for any two points t and s in time series, its autocovariance function satisfies

$$R(s-t) = E[(X_s - \mu)(X_t - \mu)] \tag{11}$$

Let m=s-t , and we have

$$R(m) = E[(X_{t+m} - \mu)(X_t - \mu)] \tag{12}$$

Specially, if m=0, the autocovariance becomes the squared deviation

$$R(0) = E[(X_t - \mu)^2] \tag{13}$$

A stationary random series is completely characterized by its mean and autocovariance. The exact values of these parameters can be calculated if the ensemble of all possible realizations is known. Otherwise, they can be estimated if multiple independent realizations are available. However, in most applications, it is difficult or impossible to obtain multiple realizations. Most available spatial data series constitute only a single realization. This makes it impossible to calculate the ensemble average. For a stationary data series, we have a natural alternative of replacing the ensemble average by the average along the time or distance axle if the stationary process is ergodic (Zhong and Hu, 1990).

The process of digitization can be considered as a stationary random sampling process when the cursor is used to trace a curve which can be modeled by a normal distribution (Keefer et al, 1988; Maffini et al 1989; Gong and Chen, 1992). This implies that the probability of the sampled points located at both sides of the curve are about the same and the sum of all the errors cancels out. However, for a random sample point, it is impossible to predict its position along the curve. Moreover, the true curve is usually unknown. It has been suggested that the epsilon band model proposed by Perkal (1966) be used to represent an uncertain zone centered at the continuous representation of the digitized curve (e.g., Blakemore, 1984). The position deviation, epsilon, times a certain number is used as the width of the uncertain zone (Figure 1).

In practice, the problem with applying the epsilon band model to indicate curve uncertainty is how to estimate the position deviation - the epsilon value. Distances between the representation of the curve and the digitized points (Bolstad et al, 1990; Gong and Chen, 1992) may be used to estimate the width of the epsilon band. Provided that the true curve can be simulated with a mathematical function, the digitizing error can be estimated from the sample

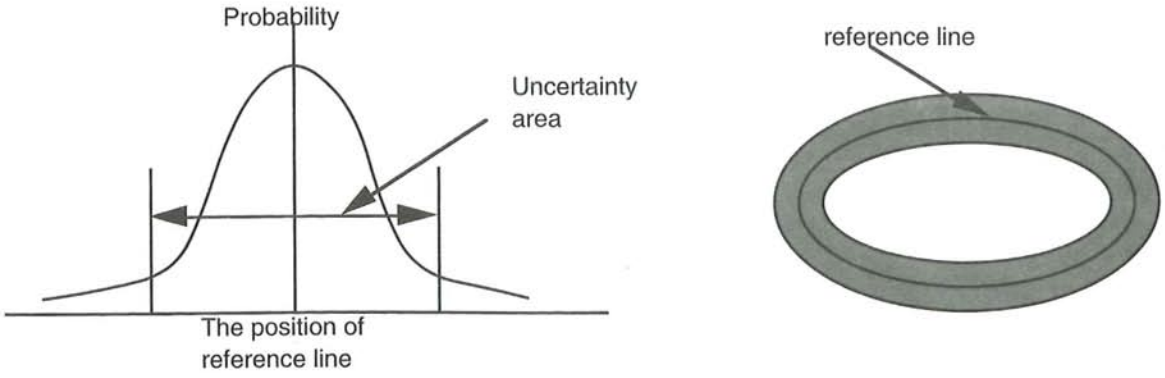


Figure 1. The distribution of digitizing points along a true boundary line and Epsilon band.

standard deviation of the stationary random process based on a series of digitized points.

$$\hat{r} = \sqrt{\hat{R}(0)} = \left[\frac{1}{n}\sum_{t=1}^{n}(X_t - f_t)^2\right]^{\frac{1}{2}} \qquad (14)$$

where f_t is supposed to be the points on the true curve, which is consistent with equation (5). Because f_t is unknown, a polynomial function can be used to represent f_t based on the available sample points.

IV. UNCERTAINTY MODELING WITH POLYNOMIAL CURVE FITTING

From equation (5), we have

$X_t = f_t + r_t$

where f_t represent the undistorted curve, and r_t is the random component with a normal stationary distribution.

There are two approaches to estimating the random error r_t, filtering and curve fitting.

The first approach is to use a filter to remove f_t from X_t, that is

$$\hat{r}_t = F\{X_t\} \qquad (15)$$

If the original curve is continuous and smooth, f_t should be a low frequency signal, which can be fitted by using a polynomial function. The component of r_t is mainly a high frequency signal, which usually has a normal distribution. To remove f_t, filtering can be applied in spatial or frequency domain.

In spatial domain, a high-pass filter is equivalent to taking the derivatives. If f_t can be represented by a K-th order polynomial function, a (K+1)th-order differential operator can remove f_t from X_t. The standard deviation of the differentiated result can be used as an estimation of the random error. The problem is that, for a discrete spatial data series, a difference operator has to be used to replace the differential operation. Difference operation is not invertible, and the operation will enhance the random component when removing the low frequency signal, which will change the magnitude of the error and affect the estimation of the parameters. High-pass filtering in frequency domain seems to be more reasonable for removing f_t and estimating the random error if a filter can be designed to remove the low-frequency part and keep the high-frequency part unchanged.

The second approach is to use a mathematical function to simulate f_t, and then subtract the estimated $\hat{f}_t$ from X_t, that is

$$\hat{r}_t = X_t - \hat{f}_t \qquad (16)$$

We use polynomials to fit a series of digitized points as a simulation of the true curve. The general form of a polynomial function is defined as

$$X = X(s) = \sum_{k=0}^{K} a_k s^k = a_0 + a_1 s + a_2 s^2 + \ldots\ldots + a_k s^k \qquad (17)$$

where K is a non-negative integer, the degree of the function, and $a_0, a_1, \cdots, a_k$ are fixed real numbers, s is the distance between point s and the origin which can be defined as the first digitizing point. The coefficients $a_0, a_1, \cdots, a_k$ can be calculated using the least squares technique based on the available sampled points.

Suppose that there are n digitized points for a curve. For the t-th point, the fitting equation is

$$\hat{X}_t = X(s_t) = \sum_{k=0}^{K} a_k s_t^k = a_0 + a_1 s_t + a_2 s_t^2 + \ldots\ldots + a_k s_t^k, \quad t=1,2,\ldots,n \qquad (18)$$

The residue is

$$R = \sum_{t=1}^{n} r_t^2 = \sum_{t=1}^{n} (X_t - \hat{X}_t)^2 \qquad (19)$$

Let

$$\frac{\partial R}{\partial a_i} = 0, \qquad i = 0,1,\ldots,k\,(k \le n) \qquad (20)$$

we have a set of k linear equations for $a_0, a_1, \cdots, a_k$. Solving these equations we can obtain the unique solution for all the coefficients. In practice, the shape of a curve to be fitted needs to be smooth and continuous and the order of the polynomials cannot be infinitely high. These are further elaborated below.

(1). Curve continuity

A curve is mathematically continuous and smooth if its various order of derivatives exist. If a curve is discontinuous or there exist sharp turning points, it should be divided into continuous and smooth segments at the broken points, and piecewise polynomial functions may be constructed segment by segment. Practically, polynomial functions are less effective for linear features that are intrinsically non-smooth or mathematically discontinuous

because more sample points may be required and a curve may have to be broken into too many segments.

(2). The order of the polynomial function

Theoretically, a polynomial function can fit any continuous and smooth curve so long as the order is sufficiently high. Usually, better results can be obtained when a higher order of polynomial function is used to fit a curve if there is a sufficient number of sample points. Practically, there is a computational problem related to the limited precision and magnitude of a computer. An exceedingly high order will cause the fitted curve vibrating around the curve because of the intrinsic ill-condition of the Vandermonde problem and the roundoff errors, which may introduce rather substantial coefficients in the leading terms of the polynomial. A reasonable order for fitting a specific curve needs to be determined.

If a k-th order polynomial function can completely represent a curve, then (k+1)th derivative operation will result in zero. This fact can be used to construct a method to determine the order of a polynomial function.

For discrete sample points, the backward difference operator can be defined as (Wei, 1990):

$$\nabla X_t = X_t - X_{t-1} = (1-B)X_t \qquad (21)$$

and

$$\nabla^k X_t = (1-B)^k X_t \qquad (22)$$

where $\nabla = 1 - B$ and B is a backward shift operator $BX_t = X_{t-1}$.

After the k-th order difference operation, we need to test the assumption of random stationary distribution for the residue (Janacek and Swift, 1993). If the assumption is true, we take k as the appropriate order for the polynomial function. If a curve can be completely represented with a K-th order polynomial function, for any P-th order polynomial function (P>K), all the coefficients $a_j \approx 0$ $(K < j \le P)$, and the accuracy of the fitting should be the same as the K-th order polynomial function. There is a more practical method to determine the order of a polynomial curve. When k is smaller than K, the fitting error or residual R_k monotonically decreases as the order increases. When k increases to K+1 and if R_k equals R_{K+1} or even is less than R_{K+1} because of the intrinsic ill-condition of the Vandermonde problem and the roundoff errors, K should be taken as the order of the polynomials.

V. MOVING LEAST SQUARES FOR CURVE FITTING

If a continuous curve changes sharply in some parts and changes gently in some other parts, it requires a high order polynomial function to fit the curve. On the other hand, because the precision limitation of a computer, a sufficiently high order of polynomial function will cause vibration of the fitted curve and hence introduce a large fitting error. To solve this problem, moving least squares can be used.

The basic idea of moving least squares is that if X is the function associated with the fitted curve, then the value of X at a point s should be most strongly influenced by the values at those points s_t that are close to s. In other words, the influence of a value at s_t on X at point s should decrease as the distance between s and s_t increases. Therefore, we can modify equation (19) to a weighted sum of squared deviations and minimize

$$R_m = \sum_{t=1}^{n} w_t(s)[X(s_t) - X_t]^2$$

$$= \sum_{t=1}^{n} w_t(s)\left[\sum_{k=0}^{K} a_k s_t^k - X_t\right]^2 \qquad (23)$$

where $w_t(s)$ is the weight function of s at point t. Choose the following function

$$w_t(s) = \exp\left(-C(s-s_t)^2\right) \qquad (24)$$

which is a monotonically decreasing function, and C is a constant for adjusting weights of neighboring points. The greater C is, the smaller is the size of the neighborhood points that have significant effect on the fitted value at position s.

To obtain the optimal coefficients according to the minimum squares of deviation, let

$$\frac{\partial R_m}{\partial a_k} = 0, \qquad k = 0, 1, \ldots, K \qquad (25)$$

Then, the normal equations are

$$a_0\left(\sum_{t=1}^{n} w_t s_t^0\right)+a_1\left(\sum_{t=1}^{n} w_t s_t^1\right)+\cdots+a_K\left(\sum_{t=1}^{n} w_t s_t^K\right)=\sum_{t=1}^{n} w_t X_t$$

$$a_0\left(\sum_{t=1}^{n} w_t s_t^1\right)+a_1\left(\sum_{t=1}^{n} w_t s_t^2\right)+\cdots+a_K\left(\sum_{t=1}^{n} w_t s_t^{K+1}\right)=\sum_{t=1}^{n} w_t s_t^1 X_t$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots$$

$$a_0\left(\sum_{t=1}^{n} w_t s_t^K\right)+a_1\left(\sum_{t=1}^{n} w_t s_t^{K+1}\right)+\cdots+a_K\left(\sum_{t=1}^{n} w_t s_t^{2K}\right)=\sum_{t=1}^{n} w_t s_t^K X_t \qquad (26)$$

For polynomial curve fitting based on least squares, the coefficients of the polynomial function is identical for any fitted point s. Therefore, a high order polynomial function may be required. With moving least squares, the solution for the coefficients $a_0, a_1, \cdots, a_k$ depends on s through the weight function $w_t(s)$. For each s, we have to solve a set of normal equations. Therefore, it is computationally prohibitive to use high order polynomials.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

Experiments were conducted to evaluate the accuracy and uncertainty of digitization and curve fitting for linear features using polynomial functions based on least squares and moving least squares. The original curves were generated with some mathematical functions. The exact position error of each point can be calculated by comparing the digitized point, the fitted curve value with the original mathematical function.

Experimental procedures are as following:

Step 1. Design a kind of mathematical function to generate a curve.

Step 2. Print out the curve and digitize the curve using a digitizer.

Step 3. Fit the curve using a polynomial function based on least squares and moving least squares

Step 4. Estimate the random error and evaluate the accuracy of curve fitting.

Curve Generation

Different mathematical functions were used to generate different shapes of curves. To evaluate the effect of curvature on digitization, circles with different radii were used. It is more difficult to fit a curve that has different curvatures in its different segments. A set of sine functions were used to represent the curves with different curvatures. The sine function used to generate curves was $y = a\sin bx$, where a and b were parameters for adjusting the shape of the curve a={1.0, 1.25, 1.5, 2.0} and b={1.0} in our experiments. The third kind of curve was an ellipse representing polygon boundaries with different curvatures. The ellipse equation is:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

where a={2, 6} and b={1, 3}.

To further test the polynomial curve fitting technique, more complicated curves were constructed by mirroring a sine curve (Figure 6) or joining two ellipses (Figure 7).

Digitization

All the curves generated with mathematical functions were digitized manually by some experienced operators at their normal speeds. The digitized data were then taken as sampled points for curve fitting and random error estimation.

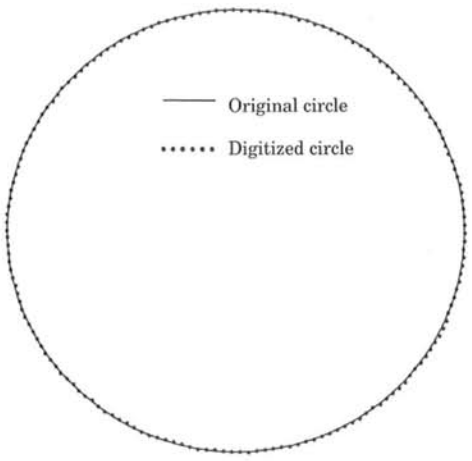


Figure 2. A digitized version of a circle (dotted) overlaid on top of the original one (solid).

Figure 2 shows a set of digitized points along one of the original circles. The radius of the circle is 5.5 cm printed with a laser printer having a resolution of 600 dots per inch (DPI). The line thickness is 0.1 mm. The digitization was done on a Summagraphics (MM II 1812) digitizer with a resolution of 1000 DPI. As can be seen from Figure 2, the digitized points are not exactly on the circle. We calculated the exact position errors and plotted the distribution of the digitizing errors in Figure 3. As expected, most of the digitized points locate along the circle and the number of sampled points decreases as the distance between the digitized points from the circle increases. Although the distribution of errors is a little skewed, it looks close to a normal distribution. In this study we assume that the digitizing error distribution is normal.

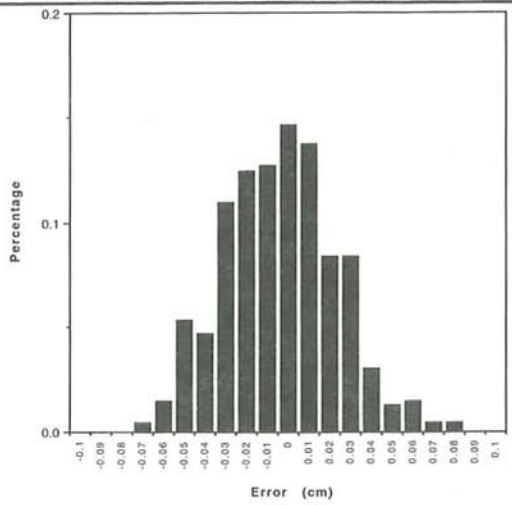


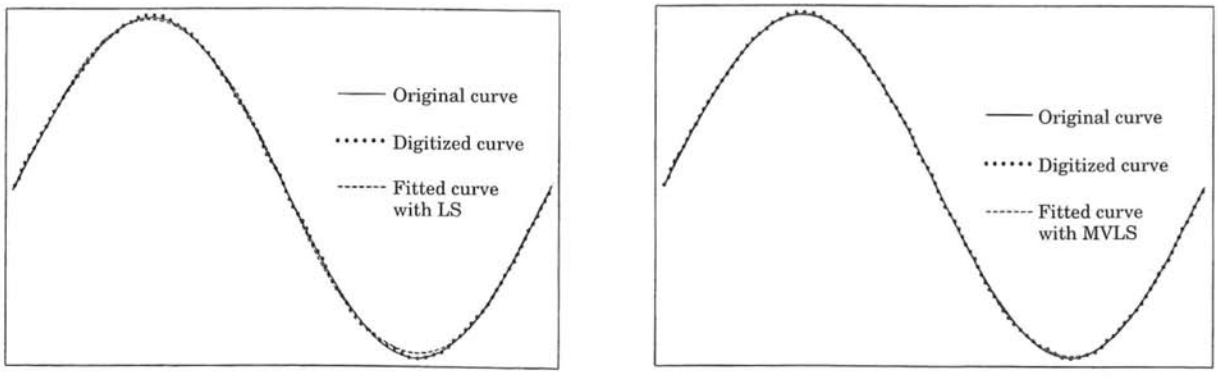
Figure 3. Digitized error distribution calculated from the example in Figure 2. Error is determined by calculating the distance of each digitized circle

Errors caused from curve plotting by a printer and point-position reading from digitizing tables are determined by the resolution of the printer and digitizer used. Since a 600 DPI printer was used, curve plotting errors should be within approximately +/- 0.022 mm while point reading errors should be within 0.013 mm. Because errors from different sources do not simply add up (Gong et al., 1995) and the error magnitudes of curve printing and point reading are one order of magnitude less than the digitizing errors at an average level of approximately 0.2 mm (Figure 3), errors caused by curve printing and point reading have been ignored in this study.

Curve fitting

The curvature of the sine curve reaches its maximum at the peak and valley positions (Figure 4). With a 9-th order polynomial function, the curve fitting errors are still largely observable. Particularly, the fitted curve does not reach the apices of the sine curve (Figure 4a). Better results were achieved from the moving least squares with an order of 5 (Figure 4b).

Figure 5 shows a comparison of the results from the least squares and the moving least squares. The four curves in Figure 5(a) include the original curve and the curves generated by polynomial functions of order 1, 5, and 9, respectively. It can be seen that the curve simulated with the 9th order polynomials is a close approximation to the original curve. Figure



(a) (b)

Figure 4. The effect of curvature on curve fitting. (a) Original curve, digitized version, and fitted curve using the 9th order polynomial functions estimated with the least squares method. (b) Fitted curve using the 5th order polynomials estimated with moving least squares.

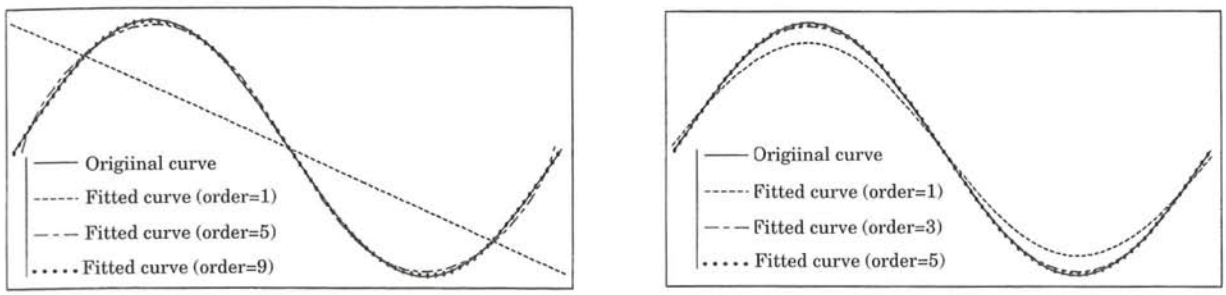


Figure 5. Curve fitting results using different orders of polynomial functions. The original is displayed using a solid line. (a) Fitted results using least squares with polynomial orders of 1, 5, and 9, respectively. (b) Fitted results using moving least squares with polynomial orders of 1, 3, and 5, respectively.

Table 1. The Comparison of the Errors of Least Squares and Moving Least Squares

Least Squares			Moving Least Squares		
Order	x (cm)	y (cm)	Order	x (cm)	y (cm)
1	0.463824	0.477288	1	0.029826	0.102721
3	0.059176	0.075351	2	0.011229	0.018969
5	0.023897	0.028138	3	0.008875	0.014013
7	0.006212	0.020102	4	0.004368	0.006309
9	0.006224	0.006068	5	0.003307	0.004847

5(b) shows the curve fitting results obtained from the moving least squares with order 1, 3, and 5, respectively. The curves generated with the 5th order polynomial functions fit well to the original curve. Table 1 summarizes some of the curve fitting accuracies. For the least squares method, when the order is 11, the sample variances along both the x and y directions are tremendously greater than those obtained from the 9th order polynomial functions. Therefore, an 11th order polynomial function may represent an over fitting to the original curve because of the intrinsic ill-condition of the Vandermonde problem and the roundoff errors.

Figure 6 shows an example when curve fitting by one single polynomial function reaches its limit in simulating curve sections containing sharp curvature changes. In this example, there are two sharp points. For the curve segment containing the left sharp corner where the sample starts and ends, the fitted curve matches the original curve well. At the other sharp point, the fitted value is smooth and cannot reach the sharp corner as shown in Figure 6. The third order polynomial functions were used. For a continuous and smooth curve as shown in Figure 7, although the curve has an intersection point that makes two closed ellipse shapes, the fitted curve can still match the original curve well with the 3rd order polynomial functions estimated using the moving least squares method with C=2.0.

Fitting accuracy and random error

Generally speaking, the order of polynomials used is directly related to curve fitting accuracies as can be seen in Table 1. Because of the limitation of computer precision, when the order was 11 or greater in our experiment, the fitting accuracy decreased dramatically using the least squares. Moving least squares resulted in higher accuracies of curve fitting with lower order polynomials.

Table 2 is a comparison of errors among the original curve, the digitized curve and the fitted curve from

Table 2. A Comparison of Position Deviations

	Position Deviation (cm)
Digitized vs. Original	0.0193
Fitted vs. Original	0.0096
Digitized vs. Fitted	0.0120

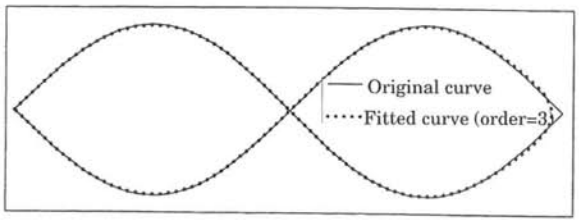


Figure 6. Fitting a curve with internal intersection and abrupt curvature changes. The fitted results were obtained using the 3rd order polynomials estimated with moving least squares.

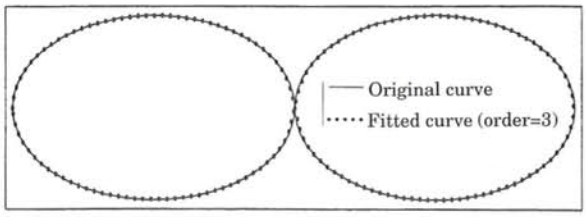


Figure 7. Fitting a curve with internal intersection with no abrupt curvature changes. The fitted version was produced by the 3rd order polynomials estimated from moving least squares.

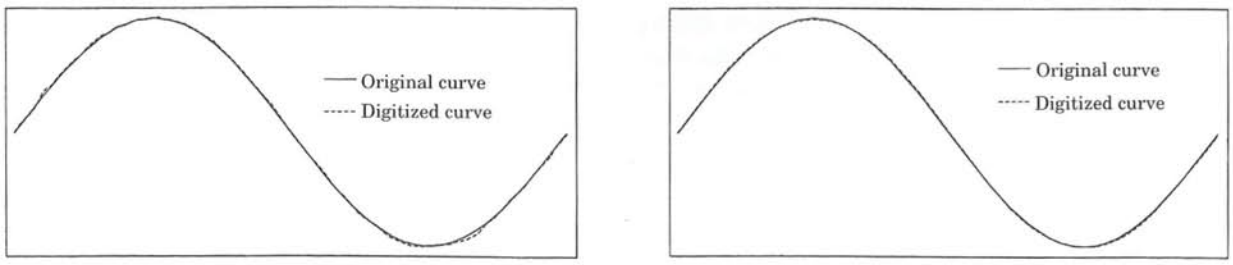
an experiment. The error of a digitized point was estimated from the minimum distance between the point to the original curve. It can be seen that the standard deviation between the digitized curve and the original curve is 0.193 mm while the standard deviation between the fitted curve and the original curve is 0.096 mm. Thus, the fitted curve has a higher accuracy than the digitized curve (Figure 8). Since the original true line may not be available in practice, the deviation of 0.120 mm between digitized points and the fitted curve may be taken as an approximation of the uncertainty introduced by digitization. Although the approximation tends to be smaller than the true digitizing error, it seems to be a proper measure of curve uncertainty for the application of the epsilon band model because the majority of the true curve will be within a 0.24 mm zone centered at the fitted curve.

Uncertainty modeling in map overlay

Map overlay is an important tool in geographical analysis. Through different operations such as intersection, matching, and merging (Pullar and Beard, 1990), two or more multisource data sets can be combined into one. Because different thematic maps are made by different people at different times using different data sources, the polygon boundaries in different digital maps do not exactly match. Even if the same polygon boundaries are separately digitized, the resultant maps will not exactly coincide due to digitization and other errors. Therefore, the operation of map overlay will generate many small spurious polygons (Goodchild, 1978). Spurious polygons are another source of uncertainty in spatial data bases.

To remove those spurious polygons, three strategies have been used: (1) randomly choose one side and delete the other side; (2) use a straight line to connect the two end points; (3) choose the line that has a higher accuracy or that is from a larger map scale and erase the other (Zhang et al, 1993).

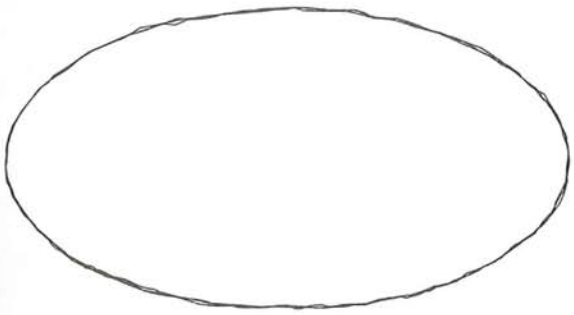
Polynomial curve fitting can be used as the fourth strategy to find a new line as an estimation of the true boundary line based on weighted least squares using all the points of a specific boundary from every layer. The new line has the minimum error if all the layers have the same accuracy. When the relative accuracies are different among layers to be overlaid, weights can be assigned to points in each layer in the weighted least squares estimation. The weight assigned to each layer should be made in



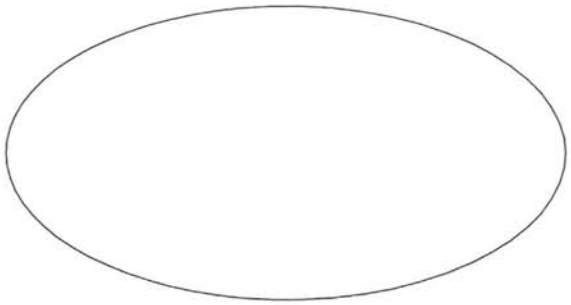
(a) (b)

Figure 8. A comparison of the accuracies between the original, a digitized, and a fitted curve. (a) The original curve and the digitized curve. (b) The original curve and the fitted curve.

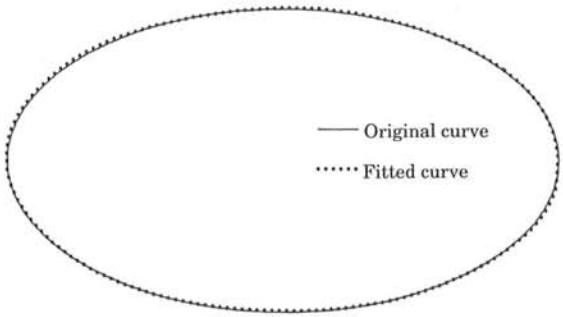
accordance to the accuracy of the layer, i.e., assign the layer of higher accuracy with a greater weight for the generation of the new boundary. The determination of weights should also be based on the scale of the source maps because maps of smaller scale tend to have less accurate boundary positions.



(b)



(a)



(c)

Figure 9. An example of multilayer curve fitting using the polynomial technique. (a) Original ellipse. (b) Three digitized versions. (c) Fitted curve with the 9th order polynomials estimated using weighted least squares technique.

Figure 9 shows some simulated results. Three curves digitized from the same original curve were regarded as three boundary lines each from a different source map. The fitted curve was calculated from the three digitized curves each having the same weight assignment. Table 3 lists the standard deviations between the fitted curve and the true curve and between the digitized points and the true curve.

An ellipse curve was produced from a mathematical equation and treated as the true curve (Figure 9a). It was digitized three times and the three digitized versions were overlaid (Figure 9b). Were they displayed at some larger scale, we would see many spurious polygons from the overlaid results along the boundary. Figure 9(c) shows the derived line using the 9th order polynomial functions based on all the points on the three different digitized curves with weighted least squares.

Table 3. A Comparison of the Position Deviations in a Map Overlay Experiment

	Position Deviation (cm)
Digitized vs. Original	0.0167
Fitted vs. Original	0.0078
Digitized vs. Fitted	0.0145

VII. SUMMARY AND CONCLUSIONS

In vector-based GIS, linear features are sampled in the form of discrete point series and represented by consecutively joined straight lines. The sampled points contain a large amount of errors and the representation method is not suitable for curve features. Few efforts have been made to improve this situation. The primary objective of this research was to seek appropriate methods to model linear features sampled in spatial databases and to determine uncertainties associated with the samples. We presented some methods based on a decomposition model implemented through parametric polynomial functions determined by least squares and moving least squares techniques to achieve the objective particularly for modeling continuous and smooth curve features.

A discrete point series can be decomposed into three components, the original data set, the systematic error, and the random error. Since the systematic error component may be detected through visualization and calibrated or removed through manual editing, we excluded it from our experiments. If the point series comes from a continuous and smooth curve such as a stream, contour, or a boundary of natural phenomena, through experiments we demonstrated that digitized

point series can be represented with polynomial functions. A single polynomial function whose coefficients are determined by the least squares technique or a group of polynomial functions whose coefficients are estimated with the moving least squares technique can be used to model a continuous and smooth curve.

In our experiments, we assumed that the random error of sampled points has a stationary normal distribution. Because a true curve is often unknown in real spatial databases, it is impossible to calculate the sample errors such as errors caused by digitization. We demonstrated the use of least squares and weighting least squares methods for estimating sample errors. The standard deviation between point data sampled along a curve and the fitted curve can be used in an epsilon band model to model uncertainties of the sample data, particularly digitized data.

The order of polynomials required to accurately model a curve is lower for moving least squares than that for the regular least squares. While the moving least squares technique gives higher curve fitting accuracies than regular least squares, it lacks computational efficiency.

These techniques may be used to estimate the uncertainties in map digitization, field survey using GPS units, multilayer map overlay, and to represent curves in spatial databases. Modeling and representing lines with polynomials have potential advantages in saving storage space, curve generalization, curve matching and object recognition.

Selecting suitable models and base functions for linear feature modeling and representation and developing appropriate uncertainty estimation methods for linear features in spatial databases warrant more research attention. Further test of the methods proposed here through experiments with curve features digitized from a map or collected in the field may provide important insights for better spatial data modeling and uncertainty estimation.

REFERENCES

[1] Blakemore, M.. 1984. Generation and error in spatial databases. *Cartographica*. 21:131-139.

[2] Bolstad, P.V., Gesler P., and Lillesand T.M.. 1990. Positional uncertainty in manually digitized map data. *International Journal of Geographical Information Systems*. 4(4):399-412.

[3] Brunsdon, C. and Openshaw, S.. 1993. Simulating the effects of error in GIS. in *Geographical Information Handling - Research and Application*. Edited by Mather, P.M.. John Wiley & Sons: England, pp. 47-61.

[4] Chrisman, N.P.. 1982. A theory of cartographic error and its measurement in digital data bases. *Proceedings, Auto-Caro 5*, pp. 159-168.

[5] Dierckx, Paul. (1993) *Curve and surface fitting with splines* New York : Oxford University Press.

[6] Dutton, G.. 1989. Modeling location uncertainty via hierarchical tessellation. in *Accuracy of Spatial Database*, Edited by M. Goodchild and S. Gopal. Taylor & Francis: New York, pp. 125-140.

[7] Gong, P., and Chen, J.. 1992. Boundary uncertainties in digitized maps I: some possible determination methods. *GIS/LIS'92*, pp. 274-281.

[8] Gong, P., Zheng, X., and Chen, J.. 1995. Boundary uncertainties in digitized maps: an experiment on digitization errors, *Geographic Information Sciences*. 1(2):65-72.

[9] Goodchild, M.F.. 1978. Statistical aspects of the polygon overlay problem. *Harvard Papers on Geographic Information Systems* (ed. G. Dutton) Vol. 6. Addison-Wesley, Reading, Mass.

[10] Goodchild, M.F.. 1990. Modeling error in spatial database, Proceedings, *GIS/LIS'90*, pp. 154-162.

[11] Graupe, D.. 1984. *Time Series Analysis, Identification and Adaptive Filtering*. Robert E. Krieger Publishing Company, Malabar, Florida.

[12] Grogan, T.A., Mitchell, O.R., Kuhl, F.P. and Chhabra, A.K. 1992. A performance comparison for global shape classification between Fourier descriptors and Walsh points methods using simulated data. *Remote Sensing Reviews*, 6(1), pp. 155-182.

[13] Gunther, O. and Wong, E.. 1990. The arc tree: an approximation scheme to represent arbitrary curved shapes, *Computer Vision, Graphics, and Image Processing*, 51:311-337.

[14] Janacek, G. and Swift, L. 1993. *Time Series: Forecasting, Simulation, Applications*.. England: Ellis Horwood Limited.

[15] Keefer, B.J., Smith, J.L. and Gregoire, T.G. 1988. Simulating manual digitizing error with statistical models, *Proceedings, GIS/LIS'88*, pp. 475-483.

[16] Lancaster, Peter. 1986. *Curve and surface fitting : an introduction*. Orlando : Academic Press.

[17] Lin, C.S. and Hwang, C.L.. 1987. New forms of shape invariants from elliptic Fourier descriptors. *Pattern Recognition*, 20(5):535-545.

[18] Maffini, G., Arno, M., and Bitterlich, W. 1989. Observations and comments on the generation and treatment of error in digital GIS data. in *Accuracy of Spatial Database*, Edited by M. Goodchild and S. Gopal. Taylor & Francis: New York, pp. 55-67.

[19] Perkal, J. 1966. On the length of empirical curves. Discussion paper 10, Ann Arbor, Michigan Inter-University Community of Mathematical

Geographers.
[20] Poiker, T.K. 1982. Looking at computer cartography. *Geojournal*. 6(3):241-249.
[21] Pullar, D. and Beard, K. 1990. Specifying and tracking errors from map overlay. Proceedings, *GIS/LIS'90*, pp. 79-87.
[22] Shumway, R.H. 1988. *Applied Statistical Time Series Analysis*. Englewood Cliffs, New Jersey: Prentice Hall.
[23] Wei, W.S. (1990). *Time Series Analysis*. Addison-Wesley Publishing Company: New York.
[24] Zhang, G., Ye, X., and Tan, F. 1993. Multi-tolerance data cleaning technique. *Proceedings of the first Symposium of CPGIS*. pp 224-231.
[25] Zhong, K and Hu, G. 1990. *Digital Signal Processing*. Beijing: Tsinghua University Press.

国家科委副主任徐冠华院士、科学院陈宜瑜副院长、自然科学基金会地学部林海副主任、中国地理信息系统协会李广源秘书长、国家基础地理信息中心陈军副主任等参加了CPGIS的五周年纪念会。

Managing Highway Accidents Involving Gaseous Hazardous Spills: A GIS Supported Simulation

Yong Lao

Social and Behavioral Sciences Center
California State University Monterey Bay, Seaside, CA 93955-8001, USA

Abstract
Managing highway incidents involving gaseous hazardous spills requires accurate assessment of potential risks to both the population and environment. Further, quick decisions must be made on how to effectively carry out emergency rescue and evacuation. In this paper we demonstrate that a Geographic Information System (GIS) provides an ideal tool to perform risk analysis and to assist emergency response. We have built a prototype GIS that integrates a dispersion model with Arc/Info to simulate gaseous hazardous spills under different circumstances. A case study is conducted based on the highway network and 1990 census data of the Greater Cincinnati Metropolitan area.

摘 要
在运送有毒气体的过程中，一旦发生事故，气体泄漏，对周围的环境及人口会造成极大的危害。处理此类事故，首先需要对有毒气体扩散的方向和危害程度进行研判，其次是安排紧急救援和人口疏散。本文阐述了如何利用地理信息系统(GIS)对上述两方面提供有效的决策支援。我们把气体扩散的数学模型与Arc/Info结合在一起来模拟各种有害气体在不同天气状况下随时间的扩散过程，并将这一成果应用于辛辛那提大都市区的一个实例研究。

I. INTRODUCTION

Over the last decade the study of hazardous material transportation has become a very popular field (for reviews, see Helander and Melachrinoudis, 1997; Erkut and Verter, 1995; William 1994; List et al., 1991). The importance of this field lies in the fact that over 1.5 billion tons of hazardous materials are shipped annually in the U.S., and there are enormous concerns expressed by the public, the government, and the industry (Lepofsky, et al., 1993). In general, managing hazardous material transportation involves several strategic tasks, as spelled out in the Hazardous Materials Transportation Uniform Safety Act of 1990: (1) assessing the risks to the population and/or environment based on such elements as type and quantities of hazardous materials, type of highways, population density, etc.; (2) siting toxic facilities and planning the best shipping routes; (3) acting on traffic division, emergency rescue, and population evacuation in the event of accidents. Many of these issues have been studied extensively with transportation engineering methods and models, but rarely they are considered and dealt with simultaneously.

As the development and application of Geographic Information Systems (GIS) continue to grow, many researchers as well as practitioners have realized the great potential of using this technology to help manage hazardous material transportation. The main idea is to integrate GIS and well developed hazardous material transportation models in a coherent environment for strategic planning, visualization, and analysis. For example, Lepofsky et al. (1993) discussed the suitability of GIS for five decision support aspects in hazardous material transportation: risk assessment, routing and scheduling, emergency preparedness, evacuation planning, and incident management. Comfort and Chang (1995) addressed the need of creating a distributed, intelligent spatial information system for disaster management. However, a closer look at current literature indicates that GIS applications in this area are mostly related to locating toxic sites and planning optimal shipping routes (for example, Anders and Olsten, 1990; List and Turnquist, 1994; Parentela and Sathisan, 1995). Not much work has been done to address the issues of how to predict the potential dangers and act on emergency rescue

1082-4006/97/0301~2-20$3.00

once an incident occurs. It is against this background that the current research project is developed.

The objective of our study is to demonstrate that a Geographic Information System (GIS) provides an ideal tool to deal with risk assessment and emergency rescue in an event of a highway accident involving gaseous hazardous material transportation. We will embed a gas dispersion model in a GIS environment so that the entire process of model input, output, display and analysis become interactive and integrated with powerful GIS functions. Specifically, we are concerned with using GIS (1) to interactively simulate the gaseous hazardous release and map the toxic area; (2) to calculate the number of population exposed and their risk level; and (3) to select routes for emergency response teams. We have chosen to focus on highway accident because it has a higher probability compared to other transportation modes. It is estimated more than 85 percent hazardous release during shipping occurs on highways (Hardwood, et al., 1989). Further, a highway incident involving hazardous material spills often cause enormous damages to people and environment, especially in the urban area. In order to effectively carry out emergency rescue and evacuation in the event of an accident, questions regarding the characteristics of hazardous material release and the severity of health consequence need be answered quickly.

The remainder of this paper is organized as follows. Section two discusses the dispersion model that predicts the toxic level of hazardous gas release. Section three demonstrates how the dispersion model is implemented in a GIS environment. Section four presents our case study. Section five concludes the paper and points out future research directions.

II. THE DISPERSION MODEL

Dispersion models are quantitative models that estimate the dispersion of atmospheric pollutants and predict the toxic corridor (see, for example, Munger, et al., 1983; Ryckman and Peters, 1983; Klug, 1984). In case of a highway accident involving gas spills, a dispersion model is employed to answer the following questions:

Where will the toxic cloud go?
When will the toxic cloud get to a place?
What are the impacted areas?
How serious will the population and/or environment be affected?

In order to correctly answer these questions, we must consider a variety of elements that may influence the outcome of a gaseous spills. These elements are:

- the characteristics of the gas (density, molecular weight, potential danger, etc.)
- the release rate and duration (sonic vs. subsonic, instantaneous vs. continuous)
- meteorological conditions (wind speed, direction, atmospheric stability, etc.)
- the topography of the surrounding area (roughness of surface, buildings, trees, etc.)

Our handling of these elements is based on the standard engineering approach, as described by William (1994). We assume the properties of various gaseous chemicals are known and have been stored in the database. Given the type of gas, users can immediately retrieve its physical and chemical attributes. The release rate is determined according to the gas dynamic theory of ideal, adiabatic compressible flows with standard equations. Depending on the pressure of the gas in the container, the flow can be classified under *sonic* or *subsonic* flow. If the criterion described in Equation (1) is satisfied, the flow is sonic (i.e., the flow velocity is equal to the sound speed in the gas), otherwise it is subsonic:

$$P / P_a \geq \left[(\gamma+1)/2\right]^{\gamma/(\gamma-1)} \qquad (1)$$

where:

P= absolute tank pressure (N/m^2)
P_a = absolute ambient pressure (N/m^2)
γ = gas specific heat ratio (usually 1.5)

The release rate for sonic and subsonic flow is calculated by Equation (2) and Equation (3) respectively (William, 1994).

$$Q = C_d A P\left[(\gamma M / RT)(2/(\gamma+1))^{(\gamma+1)(\gamma-1)}\right]^{1/2} \qquad (2)$$

$$Q = C_d A\left\{2\rho_g P(\gamma/(\gamma-1))\left[(P_a/P)^{2/\gamma} - (P_a/P)^{(\gamma+1)/\gamma}\right]\right\}^{1/2} \qquad (3)$$

where:

Q = discharge (kg/s)
C_d = orifice discharge coefficient
A = area of flow (m^2)
M = gas molecular weight (kg/kg-mole)
R = absolute gas constant (8.31 X 10^3 J/kg-mole/K)
T = absolute gas temperature in container (K)
ρ_g = gas density (kg/m^3)

The parameters required for estimating the amount of release are difficult or impossible to obtain when the accident is reported to the emergency response unit. Hence, default parameters are used to run the model until accurate information is available.

Given the gas discharge rate as input, the well tested and most widely used Gaussian model is chosen to calculate the concentration distribution of the spilled gas. The Gaussian model assumes that the atmospheric turbulence is random, which results in Gaussian distribution of toxic clouds in the vertical and horizontal downwind directions. The equation for continuous release is as follows (Gifford, 1968):

$$\chi = (Q/\pi\sigma_y\sigma_z\mu)\exp\left\{-[(y^2/2\sigma_y^2)+(h^2/2\sigma_z^2)]\right\} \quad (4)$$

The formula for instantaneous release is as follows:

$$\chi = (Q/\sqrt{2}\pi^{3/2}\sigma_x\sigma_y\sigma_z\mu)\exp\{-[((x-\mu t)^2/2\sigma_x^2)+ (y^2/2\sigma_y^2)+(h^2/2\sigma_z^2)]\} \quad (5)$$

where:

χ = concentration of the released material (g/m^3)
Q = material release rate (g/s) for continuous release or total release (g) for puff releases
$\sigma_x, \sigma_y, \sigma_z$ = standard deviations of the concentration distribution to the plume centerline (m) in the x, y, and z directions
y = horizontal distance perpendicular to the plume centerline
z = vertical distance perpendicular to the plume centerline (m)
x = distance downward (m)
h = effective release height from the ground (m)
μ = wind speed (m/s)
t = time since release (s)

$\sigma_x, \sigma_y, \sigma_z$ are diffusion parameters estimated in accord with the atmospheric stability. Atmospheric stability is usually measured by the vertical temperature gradient. An unstable temperature profile produces more turbulence whereas a stable temperature suppresses turbulence. When a highway accident is reported, the atmospheric stability for the incident location is very likely to be unknown. Therefore, as illustrated by Table 1, we adopt the Pasquill stability types which are defined according to surface wind speed and insolation conditions (Gifford 1968). The Pasquill scheme varies from extremely unstable conditions (type A) to moderately stable conditions (type F).

Once a Pasquill stability type is identified for the accident location, the formulas in Table 2 recommended by Hanna et al. (1982) are used to compute the diffusion parameters for open-country and urban conditions. The formulas are applicable to continuous toxic plumes with 10 to 30 minute exposures (William, 1994).

Finally, the concentration of the gas is converted to normally desired units of parts per million (PPM) by the formula:

$$PPM = 23691\chi/M \quad (6)$$

where M = Molecular weight of the gas.

III. THE GIS IMPLEMENTATION

The Gaussian gas dispersion model is implemented as a raster GIS model in Arc/Info. We chose using raster data because the cell-based format provides a better visual representation and analysis of gaseous spills that move continuously over space. Our objective is to simulate highway gas spills and

Table 1. Pasquill stability types

Surface Wind speed, m/sec	Daytime Insolation			Nighttime Conditions	
	Strong	Moderate	Slight	Thin overcast or ≥3/8 cloudiness	≤3/8 cloudiness
< 2	A	A-B	B		
2	A-B	B	C	E	E
4	B	B-C	C	D	E
6	C	C-D	D	D	D
> 6	C	D	D	D	D

A: Extremely unstable conditions
B: Moderately unstable conditions
C: Slightly unstable conditions
D: Neutral conditions
E: Slightly stable conditions
F: Moderately stable conditions

to perform incident analysis through friendly graphical user interface.

The Data

Four types of data need to be collected: data on the physical attributes and release characteristics of gas, meteorological data, population data, and environmental data (landuse, road network, topography, etc.). Figure 1 is the data flow diagram that shows how the spatial and attribute data are prepared and processed. First, based on the type of hazardous gas and discharge type, the attributes of the gas are used to calculate the release rate (see Equation (2) and (3)). Second, based on the location of the accident, we immediately know whether this is a predominantly urban or rural area. We could then compute the diffusion parameters according to the atmospheric stability conditions (see Table 1 and 2). Third, based on the wind speed, direction, and the time since incident, we create map layers representing location factors (direction, vertical and horizontal distance) with respect to the plume center line. In addition, population data, road network data, and other environmental data are used for query, routing and visualization purpose.

The Graphical User Interface (GUI)

A friendly graphical user interface (GUI) is designed to facilitate the dispersion modeling process. It consists of a collection of menus, buttons, and tools through which users communicate with the prototype system. As shown in the main menu (Figure 2a), the GUI inter-connects with users, the database, and all the functional elements of the prototype system. Each button in the main menu leads users to a set of sub-menus and tools for data entry, parameter selection, and model operation.

There are five major components in the main menu. The "Spill Characteristics" button opens a sub-menu (Figure 2b), which allows users to choose one special type of hazardous gas, its physical attributes, and discharge type. Similarly, the "Meteorological Conditions" button provides a sub-menu (Figure 2c) for simulating different Pasquill atmospheric stability types and wind speed. Clicking on the "Incident location/Display" button will display the highway network of the study area and a cross cursor. Users can specify the reported accident location on the highway. The system displays an error message if the point does not lie on the highway network. The "Run" button results in a sub-menu (Figure 2d) through which users could input the time since the incident, as well as the desired time increment which is needed to automatically update the display. The system prompts for confirmation before running the model for the next time increment. Finally, the "Query" button leads to the overlay and routing process which estimate the population affected by the accident and draws the

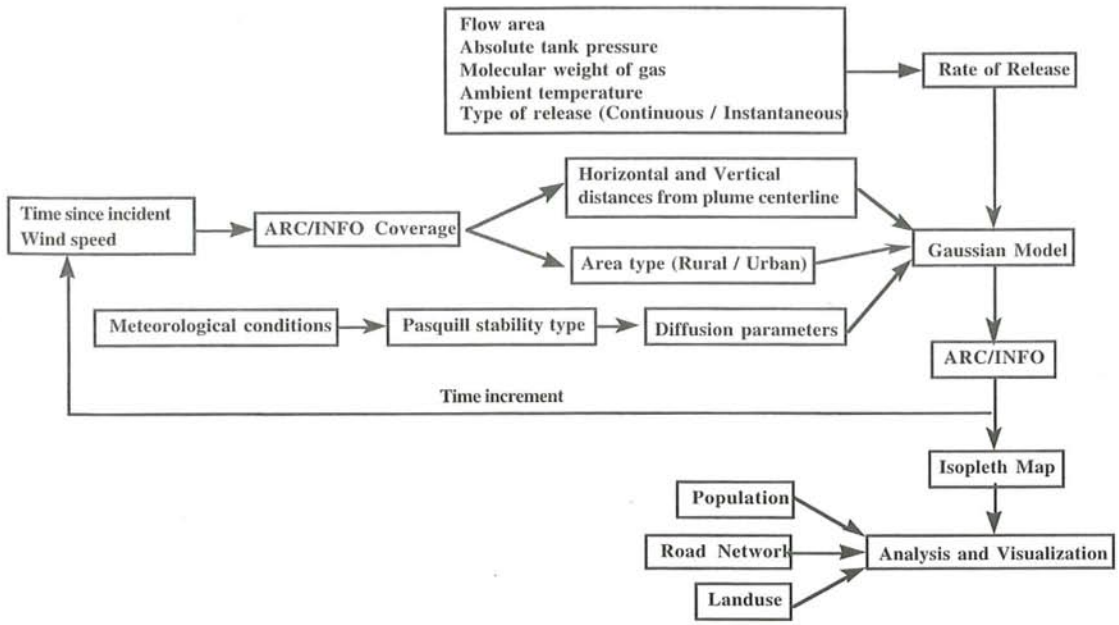


Figure 1. The data flow of the disperse modeling

Table 2. Formulas recommended by Hanna et al (1982) for $\sigma_y(x)$ and $\sigma_z(x)$ (100 m < x < 10 km)

Pasquill stability type	σ_y, m	σ_z, m
Open-Country Conditions		
A	$0.22x(1 + 0.0001x)^{-1/2}$	$0.20x$
B	$0.16x(1 + 0.0001x)^{-1/2}$	$0.12x$
C	$0.11x(1 + 0.0001x)^{-1/2}$	$0.08x(1 + 0.0001x)^{-1/2}$
D	$0.08x(1 + 0.0001x)^{-1/2}$	$0.06x(1 + 0.00015x)^{-1/2}$
E	$0.06x(1 + 0.0001x)^{-1/2}$	$0.03x(1 + 0.0003x)^{-1}$
F	$0.04x(1 + 0.0001x)^{-1/2}$	$0.016x(1 + 0.0003x)^{-1}$
Urban Conditions		
A-B	$0.32x(1 + 0.0004x)^{-1/2}$	$0.24x(1 + 0.0001x)^{-1/2}$
C	$0.22x(1 + 0.0004x)^{-1/2}$	$0.20x$
D	$0.16x(1 + 0.0004x)^{-1/2}$	$0.14x(1 + 0.0003x)^{-1/2}$
E-F	$0.11x(1 + 0.0004x)^{-1/2}$	$0.08x(1 + 0.00015x)^{-1/2}$

fastest route from the nearest emergency rescue unit to the incident location.

The Analytical Process

As illustrated in Figure 3, the modeling and visualization process is implemented by means of a collection of scripts in Arc Macro Language (AML). With user providing the incident location, type of atmospheric stability, wind speed, and the time since the release, a temporary ARC/INFO grid is created from the existing highway network coverage. The extent of the grid is determined by multiplying the current wind speed by the estimated maximum release duration time. Initially all the cells in the temporary grid coverage have the value of NODATA except for the cell representing the incident location. Then the following steps are taken to implement the dispersion model:

Step 1. The Euclidean distance function is used to calculate the distance closest to the source cell. The result is a new grid with each cell value representing its distance from the incident location. Similarly, the Euclidean direction function is used to create a new grid with each cell value representing its direction (in degrees) from the incident location. Since the plume center line follows the given wind direction, each cell's angle (in degrees) with respect to the plume center line can be easily obtained by subtracting the wind direction value from the direction grid. Based on the angle, each cell's horizontal and vertical distances (x and y) from the plume centerline are computed using the trigonometric calculations. The x and y values are stored in separate grids.

Step 2. The Pasquill stability type is selected according to the current meteorological conditions (see Table 1). The result is subsequently used to determine the horizontal and vertical diffusion parameters ($s_y(x)$ and $s_z(x)$ in Table 2).

Step 3. Depending on the type of release (instantaneous or continuous) Equation (4) or (5) as well as Equation (6) are implemented by using the grid algebra to calculate the concentration of the gas distribution (in PPM) for each cell in the grid. The concentrations are displayed in different colors in the resulting isopleth map. In the continuous release case, it is assumed that the gas will not

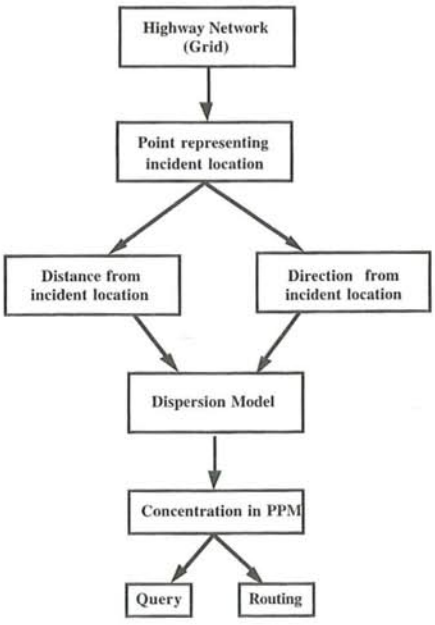


Figure 3. The logic diagram of the analytical process

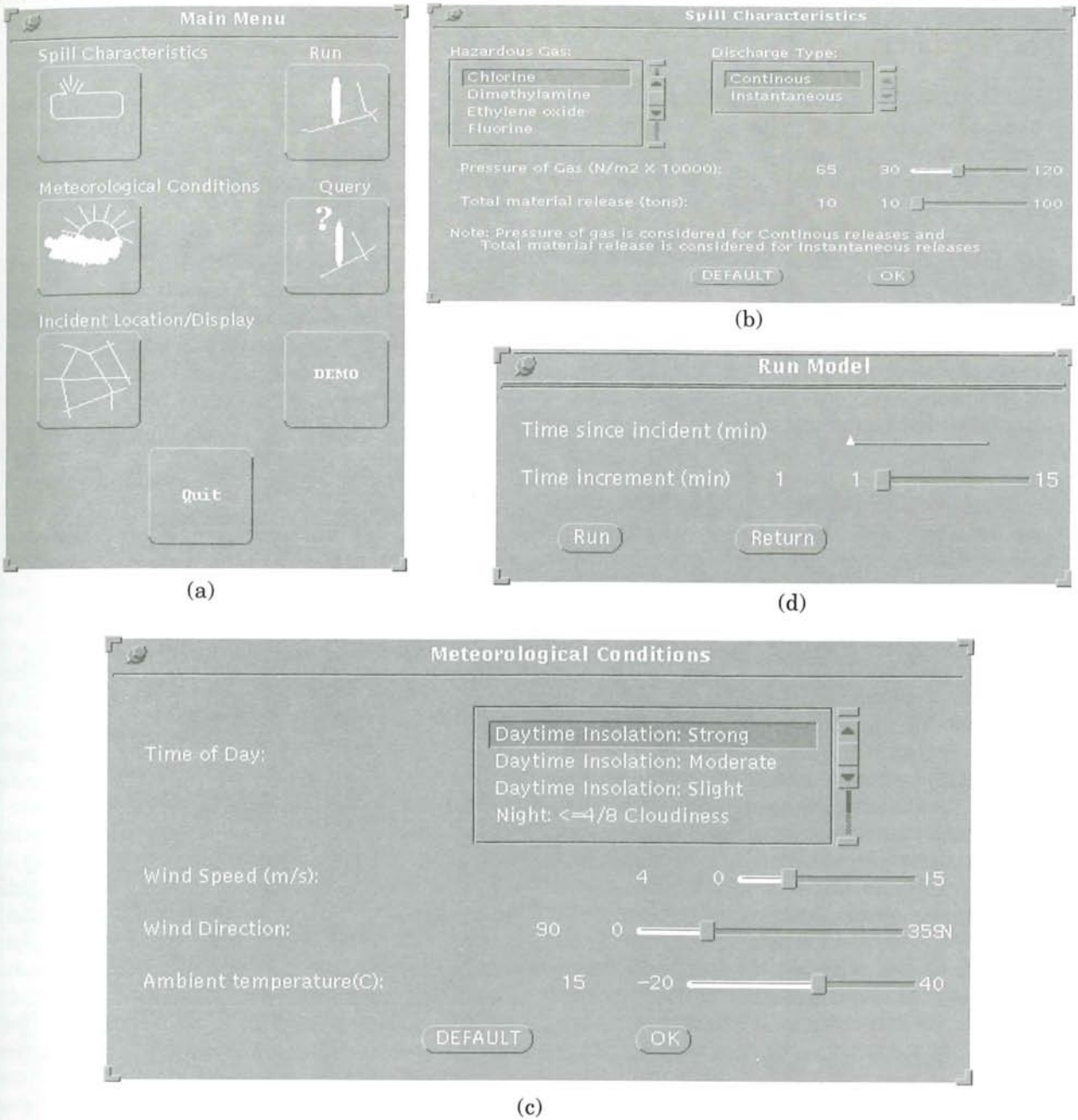


Figure 2. The graphical user interface(GUI). (a) the main menu; (b) the spill characteristics sub-menu; (c) the meteorlogical conditions sub-menu; and (d) the run model sud-menu.

spread beyond the distance covered by the wind in the same time period. As the time increases, new coordinates are sent to the dispersion model, and the display is updated accordingly.

The resulting isopleth map can be queried after each time increment. The system gives the concentration (in PPM) at any desired cell location, and the distance from that cell to the incident source. In addition, the population being affected by the spill may be estimated by carrying out the following raster overlay analysis:

Step 1. Create a grid with cell values representing the estimated number of people within each cell (or any phenomena of interest such as landuse type, property value, number of buildings, etc.). Create a binary grid with cell value equal to 1 for those that

are within the plume and 0 for the ones outside the plume.

Step 2. Multiply the above two grids to obtain the population within the toxic plume. The sum of the cell values thus represents the total population affected by the incident since the release of hazardous gas.

If users intend to know the number of people exposed to a particular range of concentration, they could simply repeat the above two steps except making the binary grid only include the cells that fit into the range of concentration. In a similar fashion, one could interactively choose any grid cell(s) at any location(s) within the plume and find out the total population in that area.

Routing analysis for the emergency response team can be done using the Arc/Info network analysis functions. The key operations are (a) to display the highway network on the top of the plume grid and to 'snap' the incident source cell to the nearest node on the network. The result is the 'destination' for the shortest path function; (b) to geocode the all the fire stations or emergency response centers in the study area so that the emergency response units that are within certain time windows to the incident location may be quickly identified and ranked (the closest, the 2nd closest, the third closest, etc.); and (c) to run the shortest path algorithm and draw the resulting paths. Since our routing analysis did not consider traffic conditions and the capacity (personnel, equipment, etc.) of fire stations, the result could be unrealistic. Therefore, it is necessary to have an expert who is able to adjust these computer selected routes based on traffic conditions, facility capacity and other unforeseen elements.

IV. A CASE STUDY

To test the prototype system, a hypothetical case study was conducted simulating a highway accident in the Greater Cincinnati Metropolitan area. Chlorine was selected as the hazardous gas for the disperse model. The GIS database includes population and highway network (intestates, US routes and state routes) for the Ohio-Kentucky-Indiana tri-state region. The cell resolution of the raster data is 25 meters.

The incident involving continuous Chlorine spills happened on Interstate highway 75 in a heavily populated suburban area in Cincinnati. Based on the user defined chemical characteristics, release types, and surface meteorological data, the dispersion model computes the size and moving direction of the plume at a 3 minute increment. Figure 4a presents the isopleth map for the distribution of Chlorine gas 15 minutes since the release. The characteristics of the release, the meteorological conditions, and the concentration level of Chlorine are displayed in the legend. The map clearly shows the movement of toxic clouds and the exposure during the downwind dispersion of the plume. By overlaying the isopleth map with the population map, we can quickly evaluate how serious the exposure is. In case of Chlorine, it is known at 1 to 5 ppm, irritation of the nose, respiratory tract, and eyes starts. At 15 ppm, irritation is considered severe (The Chlorine Institute, 1991). Therefore, as shown in Table 3, within 15 minutes of the accident, approximately 1600 people need be evacuated from the downwind area (zip code 45069 and 45246) where the predicted concentration exceeds 10 ppm. Such information is very critical to evacuation planners or emergency rescue teams. In reality, though, the calculation may be overestimated due to the fact a) the dispersion model is usually conservative in predicting the toxic level; b) there are less people in the residential area during normal business hours; c) some people may have taken evasive actions by themselves and some may stay inside a shelter.

Table 3. Population exposed to the chlorine spill 15 minutes since incident

Level of Concentration	Number of People Affected	Zip Code
> 100 ppm	741	45069
10 - 100 ppm	871	45069, 45246
5 - 10 ppm	445	45069, 45246
4 - 5 ppm	743	45069,45246
3 - 4 ppm	2084	45069, 45246
≤3 ppm	8450	45069, 45246, 45218, 45240

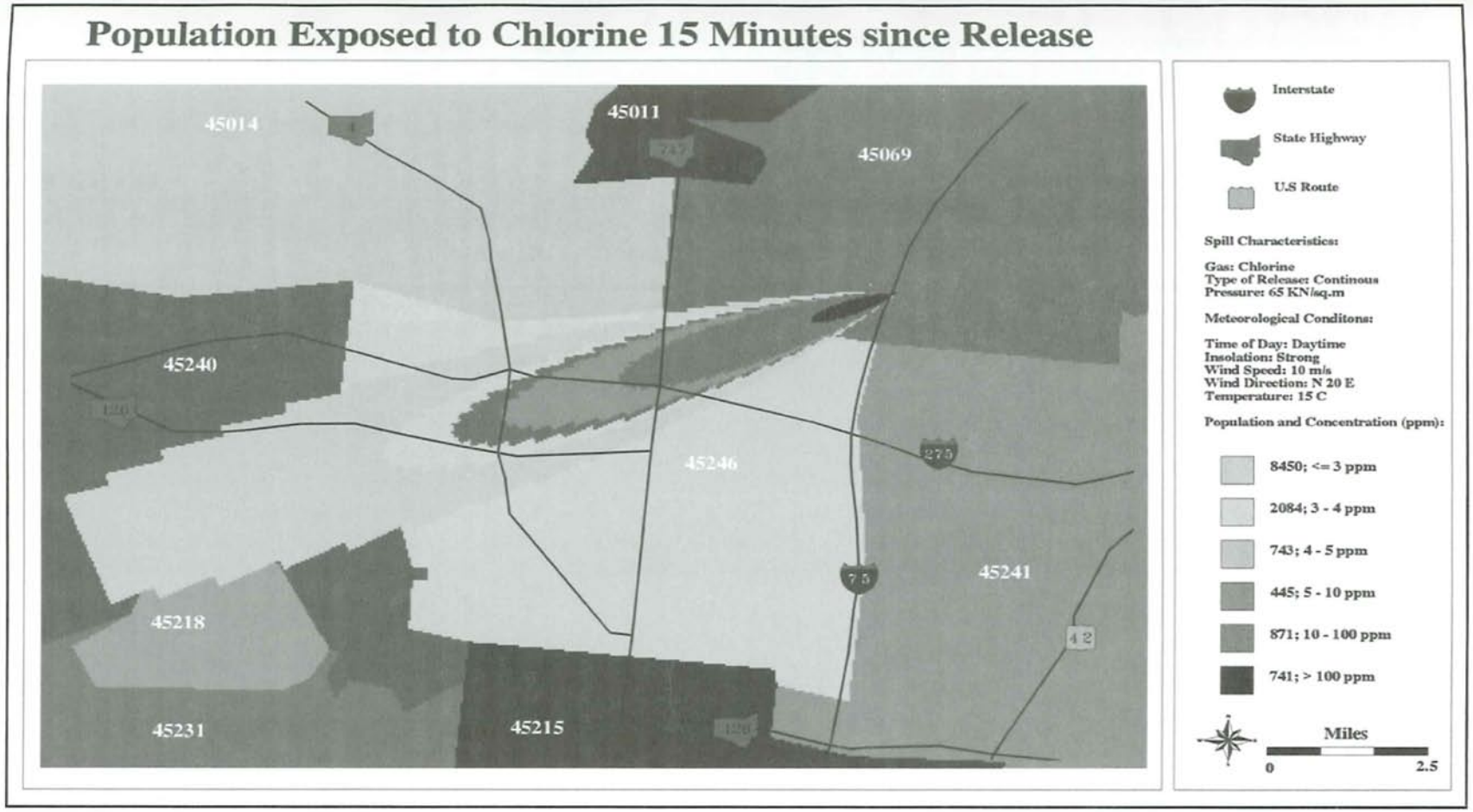


Figure 4. (a) The isopleth map for the distribution of Chlorine gas 15 minutes since the release

Figure 4b presents the route choices for emergency response vehicles. There are two emergency response units in this area, both within the 15 minute travel window to the incident location. The nearest highway entries are first identified for them, then the shortest paths are calculated from the entries to the incident location. One should notice that emergency vehicle routing is a relatively easier process compared to evacuation routing. The latter involves more sophisticated elements (see Abkowitz, 1998; L›vØs, 1998) and is beyond the scope of this paper.

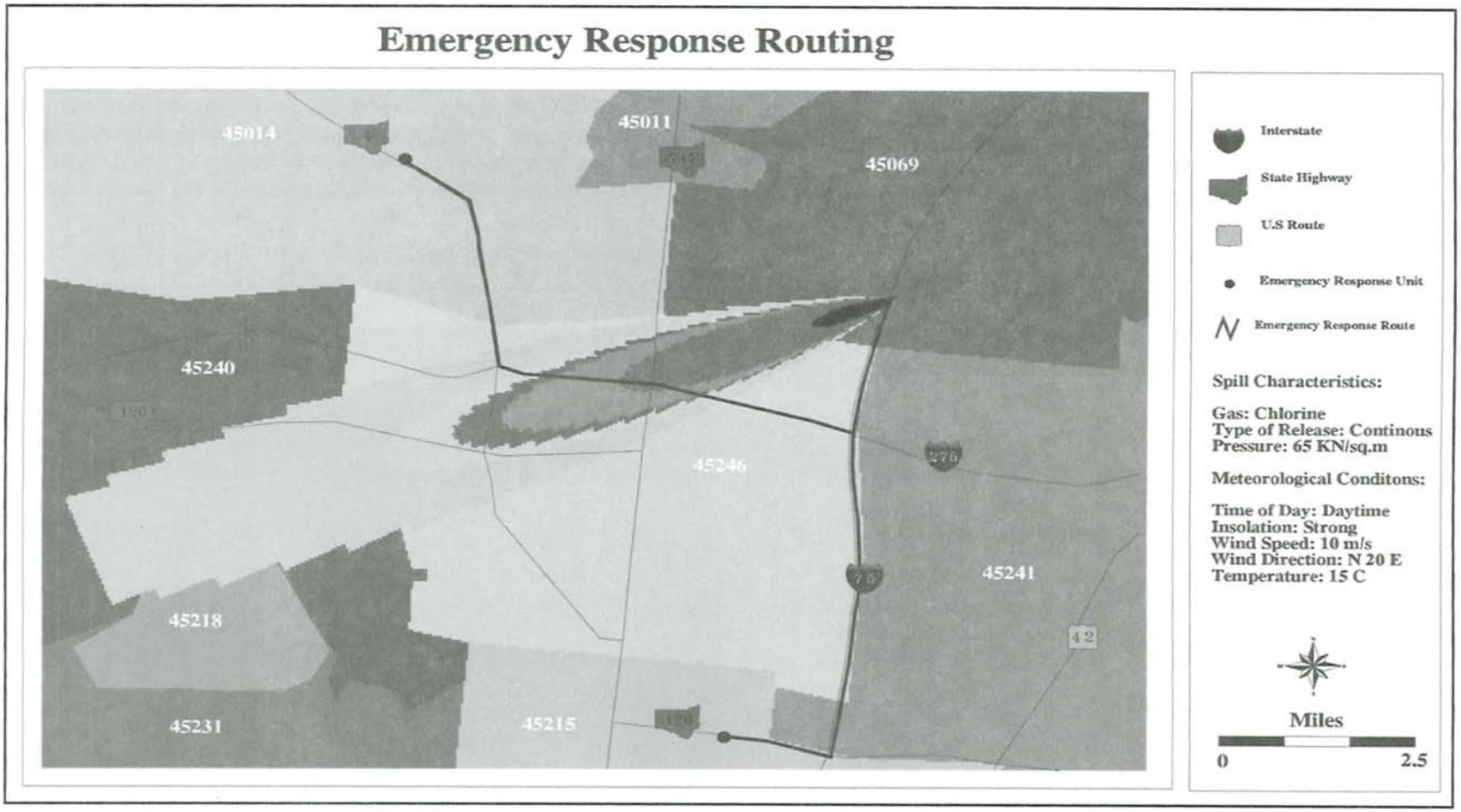


Figure 4. (b) The route choices for emergency response vehicles

V. SUMMARY

In this paper we have presented a GIS-supported approach to managing highway incidents involving hazardous gas spills. Using a simulated case study in Greater Cincinnati Metropolitan area, we demonstrated how the widely used Gaussion dispersion model can be implemented within a raster GIS environment. The result is the timely updated plume-shaped area indicating the concentration level of hazardous pollutants. Through friendly graphical user interface, one can quickly estimate the population or any phenomena of interest being affected within the given time frame. Further, routing analysis can be carried out for emergency response teams with powerful vector GIS

analysis functions.

There are two major advantages of this approach. First, it facilitates interactive model input, analysis, and visualization. Second, it allows users to experiment and simulate gaseous spill incidents at any highway location under different circumstances with respect to hazardous materials, meteorological conditions, and release types and duration. The simulated case studies can provide rich information for transportation engineers, managers, and emergency response planners.

There exists plenty of room for improvement and extension of this study. Future research should be directed to the following areas:

- Better dispersion models which are able to consider the changes in wind direction and wind speed with time; changes in terrain, obstacles etc.
- Wind direction database (one of the attributes) for the whole area which will substitute for lack of real time information.
- Routing strategies which consider capacity of fire stations and traffic volumes.
- Models to evacuate the affected population.
- Case studies that are based on real world incidents.

REFERENCES

[1] Abkowitz, M. and Meyer, E. 1998. Technological advancements in hazardous materials evacuation planning. *Transportation Research Record*, 1522: 116-121.

[2] Anders, C. and Olsten, J. 1990. GIS risk analysis of hazardous materials transport. In *State and Local Issues in Transportation of Hazardous Materials: Towards a National Strategy*, 248-261.

[3] Comfort, L. K. and Chang, S. K. 1995. A distributed, intelligent, spatial information system for disaster management: a national model. *Proceedings of ESRI User Conference*.

[4] Gifford, F. A. Jr. 1968. An outline of the theories of diffusion in the lower layer of the atmosphere. *Meteorology and Atomic Energy,* United States Atomic Energy Commission.

[5] Erkut, E. and Verter, V. 1995. Hazardous materials logistics, in *Facility Location: A Survey of Applications and Methods*, Chapter 20, Z. Drezner (ed.). Springer-Verlag, New York, 467-506.

[6] Hanna S. R , and Drivas, P. J. 1982. *Guidelines for the use of Vapor Cloud Dispersion Models*. American Institute of Chemical Engineers.

[7] Hardwood, D., Russell, E. and Viner, J. G. 1989. Characteristics of accidents and incidents in highway transportation of hazardous materials, *Transportation Research Record*, 1245: 23-33.

[8] Helander, M. and Melachrinoudis, M. 1997. Facility location and reliable route planning in hazardous material transportation. *Transportation Science*, 31(3): 216-226.

[9] Klug, W. 1984. Atmospheric dispersion modeling concepts and approaches. In W. F. Dabberdt (Ed), *Atmospheric Dispersion of Hazardous / Toxic Materials from Transport Accidents*, 63-74. New York: Elsevier Science Publishers.

[10] Lepofsky, M., Abkowitz, M. D., Cheng, P. D. M. 1993. Transportation hazard analysis in integrated GIS environment. *Journal of Transportation Engineering*, 119(2): 239-254.

[11] List, G. and Mirchandani, P. 1991. An integrated network/planar multiobjective model for routing and siting for hazardous materials and wastes. *Transportation Science*, 25(2): 146-156.

[12] List, G. and Turnquist, M. A. 1994. Routing and emergency response team siting for high-level radioactive waste shipments. Paper prepared for the special issue on Emergency Management Engineering of the IEEE Transactions on Engineering Management.

[13] LvØs, G. G. 1998. Models of wayfinding in emergency evacuations. *European Journal of Operational Research*, 105: 371-389.

[14] Munger, B., Hurt, P., Hayden, P. and Irvine, E. 1983. Integrated modeling of the release and dispersion of hazardous gases in the atmosphere. *Transportation Research Record*, 902: 25-29.

[15] Parentela, E. and Sathisan, S. 1995. GIS-based allocation of emergency response units along a transportation route. *Proceedings of ESRI User Conference*.

[16] Ryckman, M. D. and Peters, J. L. 1983. Toxic corridor projection models for emergency response. *Transportation Research Record*, 902: 25-29.

[17] William, R. R. 1994. *Hazardous Materials Transportation Risk Analysis*. New York: Van Nostrand Reinhold.

Effects of Changing Spatial Scale on the Results of Statistical Analysis with Landscape Data: A Case Study

Jianguo Wu*, Wei Gao§ and Paul T. Tueller†

*Department of Life Sciences, Arizona State University West, Phoenix, AZ 85069, USA
§Department of Geography, Arizona State University, Tempe, AZ 85287, USA
†Department of Environmental and Resource Sciences, University of Nevada, Reno, NV 89557, USA

Abstract

The effect of scale on spatial analysis has long, but sporadically, been recognized in human geography and more recently and acutely in landscape ecology. As the number of studies directly and systematically addressing scale effects is still limited, it remains unclear how results of different statistical analyses are affected by changing scale for different landscapes, or whether or not such effects can be predicted and, if so, in what situations. However, it is certain that erroneous conclusions may result if scale effects are not considered explicitly in spatial analysis with area-based data. With widespread use of remote sensing data and GIS, a better understanding of the issue of scale effects is much needed. The main purpose of this study, therefore, was to examine how results of statistical analysis respond to a systematic change in the scale of analysis. Specifically, we investigated how the relationship between landscape metrics (local landcover diversity and richness indices) and independent variables (TM bands and vegetation indices) would change with different sample sizes and mathematical representations of variables. The landscape under study is the Minden area of Nevada in the western Great Basin. Four different sample sizes (19x19, 15x15, 11x11, and 5x5 pixels) and four different representation forms (variance, mean, variance-mean ratio, and coefficient of variation) of the variables were used in all statistical analyses. We systematically examined the effects of changing sample size and representations of variables on the results of regression, analysis of variance, and correlation analysis. The results indicated that the relationship between landscape metrics and TM bands and vegetation indices was affected considerably by the change of sample size. Both the R^2 value and the level of statistical significance of the relationship tended to increase as sample size increased. In addition, the results of ANOVA showed that the relative importance of the TM bands and vegetation indices in the relationship varied with sample size as well. Although the spatial pattern of local-scale (or "neighborhood") diversity and richness of land-cover types in this Great Basin landscape could be adequately quantified using spectral information-based variables, the results and accuracy of such a analysis depended on both landscape composition and sample size. The linear response of the statistical relationship to the change in sample size over some range of scales indicated that scale effects could be readily predicted in certain cases. However, in general, because scale effects can further be complicated by the choice of variables and the idiosyncrasy of particular landscapes, the predictability of scale effects seems to be confined only to certain domains of scale. To find these domains multiple-scale or hierarchical analysis must be performed. This study further supports that the modifiable areal unit problem is a common one across the disciplinary boundaries of geography, ecology and other earth sciences. Unraveling the problem not only will improve our understanding of pattern and process in nature, but also will have important implications for appropriate use of remote sensing data and GIS.

摘 要

尺度对空间数据统计分析结果的影响（或称尺度效应），作为可塑性面积单元问题（MAUP）的一个方面，很早即为地理学家注目，但研究范畴和深度尚有局限性。近年来，尺度效应在景观生态学研究中引起极大重视。从实质上讲，景观生态学研究最突出的特点就是强调空间异质性、态学过程和尺度三者之间的关系。然而，究竟尺度如何影响对于不同景观的相同统计分析方法和对于相同景观的不同统计分析方法？尺度效应是否可以预测？这些问题仍然有待于进一步研究。可以肯定的是，未能明确地考虑尺度效应的空间统计分析或景观格局分析往往会导致错误的结论。随着遥感数据和地理信息系统（GIS）的广泛应用，对尺度效应问题的深入研究日趋重要。本文的主要目的是探究景观指数（景观多样性、优势度和丰富度）与遥感变量（Lansat TM 波段和植被指数）之间的统计关系是否随分析尺度而变化；若此，是如何变化的？这一尺度效应可否预测？文中所用景观空间数据来自地处美国大盆地（The GreatBasin）内华达州（Nevada）内的 Minden 地区。研究表明，回归和相关统计分析结果均受到尺度变化的显著影响。对于同一种统计分析方法而言，尺度效应尚随不同统计变量（如平均值，方差，均值/方差比，变异系数）而异，从而将尺度效应的研究进一步复杂化。尽管如此，尺度效应在一定尺度域（domain of scale）中有时表现出线性关系，即易预测性。要确定这些易预测尺度域，则必须要采用多尺度或等级空间分析方法。该研究进而表明，可塑性面积单元问题不是一个偶然现象，

1082-4006/97/0301~2-30$3.00

而是空间格局和过程研究中的一个普遍问题。对这一问题的深入研究将无疑会促进我们对自然界空间格局和各种生物、非生物过程及其相互关系的认识，而且对遥感数据和地理信息系统的正确应用亦具有重要意义。

I. INTRODUCTION

Spatial pattern has important effects on a variety of physical and ecological processes, including flows of energy and nutrients and dispersal and movement of plants and animals (Turner, 1989; Risser, 1990; Wiens et al., 1993; Wu et al., 1993; Hunsaker et al., 1994; Wu and Levin, 1994, 1997). To understand the interactions between pattern and process it is necessary to quantitatively characterize spatial heterogeneity over a range of scales. Because today's spatial pattern results from yesterday's dynamic processes, pattern analysis may potentially reveal critical information on properties of underlying processes. Landscape ecology, focusing on the study of the reciprocal relationship between spatial pattern and ecological processes, provides a new conceptual framework for understanding how nature works (Pickett and Cadenasso, 1995; Wu and Loucks, 1995). In recent years, numerous studies have been carried out to quantify landscape patterns using various spatial analysis methods (O'Neill et al., 1988; Turner and Gardner, 1991; Cullinan and Thompson, 1992; Plotnic et al., 1993; Wickham and Riitters, 1995; Riitters et al., 1995; Jelinski and Wu, 1996; Qi and Wu, 1996). In general, both promises and problems have been found regarding the plethora of techniques used in landscape pattern analysis (see Riitters et al., 1995, Jelinski and Wu, 1996).

Remotely sensed data and geographic information systems (GIS) have been increasingly used to facilitate large-scale studies in landscape ecology (Iverson et al., 1989; Roughgarden et al., 1991; Turner and Gardner, 1991). Landsat Thematic Mapper (TM) and NOAA satellite AVHRR data, in particular, have been widely adopted in landscape ecological studies. Based on the features of reflectance and absorption of vegetation to electromagnetic radiation, a number of vegetation indices have been developed from several TM bands (e.g., Tueller, 1989). Both the spectral values of the different TM bands and vegetation indices derived from them can be correlated with various characteristics of landscapes (e.g., Tueller, 1989, Rey-Benayas and Pope, 1995).

Landscapes are hierarchically structured in space, within which pattern and processes operate over a range of scales (O'Neill et al., 1991; Wu and Loucks, 1995). Detected spatial pattern usually varies with the scales of observation, measurement, and data analysis. Therefore, any analysis based on a single scale may provide little (or even misleading) information on the overall landscape structure under study (Wu and Loucks, 1995; Jelinski and Wu, 1996). Two concepts, grain and extent, have been particularly useful for making landscape pattern analysis scale-explicit, thus facilitating communication and comparison of the results. Grain is the "smallest unit of measure" or "the first level of spatial resolution possible with a given data set", whereas extent is the "cover" or "the total area of the study" (sensu Turner and Gardner, 1991). Studies in plant community ecology, human geography, and landscape ecology have shown that the results of spatial analysis using area-based data usually are sensitive to three kinds of related, but distinctive changes in spatial data: changes in grain size, extent (Meentemeyer and Box, 1987; Woodcock and Strahler, 1987; Turner et al., 1989; Wickham and Ritters, 1995; Qi and Wu, 1996), and aggregation zones (the zoning problem; see Openshaw, 1984; Fotheringham and Rogerson, 1993; Wu and Jelinski, 1995; Jelinski and Wu, 1996). It has been suggested, therefore, that landscape pattern should best be understood by conducting analysis on multiple scales or hierarchically (Wu and Loucks, 1995; Wu and Jelinski, 1995; Jelinski and Wu, 1996; Qi and Wu, 1996).

As a part of a research project that attempts to link spatial pattern to ecosystem properties in the Great Basin, this study examined the effects of systematically changing spatial scale on the results of particular statistical analyses. Specifically, the objectives of this study were as follows: (1) to investigate how landscape metrics such as diversity and richness relate to spectral parameters readily available from remote sensing (e.g., TM band values) and vegetation indices derived from them; and (2) to examine the effects of varying sample sizes on the results of the analysis.

II. DATA AND METHODS

The data set for this study is a land-cover map derived from empirical information on topography, vegetation distribution, and land use conditions. The data set contains fourteen land-cover types, covering the Minden area of Nevada in the western Great Basin. The geographic coordinates for the four corners are 39°9′18.3″ N and 119°51′13.7″ W, 39°6′14.2″ N and 119°30′30.0″ W, 38°54′12.3″ N and 119°54′55.8″ W, and 38°51′8.2″ N and 119°34′12.1″ W, respectively. The data set has 999 rows and 1069 columns with a linear dimension of about 30 m for each pixel, which represents a total area of 96,114 hectares (or 961.14 square kilometers). The GIS package, IDRISI™, was used for Landsat image processing and a part of the pattern analysis, while S-Plus™ was used for ANOVA, regression, and correlation analysis.

From the land-cover map, we computed three landscape metrics, diversity (H), dominance (D) and richness (R), as descriptors of landscape structure. These metrics have been widely used in landscape ecological studies (e.g., O'Neill et al., 1988, 1996; Turner, 1989; Wickham and Riitter, 1995), and are defined as follows:

Landscape Diversity

$$H = -\sum_{k=1}^{m} P_k \ln P_k$$

where H is the diversity index, m is the number of land-cover types, P_k is the proportion of the grid cells of land-cover type k (the number of pixels of the land-cover type k divided by the total number of pixels). Larger values of H correspond to more diverse landscapes which tend to have many land-cover types with similar proportions of pixels belonging to each type.

Landscape Dominance

$$D = H_{\max} - \sum_{k=1}^{m} P_k \ln P_k$$

where D is the Dominance index, H_{max} is the maximum diversity when all land-cover types are present in equal proportions (i.e.). m and P_k are defined exactly the same as in the diversity index. This index is a measure of the extent to which one or a few land covers dominate the landscape. Small values usually correspond to landscapes with a large number of land use types of similar proportions. Apparently, a simple numerical relationship exists between diversity and dominance indices, both carrying the same non-spatial, compositional information of a landscape. While they were used together in our analysis for purposes of checking computational errors and facilitating interpretation, here we will focus primarily on the results on diversity to avoid redundancy.

Relative Richness

$$R = \frac{N}{N_{\max}} \cdot 100$$

where N is the number of different land-cover types present in an area under observation, and the N_{max} is the maximum value of richness.

Although the same basic formulas are used, in this study these metrics were calculated differently from the conventional way whereby they are computed for the entire study area or non-overlapping subregions. Because we were more interested in the characteristics of local-scale (or "neighborhood") diversity and their spatial changes, the landscape metrics were computed using a 3 by 3 pixel moving window as defined by the GIS package, IDRISI. For diversity and relative richness, respectively, a value for the metric was computed for the 9 neighboring cells, and then was assigned to the central cell. The window moves on one column at a time from the up left corner of the grid, until all the grid cells received their values. This is exactly the way these metrics are calculated using the PATTERN module of IDRISI (Eastman, 1995). As a result, the values of diversity and richness formed a 2-dimensional matrix and were represented as maps.

Three vegetation indices, RVI (Ratio Vegetation Index), NDVI (Normalized Difference Vegetation Index), and TNDVI (Transformed Normalized Difference Vegetation Index) were calculated from spectral information of the Landsat TM imagery of the study area. It was one of our objectives in this analysis to determine which of these vegetation indices would be best suited for detecting changes in the Great Basin landscapes. These indices were obtained from the following formula (Richardson and Wiegand, 1977; Tucker, 1979; Huete and Jackson, 1987):

$$RVI = \frac{Red}{NIR}$$

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

$$TNDVI = \sqrt{(NIR - Red)/(NIR + Red) + 0.5}$$

The Ratio Vegetation Index is simply the ratio of red to infrared brightness values and capitalizes on the increase in brightness as one moves from the red to the infrared data space. The Normalized Difference Vegetation Index is a more complex version of this simple ratio, and has been used in numerous vegetation assessment studies. Many studies have shown that NDVI is responsive to rapidly growing highly reflective plant communities such as alfalfa fields and riparian vegetation (Tueller, 1989; Rey-Benayas et al., 1995). The transformed normalized difference vegetation index, with the addition of 0.5, avoids negative values and usually is easier to interpret (Deering et al., 1975; Richardson and Wiegand, 1977; Harlan et al., 1979).

III. ANALYSIS AND RESULTS

In previous studies (Wu et al., 1994; Wu and Jelinski, 1995; Jelinski and Wu, 1996; Qi and Wu, 1996), we have shown that, for area-based data, varying the scale of analysis (grain size) and zoning systems (orientation and configuration) of the spatial units at the same scale both may have significant effects on the results of spatial analysis. This problem has been termed the modifiable areal unit problem (MAUP) in the geography literature (Openshaw, 1984; Fotheringham and Rogerson, 1993; Amrhein, 1995; Wu and Jelinski, 1995; Jelinski and Wu, 1996). In this study, we intended to explore how systematic (or progressive) changes of the analysis scale (specifically sample size) affect the results of regression and correlation analysis based on landscape data. How do different representation forms of variables — variance, mean, variance-mean ratio (V/M), and coefficient of variation — interact with the scale effects? Do scale effects show any trends that are predictable?

We used the three landscape metrics (diversity, and richness) as dependent variables and TM3, TM4, TM7, NDVI, TNDVI, and RVI as independent variables in the statistical analysis. To examine scale effects, four sample sizes were used: 25 pixels (5x5), 121 pixels (11x11), 225 pixels (15x15), and 361 pixels (19x19). First, we cut forty-nine 5x5 pixel samples from each of the 9 images (diversity, dominance, richness, TM3, TM4, TM7, NDVI, TNDVI, and RVI), and then symmetrically increased the scale of analysis, from the center cell outward, to 11x11, 15x15, and 19x19 pixels (Figs. 1 and 2). As a result, there were 49 replicates for each sample size. Variance, mean, variance-mean ratio (V/M), and coefficient of variation (CV =) of the nine variables at each sample size (n = 49) were computed, and then used accordingly for regression analysis, analysis of variance, and correlation analysis.

Regression analysis was conducted to examine how the landscape metrics relate to TM band parameters (TM3, TM4 and TM7) and vegetation indices (NDVI. TNDVI and RVI). Variance, mean, V/M, and CV of each variable are used for each sample size, respectively. For example, at the sample size of 5 by 5 pixels, four multiple linear regression models were constructed for each of the three dependent variables (diversity, dominance, richness) in terms of their variance, mean, V/M, and CV, respectively. The analysis of variance was used to determine the relative importance of the TM band parameters and vegetation indices in the relationship. We also performed a correlation analysis to further explore the relationship between landscape metrics and TM variables. In both ANOVA and correlation analysis, only the variance of dependent and independent variables at each sample size was used as the representation form because the regression analysis had shown that variance was more sensitive to changes in the landscape metrics than mean, V/M and CV.

The results of regression analysis showed that, for the sample size of 5 by 5 pixels, there did not appear to be a linear relationship between the landscape metrics (i.e., diversity, dominance, richness) and the six independent variables (i.e., TM3, TM4, TM7, NDVI, TNDVI, and RVI). This was true for all representation forms of the variables (i.e., mean, variance, V/M, and CV). For the sample size of 11 by 11 (121 pixels), a statistically significant linear relationship was apparent between the landscape metrics and independent variables when mean, variance, and V/M, but not CV, of these variables were used for the analysis (Table 1). When the sample size increased to 15x15 and 19x19 pixels, the linear relationship of the landscape metrics with TM bands and vegetation indices became statistically significant for all four forms of measure for the variables, with progressively larger R^2 values and smaller P values (see Table 1, Figs. 3 and 4). In general, the strength of this relationship tended to increase as

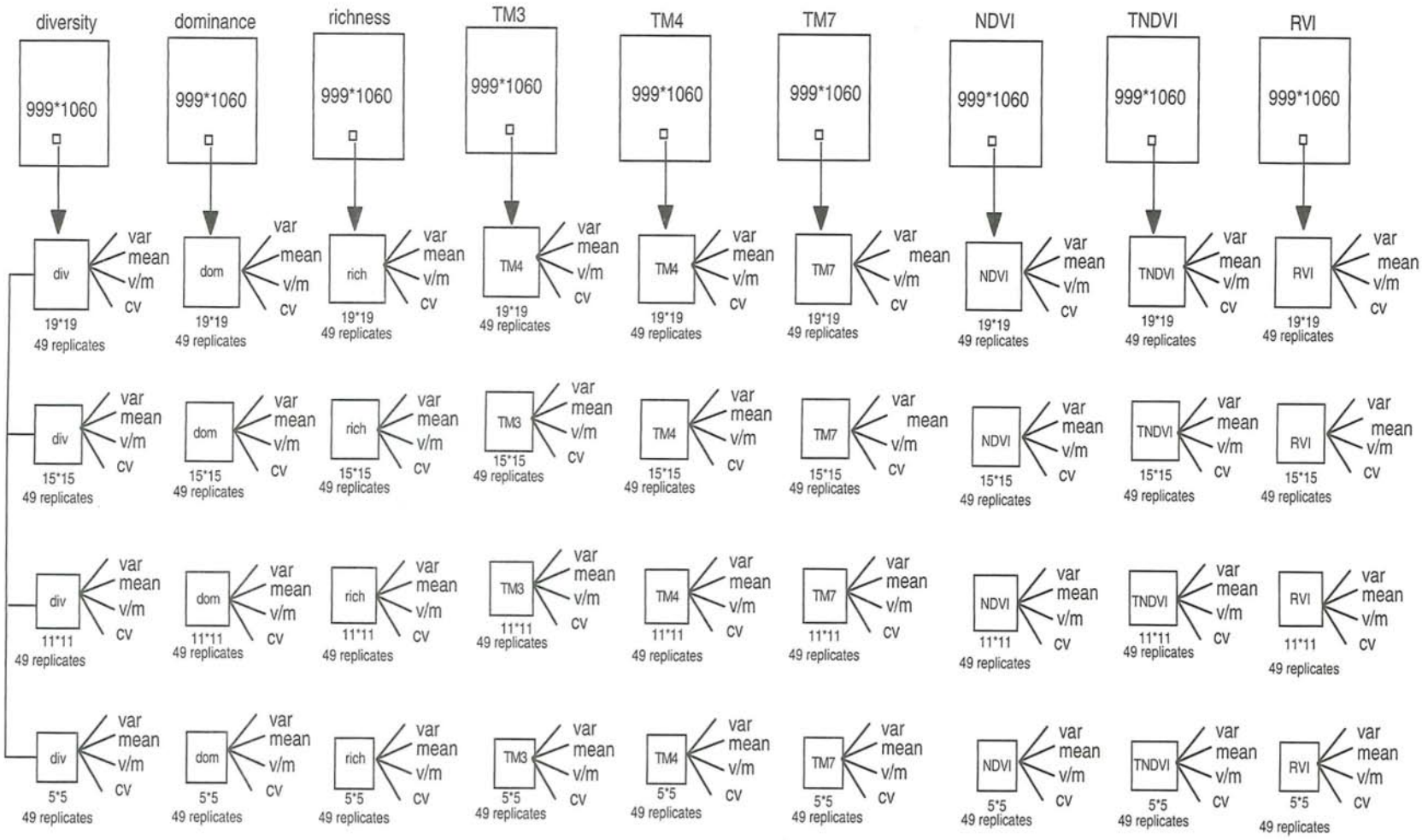


Figure 1. Schematic representation of the analysis design: 3 dependent variables (diversity, dominance, and richness); 6 independent variables (TM3, TM4, TM7, NDVI, TNDVI, and RVI); 4 different representation forms of variables (mean, variance, variance/mean ratio, and coefficient of variation); and 4 different sample sizes (5x5, 11x11, 15x15, 19x19 pixels, respectively).

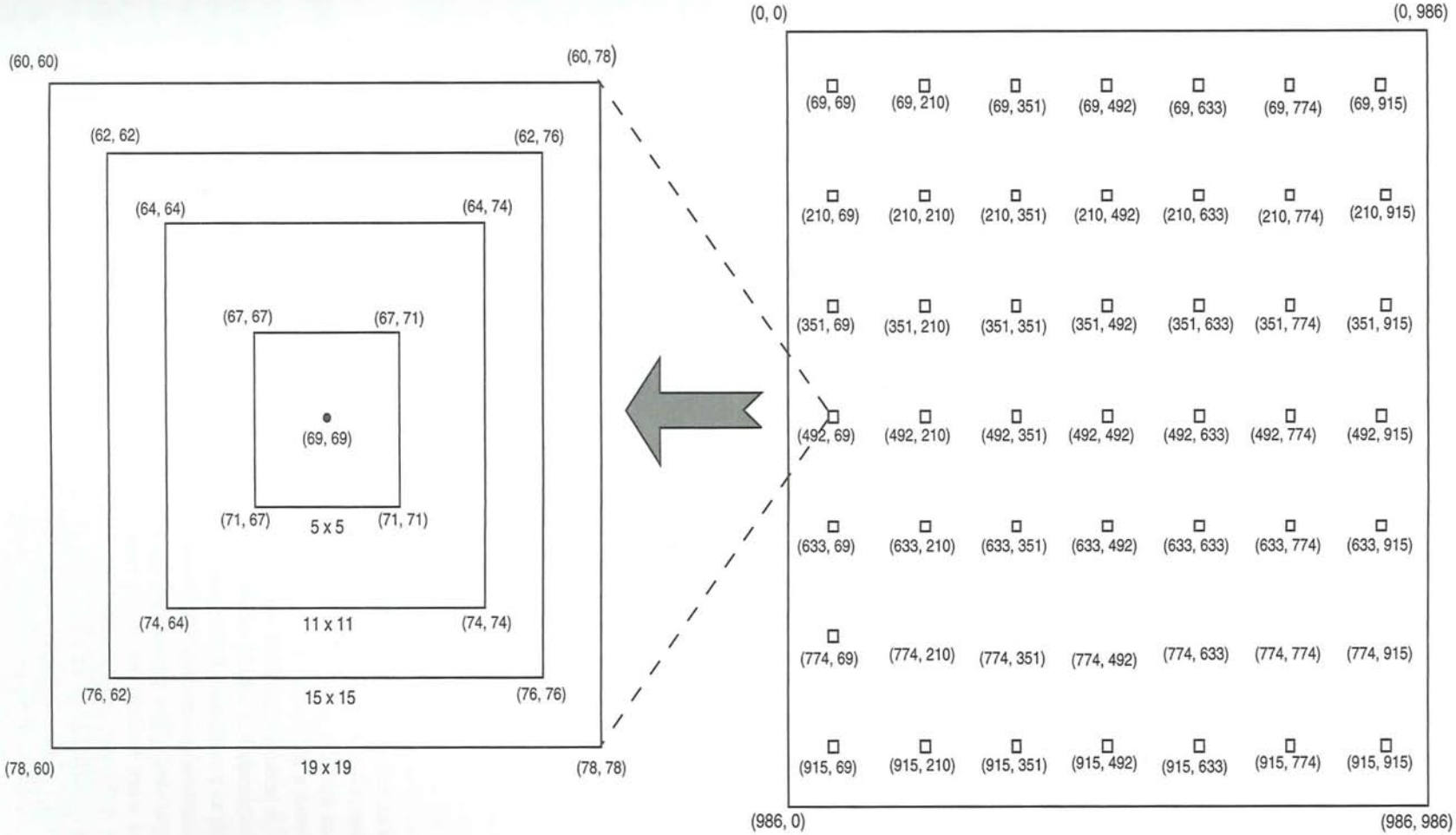


Figure 2. Illustration of the layout of samples of four different sizes: 5x5, 11x11, 15x15, and 19x19 pixels. The numbers in the parentheses in (B) are the coordinates of the center cells in each sample.

Table 1. Results of linear regression between the landscape metrics (diversity, dominance, richness) and TM3, TM4, TM7, NDVI, TNDVI, and RVI at 4 different sample sizes (5x5, 11x11, 15x15, and 19x19 pixels). Variance, mean, V/M and CV of the nine variables at each sample size are used separately in the analysis.

Measure	Landscape Indies	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value
		5*5(25 P)		11*11(121 P)		15*15(225 P)		19*19(361 P)	
Variance	Diversity	0.1637	0.3964	0.3943	0.0015**	0.4459	0.0003**	0.7679	0.0000**
	Dominance	0.2191	0.1950	0.311	0.0138*	0.4240	0.0006**	0.9999	0.0000**
	Richness	0.1192	0.6183	0.3179	0.0117*	0.5500	0.0000**	0.7049	0.0000**
Mean	Diversity	0.256	0.1117	0.3778	0.0024**	0.4553	0.0002**	0.3883	0.0018**
	Domanence	0.1473	0.4735	0.3494	0.0052**	0.4872	0.0001**	0.4232	0.0006**
	Richness	0.2471	0.1286	0.3449	0.0059**	0.4361	0.0004**	0.4022	0.0012**
V/M	Diversity	0.2505	0.1218	0.2531	0.0492*	0.4166	0.0163*	0.3889	0.0174*
	Domanence	0.2293	0.1682	0.3231	0.0103*	0.5485	0.0000**	0.3181	0.0116*
	Richness	0.2315	0.1534	0.2933	0.0211*	0.5731	0.0000**	0.4207	0.0007**
CV	Diversity	0.1278	0.5725	0.1568	0.2919	0.2214	0.0966	0.5232	0.0001**
	Domanence	0.1138	0.6469	0.1808	0.1994	0.2180	0.1031	0.2885	0.0236*
	Richness	0.1359	0.5305	0.1448	0.3478	0.2511	0.0534	0.2514	0.0531

* $P \leq 0.05$, ** $P \leq 0.01$

sample size increased for all four forms of measure (Fig. 3). However, a closer look reveals that R^2 values actually peaked at the sample size of 15x15 pixels in the cases of mean and V/M (Fig. 4).

The results of analysis of variance showed that, when variance was used as the representation form for the variables, the independent variables differed in terms of the level of significance in the relationship with landscape metrics as sample size increased (Table 2). For all the three landscape metrics, all independent variables were found insignificant at the sample size of 5x5 pixels. TM3 was statistically significant in the relationship for all the three landscape metrics at sample sizes of 11x11 pixels and larger, NDVI was significant for sample sizes of 15x15 and 19x19 pixels, and TM7 was only significant for the sample size of 19x19 pixels. The number of the spectral variables that were significant in the regression relationship increased as the sample size expanded. The results of the analysis of variance also were indicative of the relative importance of the different independent variables in the regression relationship at each sample size. Although a certain variable might be important at several sample sizes, its P value tended to decrease with the sample size (Table 2).

The results of correlation analyses, using variance as the representation form of all variables, showed that TM7 was significantly correlated with all the three landscape metrics at all four sample sizes, whereas TM 3 and TM4 were significantly correlated with these metrics when sample size was bigger than 5x5 pixels (Table 3). For all the three TM bands, R^2 values increased and P decreased as sample size expanded, indicating that the correlation between the landscape metrics and the TM bands became more significant with increasing sample size.

IV. DISCUSSION AND CONCLUSIONS

The results of our study have shown that the spatial pattern of local-scale or neighborhood diversity and richness in the Minden landscape could be characterized using TM spectral data. But sample size or the scale of analysis played an important role in relating the landscape metrics to TM spectral variables. With explicit specification of this scale effect, it seems feasible to use TM spectral information or vegetation indices to quantify and monitor spatial changes in the Great Basin landscape. However, several points are worth further discussion.

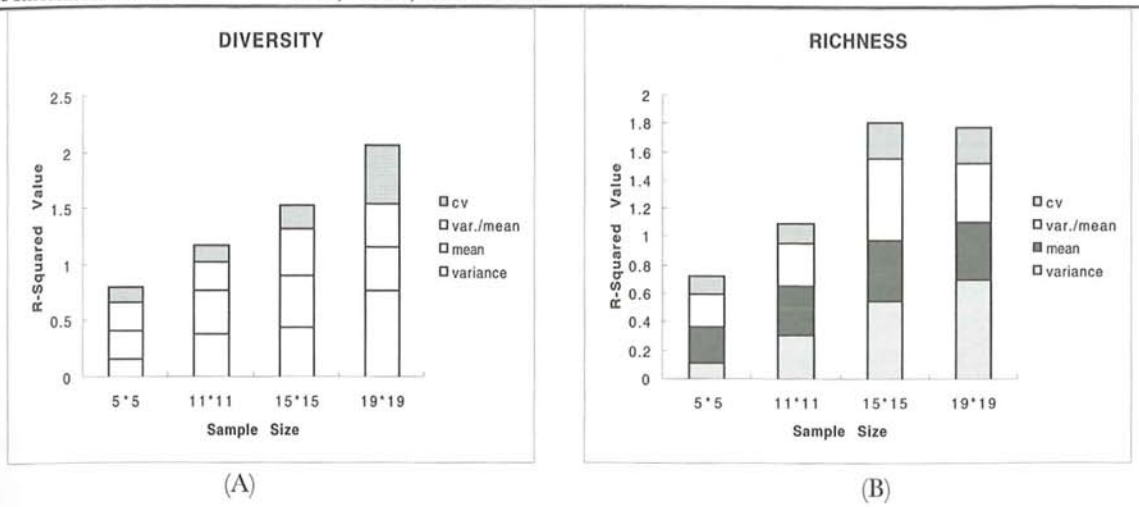


Figure 3. Accumulative R^2 values for the multiple linear regression between landscape metrics and spectral variables as a function of increasing sample sizes (5x5, 11x11, 15x15, and 19x19 pixels). Dependent variables are diversity (A), dominance (not shown here), and richness (B), and independent variables are TM3, TM4, TM7, NDVI, TNDVI, and RVI. Variance, mean, V/M and CV of the nine variables at each

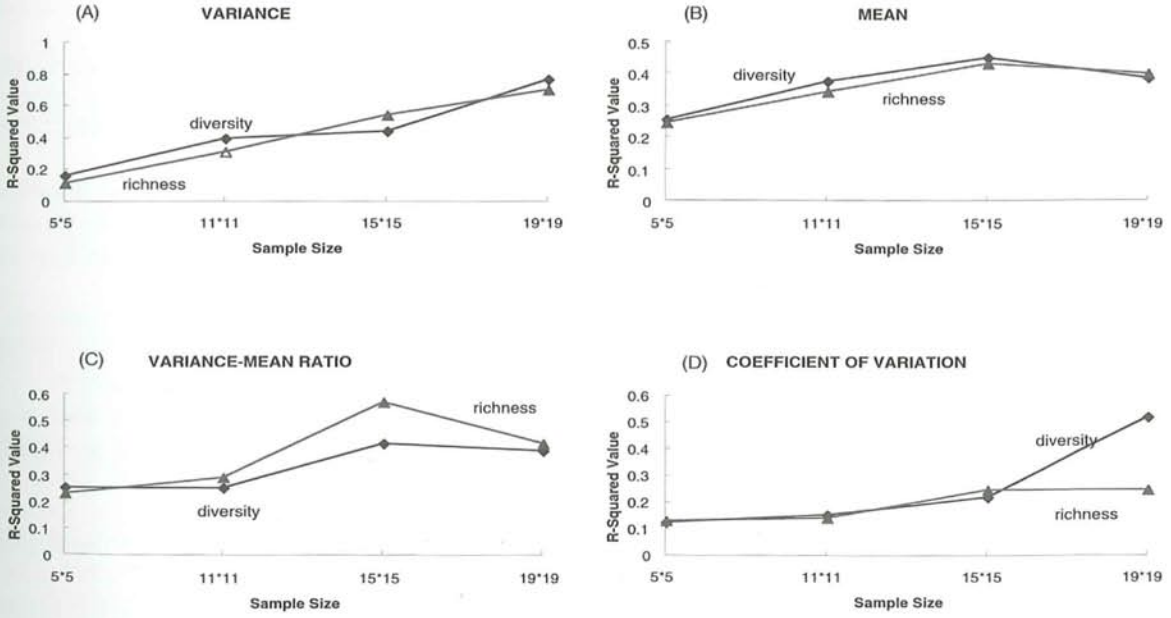


Figure 4. R^2 values for the multiple linear regression between landscape metrics and spectral variables as a function of increasing sample sizes (5x5, 11x11, 15x15, and 19x19 pixels). Dependent variables are diversity, dominance (not shown here), and richness, whereas independent variables are TM3, TM4, TM7, NDVI, TNDVI, and RVI. Variance (A), mean (B), V/M (C) and CV (D) of the nine variables at each sample size are used separately in the analysis. Also refer to Table 1 for numerical values.

Scale effects

Several studies have shown that changing scale may significantly affect the pattern quantification of an entire landscape or its subregions using, for example, richness and information theory-based metrics (Turner et al., 1989; Wickham and Ritters, 1995; O'Neill et al., 1996) and spatial autocorrelation indices (Legendre and Fortin, 1989; Jelinski and Wu, 1996; Qi and Wu, 1996). Specifically, the scale be-

Table 2. Results of analysis of variance between the landscape metrics (diversity, dominance, richness) and TM3, TM4, TM7, NDVI, TNDVI, and RVI at 4 different sample sizes. The variance value of each variable at each sample size is used in the analysis.

Sample size	Diversity vs.	P value	Dominance vs.	P value	Richness vs.	P value
5X5	TM3	0.08091	TM3	0.42235	TM3	0.17482
	TM4	0.59403	TM4	0.61092	TM4	0.76886
	TM7	0.41344	TM7	0.74131	TM7	0.57186
	NDVI	0.94461	NDVI	0.36995	NDVI	0.80863
	RVI	0.15879	RVI	0.05845	RVI	0.23338
	TNDVI	0.69449	TNDVI	0.06810	TNDVI	0.44635
11X11	TM3	0.00015**	TM3	0.01066*	TM3	0.00077**
	TM4	0.39013	TM4	0.09552	TM4	0.72986
	TM7	0.64337	TM7	0.68709	TM7	0.18886
	NDVI	0.03771*	NDVI	0.02759*	NDVI	0.16830
	RVI	0.06890	RVI	0.06370	RVI	0.15953
	TNDVI	0.71911	TNDVI	0.38048	TNDVI	0.97257
15X15	TM3	0.00002**	TM3	0.00249**	TM3	0.00000**
	TM4	0.55443	TM4	0.21824	TM4	0.79383
	TM7	0.20461	TM7	0.08522	TM7	0.11359
	NDVI	0.00206**	NDVI	0.01068*	NDVI	0.01075*
	RVI	0.04888*	RVI	0.01108*	RVI	0.01462*
	TNDVI	0.84668	TNDVI	0.16059	TNDVI	0.95484
19X19	TM3	0.00000**	TM3	0.00000**	TM3	0.00000**
	TM4	0.63878	TM4	0.00000**	TM4	0.97097
	TM7	0.00006**	TM7	0.00000**	TM7	0.00025**
	NDVI	0.00000**	NDVI	0.00000**	NDVI	0.00000**
	RVI	0.42321	RVI	0.31253	RVI	0.27869
	TNDVI	0.94307	TNDVI	0.56559	TNDVI	0.50972

* $P \leq 0.05$, ** $P \leq 0.01$

ing changed in our study is sample size, or may be regarded as extent with 49 replicates (see Fig. 2). Our study further has suggested that statistical analyses like regression, ANOVA, and correlation analysis with landscape data are also affected by changing scale. The effect of changing sample size on these analyses can be considerably large (Fig. 4). Of particular interest was that R^2 values increased monotonically in the variance and CV graphs (A and D in Fig. 4), whereas a peak became apparent at the 15x15 sample size in both mean and V/M graphs (B and C in Fig. 4). Further studies are needed to confirm whether this peak was indicative of a characteristic scale at which a real structural change in the landscape takes place. Because of scale effects, ecological conclusions based on such analyses should be made with explicit specification of scales (grain size and extent). Our results seem to suggest that this effect may be predictable within a certain domain of scales in some cases (see Fig. 4 for regions that correspond to nearly linear change in R^2 values).

Effects of different representation forms of variables

Scale effects were further complicated by the effect of different representation forms of variables used for the landscape analysis. For example, the four representation forms (variance, mean, V/M, and CV) for the 9 variables in this study resulted in somewhat distinctive patterns of change in R^2 values with increasing sample size (Fig. 4). For example, while diversity and richness seemed to exhibit similar patterns for each representation form at finer scales, variance was most sensitive to changes in diversity and richness pattern. The higher sensitivity of variance to change in the analysis scale is attributable,

Table 3. Results of correlation analysis between the landscape metrics (diversity, dominance, richness) and TM3, TM4, and TM7 at 4 different sample sizes. The variance value of each variable at each sample size is used in the analysis.

Sample size	TM (Variance)	Diversity		Dominance		Richness	
		R^2	P	R^2	P	R^2	P
5x5 pixels	TM3	0.2866	0.0365*	-0.125	0.7789	0.2266	0.0799
	TM4	0.0772	0.3180	0.0018	0.4955	0.0775	0.3174
	TM7	0.3205	0.0219*	-0.096	0.7221	0.2473	0.0019*
11x11 pixels	TM3	0.5082	0.0001**	0.3620	0.0154*	0.4769	0.0004**
	TM4	0.3602	0.0060**	0.3537	0.0069**	0.2994	0.0238*
	TM7	0.4849	0.0004**	0.3223	0.0127*	0.3512	0.0094**
15x15 pixels	TM3	0.5697	0.0041**	0.3820	0.0248*	0.6361	0.0013**
	TM4	0.3664	0.0052**	0.3312	0.0107*	0.3672	0.0051**
	TM7	0.5478	0.0000**	0.4496	0.0007**	0.6043	0.0000**
19x19 pixels	TM3	0.6020	0.0000**	0.4048	0.0050**	0.6677	0.0000**
	TM4	0.3680	0.0020**	0.3374	0.0095**	0.5114	0.0015**
	TM7	0.6728	0.0000**	0.4204	0.0001**	0.7118	0.0000

* $P \leq 0.05$, ** $P \leq 0.01$

at least in part, to the fact that its values are larger than those of V/M or CV in which variance is "scaled down" by mean.

Relationship between TM bands/derived vegetation indices and spatial pattern of land-cover richness and diversity

The results of regression analysis indicated that neighborhood diversity and richness were significantly correlated to TM band parameters and vegetation indices. The strength of the correlation seemed to increase with sample size (or calculation scale). This was evidenced by the increasing R^2 values and decreasing P values for the regression relationship, as well as by ANOVA and correlation analysis. In particular, the results suggested that the selected TM bands and vegetation indices could detect and predict changes in local-scale diversity and richness at sample sizes from 11x11 to 19x19 pixels with increasing accuracy. Clearly, use of variance as the representation form of variables at the 19x19 sample size gave the best result (R^2 larger than 0.7 for all three metrics; see Table 1 and Fig. 4). The results of both ANOVA and correlation analysis further suggested that TM3 and NDVI were the most consistent and best predictor variables.

TM3 band has been shown to be a good indicator of green vegetation (e.g., Tucker, 1979; Baret and Guyot, 1991). Rey-Benayas and Pope (1995) indicated that TM spectral data have the potential of measuring landscape diversity. While our results seem to support this claim, the choice of appropriate sample size will be critically important to achieve high accuracy. On the other hand, vegetation indices derived from several bands using different mathematical formulations may indicate quantitative and qualitative differences in the properties of vegetation because significant differences in reflectance and absorption of radiation exist between vegetation and other geographical characteristics of the landscape (Tueller, 1989). According to our analysis, normalized difference vegetation index (NDVI) appeared to be better than RVI and TNDVI for characterizing local-scale diversity and richness pattern in this particular desert landscape (Table 2). Numerous studies have shown that NDVI is a sensitive indicator of green biomass (Tucker, 1979, Tueller, 1989). Out study suggested that, together with TM3 and TM7, NDVI was a good predictor of diversity and richness in the landscape of our study. However, it is worth emphasizing again that the accuracy of these variables as predictors of land-cover diversity and richness not only depends on landscape composition, but also on sample size.

In conclusion, we emphasize that scale effects represent an important and challenging issue that must be considered explicitly in all landscape analysis.

Based on this and previous studies it seems unlikely to find "universal" rules that can be used to accurately predict scale effects over a wide range of scales or across different types of analysis and landscapes. This is in part because scale effects are further complicated by the choice of variables and the idiosyncrasy of particular landscapes. Yet, as this study suggests, responses of the statistical relationship to changes in analysis scale may exhibit simple (e.g., linear or monotonic) patterns over some ranges of scale, implying that scale effects could be readily predicted within these domains of scale. To find scale domains where predictions or extrapolations can be readily made, multiple-scale or hierarchical analysis must be performed. This study further supports that the modifiable areal unit problem is common across the disciplinary boundaries of geography, ecology and other earth sciences. Unraveling the problem will not only improve our understanding of pattern and process in nature, but also will have important implications for appropriate use of remote sensing data and GIS.

ACKNOWLEDGMENTS

This research was supported by research grants from the United States Department of Agriculture (USDA-NRICGP 95-37101-2028) and Arizona State University (FGIA HBR H044 and SRCA HB15001). The assistance with data collection and analysis by Ellen Ellis, Mingxi Jiang, and Michael Limb is gratefully acknowledged. We also thank Gong Peng, Ye Qi, and an annonymous reviewer for their comments on the manuscript.

REFERENCES

[1] Amrhein, C. G., 1995. Searching for the elusive aggregation effect: evidence from statistical simulations. *Environment and Planning A*, 27:105-119.

[2] Baret, F., and G. Guyot, 1991. Potential and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, 35:161-173.

[3] Cullinan, V. I., and J. M. Thomas, 1992. A comparison of quantitative methods for examining landscape pattern and scale. *Landscape Ecology*, 7(3):211-227.

[4] Deering, D. W., J. W. Rouse, R. H. Haas, and J. A. Schell, 1975. Measuring "forage production" on grazing units from LANDSAT MSS data. *Proceedings of the 10th International Symposium on Remote Sensing of Environment, Volume II*:1169-1178.

[5] Eastman, J. R., 1995. *Idrisi For Windows, User's Guide Version 1.0*. Idrisi Production 1987-1995, Clark University.

[6] Fotheringham, A. S., and P. A. Rogerson, 1993. GIS and spatial analytical problems. *International Journal of Geographical Information Systems*, 7:3-19.

[7] Harlan, J. C., D. W. Deering, R. H. Haas, and W. E. Boyd, 1979. Determination of range biomass using LANDSAT. *Proceedings of 13th International Symposium on Remote Sensing of Environment, Volume I*:659-673.

[8] Huete, A. R., and R .D. Jackson, 1987. Suitability of spectral indices for evaluating vegetation characters on arid rangelands. *Remote Sensing of Environment*, 23:213-232.

[9] Hunsaker, C. T., R. V. O'Neill, B. L. Jackson, S. P. Timmins, and D. A. Levine, 1994. Sampling to characterize landscape pattern. *Landscape Ecology*, 9:207-226.

[10] Iverson, L. R., R. L. Graham, and E. A. Cook, 1989. Applications of satellite remote sensing to forest ecosystems. *Landscape Ecology*, 3: 131-143.

[11] Jelinski, D. E., and J. Wu, 1996. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, 11:129-140.

[12] Legendre, P., and M.-J. F. Fortin, 1989. Spatial pattern and ecological analysis. *Vegetation*, 80:107-138.

[13] Meentemeyer, V. and E. O. Box, 1987. Scale effects in landscape studies. In *Landscape Heterogeneity and Disturbance*, Edited by M. G. Turner, Springer-Verlag, New York, pp. 15-34.

[14] O'Neill, R. V., R. H. Gardner, B. T. Milne, M. G. Turner, and B. Jackson, 1991. Heterogeneity and Spatial Hierarchies. In *Ecological Heterogeneity*, Edited by J. Kolasa and S. T. A. Pickett, Springer-Verlag, New York, pp. 85-96.

[15] O'Neill, R. V., B. T. Milne, M. G. Turner, and R. H. Gardner, 1988. Resource utilization scales and landscape pattern. *Landscape Ecology*, 2:63-69.

[16] O'Neill, R. V., C. T. Hunsaker, S. P. Timmins, B. L. Timmins, K. B. Jackson, K. B. Jones, K. H. Riitters, and J. D. Wickham, 1996. Scale problems in reporting landscape pattern at the regional scale. *Landscape Ecology*, 11:169-180.

[17] Openshaw, S., 1984. The modifiable areal unit problem. *CATMOG 38*. GeoBooks, Norwich.

[18] Plotnic, R. E., R. H. Gardner, and R. V. O'Neill, 1993. Lacunarity indices as measures of landscape texture. *Landscape Ecology*, 8:201-211.

[19] Pickett, S. T. A., and M. L. Cadenasso, 1995. Landscape ecology: spatial heterogeneity in ecological systems. *Science*, 269:331-334.

[20] Qi, Y. and J. Wu, 1996. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. *Landscape Ecology*, 11:39-50.

[21] Rey-Benayas, M. Jose, and K. O. Pope, 1995. Landscape ecology and diversity patterns in he seasonal tropics form Landsat TM imagery. *Ecological Applications*, 5:386-394.

[22] Richardson, A. J., and C. L. Wiegand, 1977. Distinguishing vegetation from soil background information. *Photogrammetric Engineering & Remote Sensing*, 43:1541-1552.

[23] Riitters, K. H., R. V. O'Neill, C. T. Hunsaker, J. D. Wickham, D. H. Yankee, K. B. Timmins, and B. L. Jackson, 1995. A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*, 10:23-39.

[24] Risser, P. G., 1990. Landscape Pattern and Its Effects on Energy and Nutrient Distribution. In *Changing Landscapes: An Ecological Perspective,* Edited by I. S. Zonneveld and R. T. T. Forman, Springer-Verlag, New York, pp. 45-56.

[25] Roughgarden, J., S. W. Running, and P. A. Matson, 1991. What does remote sensing do for ecology? *Ecology*, 72: 1918-1922.

[26] Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8:127-150.

[27] Tueller, P. T., 1989. Remote sensing technology for rangeland management applications. Journal of Range Management, 42:442-452.

[28] Turner, M. G., 1989. Landscape ecology: The effect pattern on process. *Annual Review of Ecology and Systematics*, 20:171-197.

[29] Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne, 1989. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, 3:153-162.

[30] Turner, M. G. and R. H. Gardner, 1991. *Quantitative Methods in Landscape ecology*. Springer-Verlag, New York.

[31] Wickham, J. D, and K. H. Riitters, 1995. Sensitivity of landscape metrics to pixel size. *International Journal of Remote Sensing,* 16:3585-3595.

[32] Wiens, J. A., N. C. Stenseth, B. V. Horne, and R. A. Ims, 1993. Ecological mechanisms and landscape ecology. *Oikos*, 66:369-380.

[33] Woodcock, C. E., and A. H. Strahler, 1987. The factor of scale in remote sensing. *Remote Sensing of Environment*, 21:311-332.

[34] Wu, J., and D. E. Jelinski., 1995. Pattern and scale in ecology: The modifiable areal unit problem. In *Lectures in Modern Ecology*, Edited by Li Bo, Science Press, Beijing, pp.1-9. (In Chinese)

[35] Wu, J., and S. A. Levin, 1994. A spatial patch dynamic modeling approach to pattern and process in an annual grassland. *Ecological Monographs*, 64(4): 447-464.

[36] Wu, J., and S. A. Levin, 1997. A patch-based spatial modeling approach: conceptual framework and simulation scheme. *Ecological Modelling*, 101:325-346.

[37] Wu, J., and O. L. Loucks, 1995. From balance-of-nature to hierarchical patch dynamics: A paradigm shift in ecology. *Quarterly Review of Biology*, 70:439-466.

[38] Wu, J., J. L. Vankat, and B. Barlas, 1993. Effects of patch connectivity and arrangement on animal metapopulation dynamics: a simulation study. *Ecological Modelling*, 65:221-254.

Restructuring the SQL Framework for Spatial Queries

Bo Huang and Hui Lin

Department of Geography & Joint Laboratory for Geoinformation Science
The Chinese University of Hong Kong
Shatin, NT, Hong Kong

Abstract

This paper presents an approach to designing a spatial query language, called GeoSQL, in terms of the conventional spatial query and implementation process. A critical factor to the design is how to accommodate spatial operators in an appropriate form, while being compatible with the Structured Query Language (SQL) standard. To achieve this, the FROM clause of SQL is restructured to contain spatial operators via a subquery so that the results of spatial operations can be easily fed into both the SELECT and WHERE clauses. The subquery in the FROM clause creates an intermediate relation, on which the selection in terms of certain criteria is conducted. This is a distinct characteristic of GeoSQL. The syntax and semantics of GeoSQL are described, and a set of examples for testing the expressiveness of the language is given. The interface of the language is also designed with the introduction of visual constructs (e.g., icons and ListBoxes) to aid the entry of query text. This distinguishes GeoSQL's interface from the previous extended SQLs, which only employ pure text for constructing a query. After this, an implementation of GeoSQL is discussed. This paper finally suggests further extending GeoSQL for temporal and fuzzy queries.

摘 要

本文根据常规的空间分析过程，设计了空间查询语言：GeoSQL。GeoSQL采用了子查询(subquery)，重组了FROM语句，这是GeoSQL的主要特色之一。其次，GeoSQL的界面设计融入了Icon，ListBox等可视化部件，使查询文本输入变得容易，同时减少了语法错误，这是GeoSQL的另一特色。本文介绍了GeoSQL的表达形式、界面设计和实现方法。

I. INTRODUCTION

The need for a formal spatial query language has been widely identified in GIS community [8, 10]. Therefore, several approaches to devising a spatial query language have been developed [2, 4]. Of them, extending the relational database languages, primarily SQL, is a major one.

SQL, very suitable for the retrieval of lexical data, has been the standard query language for relational databases [1]. However, based on the underlying power of relational algebra, SQL has proved to be insufficient for the queries involving spatial proprieties such as metric and topology [7]. Hence a variety of extended SQLs were addressed (e.g., [8, 12, 14, 16, 21]). For GIS requirements, the main extensions to SQL include the introduction of spatial data types such as point, line and polygon, as well as spatial operators such as distance, direction, intersection and buffer. Given that spatial operators in these languages are applied in either the SELECT or WHERE clause, it becomes difficult to apply the results of spatial operators occurring in one clause (e.g., WHERE clause) to another clause (e.g., SELECT clause), and to implement further conditions on these operators such as temporality.

Different from the above approach, Gadia [11] proposed a spatial SQL in the form of "SELECT ... RESTRICTED TO · FROM ... WHERE". The condition in the WHERE clause only includes non-spatial attributes, while the condition in the augmented RESTRICTED TO clause deals with the spatial data. Huang [15] designed an extended SQL by incorporating spatial operators in the FROM clause, while the other clauses remain intact. The modified FROM clause creates a new relation with derived attributes representing the results of spatial operations. These attributes, just as other attributes in the source relations, can then be applied in both the SELECT and WHERE clauses. The direct incorporation of spatial operators, however, does not comply with the general representation of the FROM

1082-4006/97/0301~2-42$3.00

clause, as has been done by many variants of SQL.

In order to overcome this problem, this paper attempts to redesign the FROM clause via a subquery (i.e., a nested SQL), and in the meanwhile, incorporate more spatial operators including those for complicated spatial analyses such as INTERSECTION, UNION and DIFFERENCE. It should be clear that the ongoing SQL/MM [17] and SQL3 do not impose a unique form for representing spatial queries. The design of GeoSQL is conducted strictly within their framework.

The remainder of this paper is organized as follows. Section II describes the spatial data types and spatial operators in GeoSQL. The syntax and semantics of GeoSQL are discussed in Section III, with a stress on the representation of the FROM clause. Section IV gives several examples to illustrate how spatial queries are represented by GeoSQL. The interface of GeoSQL is presented in Section V, which introduces some visual constructs such as icons and ListBoxes, thereby increasing its user friendliness. Following this is an implementation of GeoSQL in Section VI. Finally, in Section VII, this paper concludes with some comments on the characteristics of GeoSQL and its future development.

II. SPATIAL DATA TYPES AND OPERATORS

Spatial Data Types

Generally, there are two kinds of spatial data models: feature-based and layer-based in GIS. The former one models spatial features while the latter one models map or a set of thematic maps [24]. The feature-based data model is currently adopted by many GIS packages such as ESRI's ArcView, MapInfo Inc.'s MapInfo and Intergraph's Modular Graphical Environment (MGE).

In a feature-based model, a spatial feature, e.g., a road, school or region, is represented as a geometric object with spatial attributes such as coordinates and topological relationships, as well as non-spatial attributes such as name, type and size. Usually, a class of features having a similar thematic property (e.g., roads, rivers or landuse) is represented by a spatial relation, and a feature corresponds to a tuple in the spatial relation. The spatial relation extends the conventional relation with an Abstract Data Type (ADT) , i.e., using GEO attribute for spatial representation. In other words, spatial attributes appear at the same conceptual level as the non-spatial attributes [22]. The basic relational operations such as projection and Cartesian product are considered applicable to spatial relations. The GEO attribute can be of point, line or polygon type.

Using the above method, the following schemas related to Hong Kong region are defined:

region (ID, name, population, GEO)
landuse (ID, type, GEO)
parcel (ID, address, GEO)
road (ID, name, class, GEO)
building (ID, name, owner, GEO)
university (ID, name, studentnum, GEO)

The features of region, landuse, parcel, building and university are of polygon type, while those of road are of line type. These tables are to be used along this paper.

Spatial Operators

Spatial operators are the methods pertaining to spatial features, which are employed to extract information from spatial features, as well as to create new spatial features from existing ones [23].

Several sets of spatial operators have been defined to query spatial database [2, 6]. Both SQL/MM and Spatial Database Engine (SDE) [9] have also defined their sets of spatial operators. Based on these, four groups of typical spatial operators are defined to illustrate how they are applied in GeoSQL.

(1) Unary spatial operators

The unary spatial operators are often used to obtain a scalar value, arcs or centroid of a spatial feature such as

ARCS(Pgn) gets the arcs from the polygon Pgn
AREA(Pgn) calculates the area of the polygon Pgn
LENGTH(L/Pgn) calculates the length of the line L or the polygon Pgn (perimeter)
CENTROID(Pgn) gets the center point of the polygon Pgn
VORONOI(Pnts) gets the VORONOI map of a pointset Pnts
BUFFER(SP) gets the buffer area of a spatial feature

Some of the above operators are type-specific, e.g., AREA, which can only take spatial features of poly-

gon type as its operands while line and point type of features are not applicable. But, some others are generic, e.g., BUFFER(SP), which can operate on one or more data types,. In this case, the spatial data type is defined as SP.

(2) Binary geometric operators

The following are the two main geometrical operators:

DISTANCE(SP1, SP2) calculates the minimum Euclidean distance between two spatial features.
DIRECTION(Pnt1, Pnt2) calculates the angle of the line connecting the points Pnt1 and Pnt2.

(3) Binary topological operators

Topological operators determine the topological relationship between two spatial features and return a Boolean value. If the topological relationship defined by an operator holds between its arguments, the operator returns the value TRUE; else FALSE.

According to [5, 6], the topological relationships usually include DISJOINT, CONTAINS, TOUCH, WITHIN, OVERLAP, CROSS, INTERSECTS and EQUALS.

(4) Binary construction operators

Construction operators may create new spatial features if a certain topological relationship holds between two spatial features. The main construction operators are:

UNION(SP1, SP2) gets all the primitive lines or polygons
INTERSECTION(SP1, SP2) gets the common part of two spatial features
DIFFERENCE(SP1, SP2) gets the different part of two spatial features

This group of operators represent the most difficult type of spatial operators to define directly in SQL [14]. However, like other types of operators, these operators in GeoSQL are represented in the way as other operators. This is discussed below.

III. REPRESENTATION OF GEOSQL

Conventional Spatial Query and Implementation Process Using GIS

When a query involves several spatial operations, we often first do spatial operations using unary spatial operators, binary spatial operators or both of them, and obtain an intermediate result. Then a selection in terms of certain criteria is carried out on this intermediate result. Finally, the desired attributes are projected. This procedure is shown in the left part of Figure 1, which provides a basis for the design of GeoSQL.

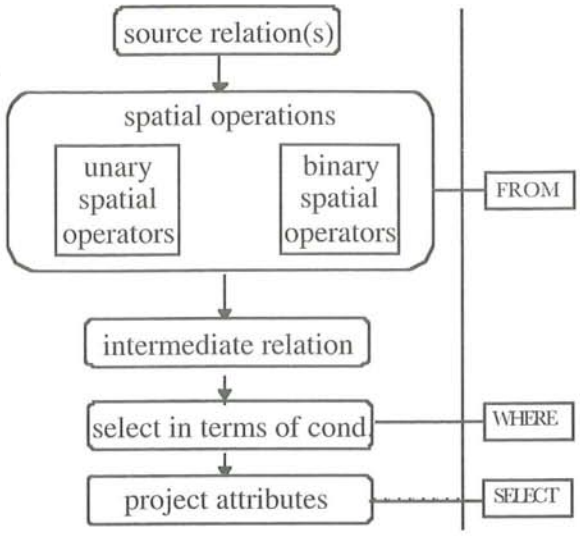


Figure 1. The conventional spatial query and implementation process using current GIS packages

Syntax of GeoSQL

Generally, a SQL statement is as follows:

SELECT $A_1, ..., A_m$
FROM $R_1, ..., R_n$
WHERE F

It is described in relational algebra as

$$\prod_{A_1,...,A_m}(\sigma_F(R_1 \times ... \times R_m))$$

Hence the FROM clause implies the Cartesian product of the given relations. In any case, the FROM clause needs to finally define a single relation, because it is the relation to which the selection in the WHERE clause and the projection in the SELECT clause are applied.

Enlightened from this, if the FROM clause is restructured to create an intermediate spatial relation resulted from spatial operations, the SQL statement can then be easily adapted to the above spatial query and implementation process (Figure 1).

An approach to implementing this is to append new attributes derived from spatial operations to the Cartesian product of source relations. Since the direct insertion of spatial operators into the FROM clause will not be consistent with the ongoing SQL3 standard, it is necessary to employ a subquery, i.e., embedding a SQL statement, in the FROM clause. The spatial operators are, therefore, applied in the nested SELECT clause for projection of their results, for example,

```
SELECT .....
FROM
    (SELECT *, CONTAINS(r. GEO, u. GEO) AS
          contval, AREA(urge) AS areaval
     FROM  region AS r, university AS u)
WHERE  .....
```

The result of CONTAINS operation is represented by the attribute "contval", and the result of AREA operation by the attribute "areaval". The subquery in the above FROM clause creates an intermediate relation, whose schema is shown in Table 1. The selection with certain conditions can then be carried out on this intermediate relation. In effect, such a relation is a virtual result because it can be optimized in terms of projection items and selection conditions during the implementation process.

This FROM clause can be described in geo-relational algebra [13] as

```
region university product
extend [contval: CONTAINS(region.GEO,
university.GEO), areaval:
          AREA(university.GEO)]
select [·]
project [·]
```

The "extend" operation above is to add new attributes to the Cartesian product of region and university.

The syntax of GeoSQL is based on the conventional SQL, whose basic form is

```
SELECT <select-clause>
FROM  <from-clause>
WHERE <where-clause>
```

Since only the FROM clause is different from that in the conventional SQL, its BNF form is described below:

```
<from-clause> :: = <relations> | <nested SQL>
<relations> :: = relation {, <relations>}
<nested SQL> :: = SELECT <sub-select clause>
    FROM <relations>
<sub-select clause> :: = *, <spatially derived at-
    tributes>
<spatially derived attributes> :: = <spatial op-
    erators> AS <attribute name> {, <spatially de-
    rived attributes> }
<spatial operators> :: = <unary spatial opera
    tors> | <binary geometric operators>
    |<binary topological operators> | <binary
    construction operators>
```

Conceptually, the subquery in the FROM clause of GeoSQL is just seen as an intermediate relation, which includes new derived attributes representing the spatial operation results. The derived attributes are taken as the same as those in the source relations, and thus can be easily applied as constraints in the WHERE clause, and referenced in the main SELECT clause for further statistical analysis or graphical display.

IV. QUERY EXAMPLES

A set of database schemas related to Hong Kong region has been defined in Section II. The following gives a group of examples to illustrate how spatial queries are represented by GeoSQL.

Example 1. Display the commercial landuse in Hong Kong.

```
SELECT  GEO
FROM    landuse
WHERE  type = 'commercial'
```

The GEO attribute occurring in the SELECT clause is to display the selected features in a map.

Example 2. Find the residences less than 2KM away from the Hong Kong Polytechnic University (HKPU), and show the distance.

Table 1. The virtual schema of the above FROM clause

Attributes of region (r)				Attributes of university(u)				New attributes	
r. ID	r. name	r. population	r. GEO	u. ID	u. name	u. studentnum	u. GEO	contval	areaval

```
SELECT  b.name, distval
FROM
   (SELECT *, DISTANCE(u.GEO, b.GEO) AS
         distval
    FROM university AS u, building ASb)
WHERE  u.name = 'HKPU' and b.type = 'resi
   dence' and  distval <= 2
```

The attribute "distval" derived in the FROM clause enables it to be applied in both the SELECT and WHERE clauses.

Example 3. Find the universities with more than 10,000 students, and indicate they are inside or outside the Kowloon region in Hong Kong.

```
SELECT  u.name, contval
FROM
   (SELECT *, CONTAINS(r.GEO, u.GEO) AS
         contval
    FROM region AS r, university AS u)
WHERE  r.name = 'Kowloon' and u.studentnum
         > 10,000
```

The derived attribute "contval" occurring in the SELECT clause shows a Boolean value indicating a selected university is inside or outside the specified region.

Example 4. Display the built-up area in Kowloon and New Territories, and sum the area.

```
SELECT  IGEO, sum(areaval)
FROM
   (SELECT *, INTERSECTION(rg.GEO,
         lu.GEO) AS IGEO, AREA(IGEO) AS
         areaval
    FROM  region AS rg, landuse AS lu)
WHERE  rg.name = 'Kowloon' or rg.name = 'New
         Territories' and lu.type = 'built-up'
```

Since landuse parcels and the regions may overlap, an intersection operation is required. Its result is specified by the spatial attribute "IGEO".

Example 5. Display the residences in Kowloon region that are nearer to HKPU than other universities.

```
SELECT  b.GEO
FROM
   (SELECT *, VORONOI(u.GEO) AS VGEO,
         INTERSECTION(VGEO, r.GEO) AS
         IGEO CONTAINS(IGEO, b.GEO) AS
         contval
    FROM  region AS r, building AS b, university
         AS u)
WHERE  r.name = 'Kowloon' and b.type = 'resi
         dence' and u.name = 'HKPU' and
         contval = TRUE
```

Voronoi operation, whose result is specified by the spatial attribute "VGEO", can meet the requirement of "nearer" function.

Example 6. Which are the roads crossing Kowloon region such that the total distance is between 30 KM and 50 KM?

```
SELECT  rd.name, rd.GEO, lval
FROM
   (SELECT *, INTERSECTION(rg.GEO,
         rd.GEO) AS IGEO,
         LENGTH(IGEO) AS lval
    FROM  region AS rg, road AS rd)
WHERE   rg.name = 'Kowloon'
GROUP BY rd.name
HAVING sum(lval) > 30 and sum(lval) < 50
```

This query shows that the result of spatial operation can also be applied to other clauses like HAVING clause besides the SELECT and WHERE clauses.

V. INTERFACE DESIGN

Although GeoSQL, an extended SQL, belongs to the textual language that is often considered incompatible with the visual language, the advantages of visual query languages such as intuitiveness and easiness [3, 19] can still be introduced in the interface design of GeoSQL to facilitate text input and reducing syntactic errors. In particular, with the development of windows programming, various visual constructs such as icons, ListBoxes and ComboBoxes can be employed in building such a user interface (Figure 2).

The interface is composed of five windows: (1) the text window, (2) the control window, (3) the ComboBoxes window, (4) the icons window, and (5)the settings window.

The text window is where the user enters the GeoSQL text following the SELECT ... FROM (SELECT - FROM) ... WHERE block. The control window contains four command buttons. The execute button passes the GeoSQL text to the implementation program. After execution, the map and attribute table are popped up to show the query result. The verify button checks the text with GeoSQL gram-

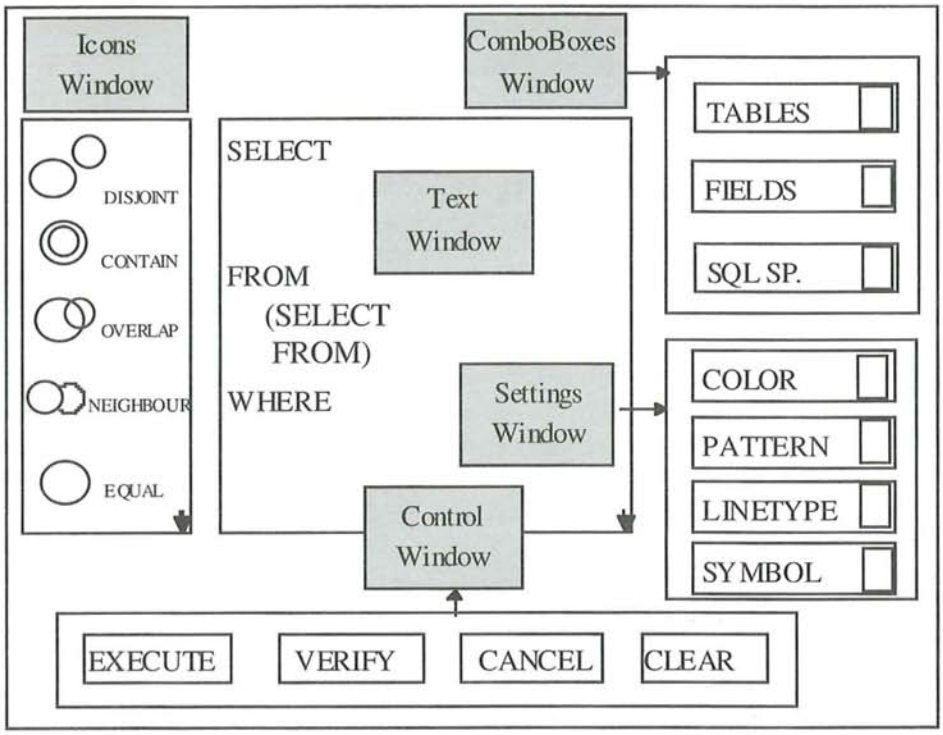


Figure 2. Window layout of GeoSQL

mar. If there is any mistake, a message is shown in a temporary pop-up window. The clear button clears the text window, and causes the already selected part of spatial objects to disappear in both the map and attribute display windows.

The ComboBoxes window contains three ComboBoxes. The "tables" stores both spatial and non-spatial tables. Once an item of the table ComboBox is selected, its corresponding fields will be added to the fields ComboBox. If two or more tables are selected, the field name will be in the form of table_name.attribute_name.

The SQL Special ComboBox lists all the predicates and operators in standard SQL. If any item in the above four ComboBoxes is selected, its corresponding text expression will occur in where the cursor locates in the text window so that the user input by keyboard is saved and the likely syntactical errors are reduced (Figure 3).
The icons window lists spatial operators in GeoSQL. The item in this ListBox differs from the ordinary ListBox in that it can combine an icon with its textual description. The ListBox is scrollable so that the user can select all the icons. It is apparent that the icons window and the ComboBoxes window are used together to assist the entry of GeoSQL text.

The settings window sets the graphical output of the query with colors, patterns, linetypes or symbols, which are realized by the four ComboBoxes respectively. The patterns for spatial data output are shown in Figure 3.

VI. AN IMPLEMENTATION

In our prototype system, GeoSQL is implemented using Oracle, Open Database Connectivity (ODBC) and Visual C++ (VC). The user interface is programmed in Visual C++ language (VC). The spatial data are stored in the Binary Large Object Block (BLOB) item in Oracle database, which are accessed through the ODBC embedded in VC programs.

After checking syntactic errors in the query text and optimization, the query processor including a spatial ODBC executes spatial operators and yields results. The spatial ODBC is composed of ODBC Application Program Interfaces (API) and APIs for implementing spatial operators. The whole implementation process corresponding to the "execute" command in the user interface is shown in Figure

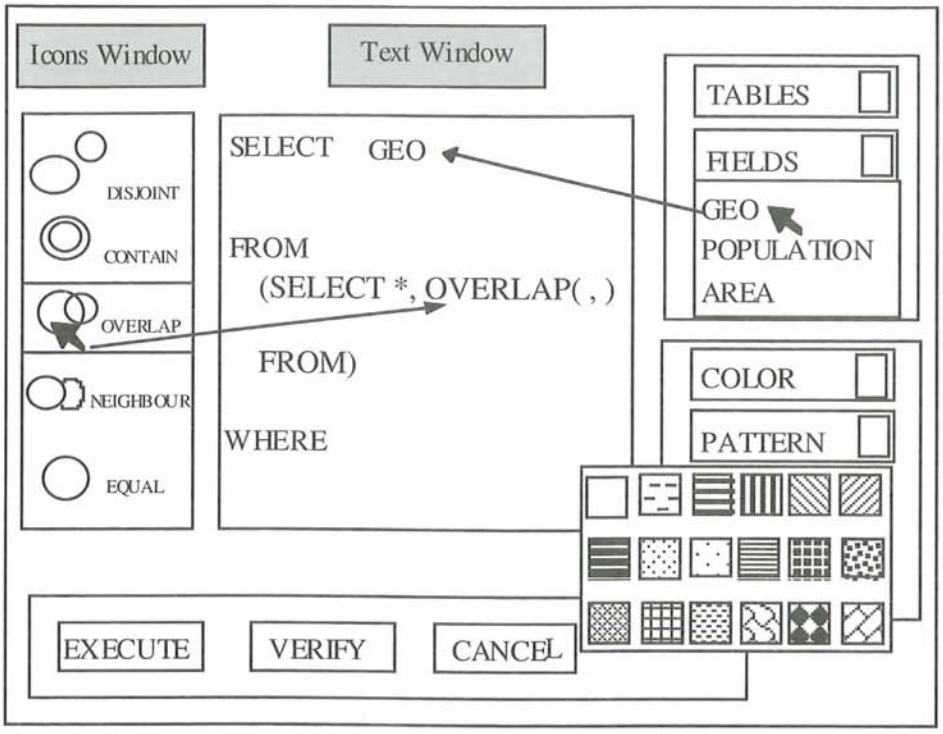


Figure 3. The assistance of visual constructs for GeoSQL text entry and the settings for graphical output

4. Using such a method, the example 4 in Section IV is implemented and the result is shown in Figure 5.

VII. CONCLUSIONS

This paper describes a different approach to designing an extended SQL in terms of the conventional spatial query and implementation process. By incorporating spatial operators in the FROM clause via a subquery, GeoSQL is well adapted to the conventional SQL design principles. More importantly, it becomes possible to apply the results of spatial operations in the SELECT, WHERE and HAVING clauses. Because the result of a spatial operation in GeoSQL is treated as a derived attribute, the further conditions such as temporality (e.g., valid-time) and fuzziness (e.g. "very far" and "overlap significantly") can be acted on it. In this sense, GeoSQL holds promise in expressing the spatio-temporal and fuzzy queries.

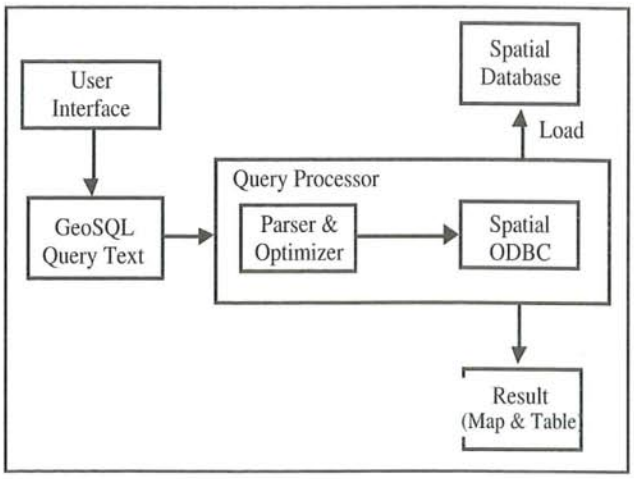


Figure 4. An implementation of GeoSQL

The interface design of GeoSQL introduces a set of visual constructs, which aid the entry of query text and reduce the possible syntactical errors. Such an approach overcomes the problems of previous extended SQLs, which compose a query only by pure text input.

The implementation of GeoSQL is a non-trivial task. Currently only part of the spatial operators can be implemented. Since SDE is now available in our laboratory, the spatial operators are to be realized

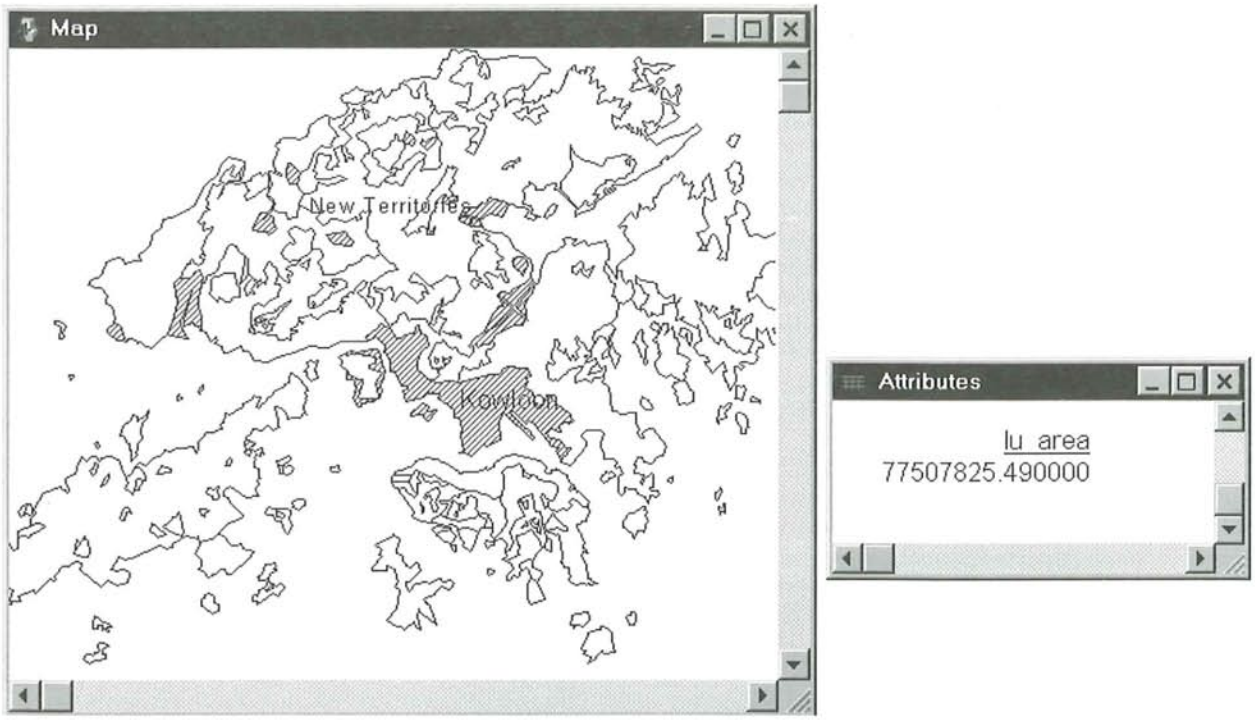


Figure 5. Experimental result of example 4 in Section IV

by it. The connection of SDE with our programs still needs to be done. In addition, more attention is also needed to the optimization of the language during its implementation.

ACKNOWLEDGMENTS

This research is partially supported by the Research Grants Council of HKSAR government under RGC earmarked research grant No. CUHK 150/96H, CUHK RAC under grant No. 4720401, and the National Natural Science Foundation (NNSF) of China under grant No. 49501013. We are grateful to Professors Guanhua XU and Shouyong Yan for their thoughtful suggestions at the early stage of this research.

REFERENCES

[1] American National Standards Institute (ANSI), 1989, *X3.135-1989 Database Language SQL*.

[2] Boursier, P. and M. Mainguenaud, 1992, Spatial query languages: extended SQL vs. visual language vs. hypermaps. *5th International Symposium on Spatial Data Handling*, Charleston, SC, USA.

[3] Calcinelli, D. and M. Mainguenaud, 1992, Cigales: a visual query language for Geographical Information System: the user interface. *Journal of Visual Languages and Computing*, 5, 113-132.

[4] Chu, T.H. and Y.T. Lung, 1995, Design a spatial language applying behavioral approach. In *Proceedings of GeoInformatics' 95*, Hong Kong, pp. 851-867.

[5] Clementini, E., D. Paolino and P. Oosterom, 1993, A small set of formal topological relationships suitable for end-user interaction. In D.J. Abel and B.C. Ooi (eds.), *Advances in Spatial Databases*, SSD'1993, Singapore, pp. 277-295.

[6] Egenhofer, M. and R. Franzosa, 1991, Point-set topological spatial relations. *International Journal of Geographical Information Systems*, 5(2): 161-174.

[7] Egenhofer, M., 1992, Why not SQL! *International Journal of Geographical Information Systems*, 6(2): 71-85.

[8] Egenhofer, M., 1994, Spatial SQL: a query and presentation language. *IEEE Transactions on Knowledge Engineering and Data Engineering*. 6(1): 86-95.

[9] ESRI, 1996, *SDE Developer's Guide Version 2.1*, Environmental Systems Research Institute, Inc.

[10] Frank, A., 1982, Mapquery - Database query languages for retrieval of geometric data and its graphical representation. *ACM Computer Graphics*, 16(3): 199-207.

[11] Gadia, S.K., 1993, Parametric databases: seamless integration of spatial, temporal, belief, and ordinary data. *SIGMOD Record*, 22(1): 15-20.

[12] Goh, P.C., 1989, A graphic query language for cartographic and land information systems. *International Journal of Geographical Information Systems*, 3(1):

245-255.

[13] Guting, R.H., 1988, Geo-relational algebra: a model and query language for geometric database system. In J.W. Schmidt, S. Ceri and M. Missikoff (eds.), *Proceedings of the International Conference on EDBT*, Venice, pp. 506-527.

[14] Herring, J.R., R.C. Larsen and J. Shivakumar, 1988, Extensions to the SQL query language to support spatial analysis in a topological database. In *Proceedings of GIS/LIS' 88*, San Antonio, 30 November-2 December.

[15] Huang, B., 1997, GeoSQL: a Visual Spatial SQL for Topological Relationships in GIS. In *Proceedings of the Workshop on Dynamic and Multi-dimensional GIS*, The Polytechnic University of Hong Kong, Hong Kong, August.

[16] Ingram, K.J. and W.W. Phillips, 1987, Geographic information processing using a SQL based query language. In *Proceedings of AutoCarto 8*, Baltimore, Maryland, pp. 326-335.

[17] ISO, 1995, *SQL Multimedia and Application Packages (SQL/MM) Part3: Spatial*, ISO Working Draft, September.

[18] Kruglinski, D.J., 1993, *Inside Visual C++*. Microsoft Press, USA.

[19] Lee, Y.C. and F.L. Chin, 1995, An iconic query language for topological relationships in GIS. *International Journal of Geographical Information Systems*, 9(1): 25-46.

[20] Ooi, B.C., 1988, Efficient query processing for Geographic Information System. PhD thesis, Monash University, Victoria, Australia.

[21] Roussopoulos, N., C. Faloutsos and T. Sellis, 1988, An efficient pictorial database system for PSQL. *IEEE Transactions on Software Engineering*, 14(5): 639-650.

[22] Samet, H. and W. Aref, 1994, Spatial data models and query processing. In W. Kim (ed.), *Modern Database Systems: The Object Model, Interoperability, and Beyond,* Addison Wesley/ACM Press, Reading, MA.

[23] Svensson, P. and Z.X. Huang, 1991, A query language for spatial data analysis. In *2nd Symposium on Large Spatial Databases*, Zurich, Switzerland.

[24] Tang, A.Y., T.M. Adams and E.L. Usery, 1996, A spatial data model design for feature-based Geographical Information Systems. *International Journal of Geographical Information Systems*, 10(5): 643-659.

在五周年庆祝会上，CPGIS 成员与中国科学院遥感所部分领导合影。从左到右：林珲、宫鹏、郭华东(遥感所所长)、孙以义(CPGIS 上海站经理)、丁跃民、柳林、李斌、周启鸣、杨崇俊(中科院百人计划入选者)、王超(遥感所副所长)、沈小平、周成虎(地理所资源与环境信息系统国家重点实验室主任) 和关蔚禾。

Road Network Extraction from High Resolution Airborne Digital Camera Data

Peng Gong* and Jinfei Wang†

*Center for Assessment and Monitoring of Forest and Environmental Resources
Department of Environmental Science, Policy and Management,
University of California, Berkeley, CA 94720-3310, USA
†Department of Geography
University of Western Ontario, London, Ontario, Canada, N6A 5B7

Abstract

Most road network extraction algorithms developed are based on linear analysis methods. These methods involve search of edges through edge filtering, morphological filtering, or gradient modelling. As image resolution increases from 10-30 m to 0.5-2 m, road networks will appear to be narrow areas rather than thin lines. This becomes a challenge for traditional linear analysis methods based on mask operations but creates an opportunity for classification based methods. We experimented with an advanced linear analysis, gradient direction profile analysis, and a few classification algorithms including a maximum likelihood classification, clustering and a contextual classifier for road network extraction using airborne digital camera data acquired over Livermore, California with approximately 1.6 m spatial resolution.

Results indicate that both the linear extraction and image clustering algorithms worked reasonably well. The linear extraction method requires some preprocessing such as filtering of the original image. Best road network results have been obtained by applying the linear extraction algorithm to a morphologically filtered image that was generated by combining the near infrared (NIR) and red (R) image bands through NIR/R+NIR. With this method, the correctly extracted road pixels account for 78.7% of the total road pixels obtained from image interpretation with field verification. The image clustering method resulted in 74.5% correctly extracted road pixels. The contextual classification resulted in relatively noise-free road networks in new residential areas but omitted some roads at older residential areas. When experimenting with the images resampled at approximately 3 m and 5 m resolution, the best overall accuracies for road extraction decreased to 74.6% and 61.6%, respectively.

摘 要

大部分道路提取的算法基于线性特征提取。这类方法大致由边界搜索、形态滤波或梯度模拟。当图象空间分辨率从10-30米提高到0.5-2米时，道路则由线状特征变成较狭窄的面。这既对传统线性提取算法带来了挑战，又给分类算法制造了机会。本文使用在加州 Livermore 市获取的大约 1.5 米分辨率的多光谱航空数字摄影图象，试验了一种先进的线性特征提取方法－梯度方向剖面分析法，并与一些分类算法进行比较，其中包括最大似然法、簇分析法和一种上下文分类的方法。

试验结果表明线性提取和簇分析法对道路提取均有成效。使用线性提取法之前需对原始图象进行一些预处理。最佳结果是先对原始图象的近红外(NIR)和红波段(R)进行NIR/R+NIR 运算。然后使用亮度扩展的数学形态学方法对上述运算所得图象进行处理。再从该处理图象上运用梯度方向分析法进行道路提取，精确度可达78.7%。簇分析法所得道路网的精确度可达 74.5%。当图象空间分辨率降低至大约 3 米或 5 米时，最佳精确度分别降低 74.6% 和 61.6%。

I. INTRODUCTION

Road network changes constantly at many rural-urban fringe areas due to urban expansion. Urban planners and decision makers on land use development often have obsolete land use information because operational mapping methods based on manual interpretation of aerial photographs usually take a year or two to complete from the time of aerial photography. Research efforts have been made to develop computer analysis algorithms for road network extraction (e.g., Wang et al., 1992; Gruen and Li, 1995; Wang et al., 1996) and land-use mapping (e.g., Gong and Howarth, 1992) from satellite images. On 10-30 m spatial resolution satellite images, roads are linear features

1082-4006/97/0301~2-51$3.00

represented by valleys or ridges of brightness. Wang and Liu (1994) grouped 4 types of line extraction methods that could be applied to road network extraction. They are (1) gradient operator and mask convolution method, (2) gradient direction profile analysis (GDPA) method, (3) mathematical morphology analysis method, and (4) knowledge-based method. Because the contrast between a road and the image background varies both spatially and spectrally, the use of multispectral data helps reduce road ambiguity (Wang and Liu, 1994).

Among various satellite and airborne sensing technologies, low-cost high spatial resolution digital CCD (Charge-Coupled Device) cameras are developing rapidly. The interest in multispectral imaging with digital CCD cameras has been increasing (King, 1995). High precision georeferencing techniques are available through integrating GPS (global positioning systems) and INS (inertial navigation systems) with digital photography (e.g., Schwarz et al., 1993). It is now possible to have high geometric and radiometric quality digital camera images on airborne platforms with spatial resolutions at the sub-meter level. In addition, 1-4 m resolution satellite imagery will soon become available (Fritz, 1996) and high spatial resolution digital image can be obtained by scanning aerial photographs. In a study of road networks extraction from scanned color-infrared films from aerial photography, Benjamin and Gaydos (1990) claim that 3 m spatial resolution is most suitable for road network extraction in Cupertino, California. They applied clustering and editing instead of the more sophisticated line extraction algorithms to scanned data resampled to different spatial resolution (1-5 m). Since most roads in urban areas are wider than 5 m, road networks become narrow areas rather than brightness valleys or ridges on images with pixel sizes smaller than 5 m by 5 m. At a spatial resolution better than 5 m, it is possible to extract road network with classification methods.

Our questions are:

- how well can traditional linear extraction algorithms perform when applied to those high spatial resolution images?
- to what extent, can classification methods be used for road network extraction purposes?
- how can methods in the two different paradigms be used to complement each other for improved road network extraction and land use classification?

This paper presents some of our efforts toward answering the first two questions. The objective was to compare classification methods with line extraction algorithms for road network extraction from digital camera images resampled at different spatial resolutions. A supervised per-pixel classification, a clustering, and a cover-frequency based contextual classification method were applied to classify an airborne digital camera image. A few road cover types were included in each classification. A gradient direction profile analysis algorithm was also applied to the same image. Road extraction results are presented and discussed.

II. STUDY SITE AND DATA

The study site is located on the east border (121°43' W, 37°41' E) of City of Livermore, California (Figure 1). The 1:24,000 USGS topographic map was last updated in 1981. Since then, new residential areas and road networks have been built. On June 30, 1995, an imaging system consisting of 4 Kodak DCS-200 cameras was used on board of an aircraft to acquire multispectral images over the study area. The four cameras simultaneously acquired images at 450 nm, 550 nm, 650 nm, and 850 nm, respectively, with a band width of 80 nm. The single-band images from individual cameras were then geometrically corrected and resampled to form a multispectral data set. Part of the red band image is shown in Figure 2. The spatial resolution is approximately 1.6 m.

Roads in the study area are mostly asphalt paved. Due to the fact that pavement ages are different, the asphalt road surface looks different in brightness. Older pavements appear brighter than the background among the visible bands but darker in the near-infrared band. Sidewalks are the brightest in the green and red bands of the image. There are also unpaved roads in the non-irrigated grass land in the study area. Grass lands in the build-up areas are usually irrigated.

III. ROAD EXTRACTION METHODS AND ACCURACY ASSESSMENT

Road network extraction methods generally involve five steps: image preprocessing, obtaining initial road network, noise removal, thinning, and pruning (Figure 3). The purpose of image preprocessing is to enhance road network features for subsequent analysis. This includes spatial filtering such as the

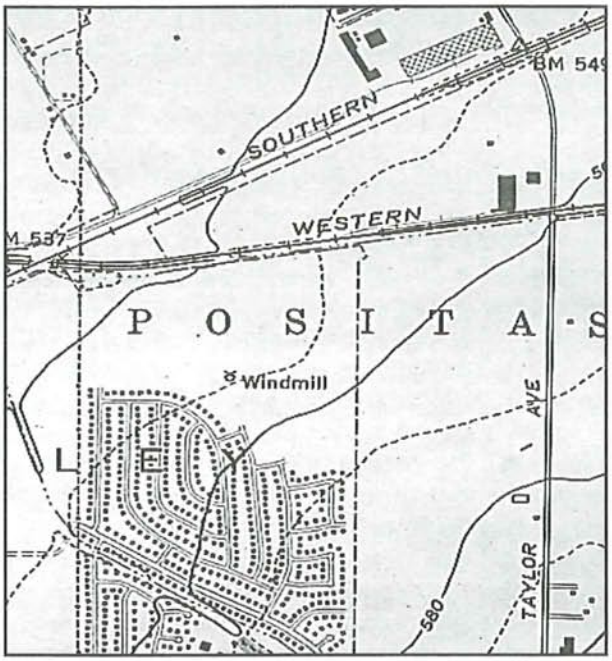


Figure 1. 1981 USGS 1:24,000 topographic map of the study area (note: presented map sacle may not be 1:24,000)

use of a median filter to preserve edges and remove speckle noise, morphological filtering, band ratioing and linear transformation such as the use of principal component analysis to enhance linear features in the original image (Jensen, 1996). Morphological filtering is usually conducted on binary images (Wang et al., 1996; Destival, 1986; Lee et al., 1987). We undertook grey-level dilation filtering to the original image. It is essentially an operation in search of maximum from a local neighborhood defined by the structuring element (Pratt, 1991; Sternberg, 1986). Moving a 3 X 3 kernel over a grey-level image, we assign the maximum grey-level value to the pixel at the kernel center. This is equivalent to a special version of the dilation filtering using a 3 X 3 structuring element with all its kernel values being 1. Noise removal is to remove from the initial road networks relatively small patches of pixels that have been identified as road segments. Pixel patches smaller than a certain size are removed from the initial road network image. Thinning reduces the detected road network from a few pixels wide to one pixel in width. Pruning removes short-branches of dead-end roads according to their lengths.

Gradient Direction Profile Analysis

The GDPA algorithm used in Wang et al. (1992) was

(a) (b)

Figure 2. Digital camera image acquired over Livermore, California on June 30, 1995. (a) Red band, and (b) Near infrared band.

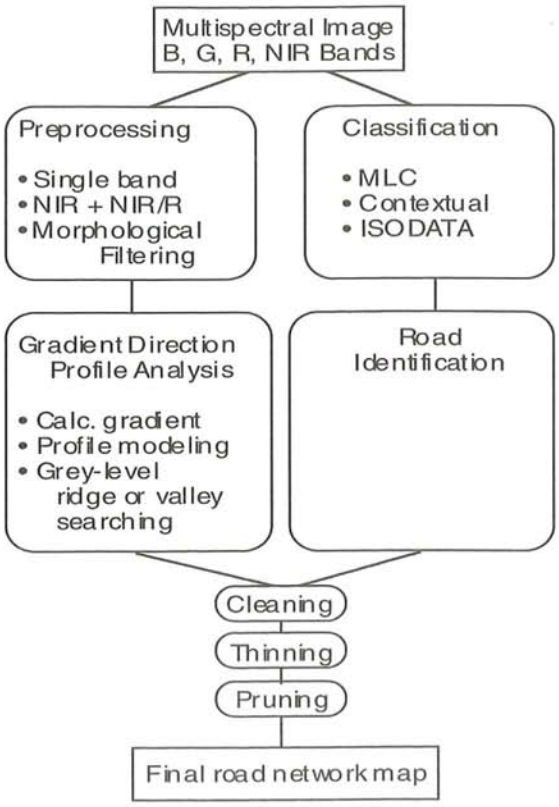


Figure 3. Procedures used in this study.

selected for use in this study. This algorithm first finds the greatest gradient direction for each pixel using the brightness values in a 3 by 3 neighborhood. A pixel is considered as a candidate road pixel if its greatest gradient exceeds a limit, T_d, specified by the analyst. The algorithm then searches among the candidate road pixels for pixels at ridge tops or valley bottoms of grey levels by modeling grey level profiles along the greatest gradient direction centered at each candidate pixel. The modeling is achieved through polynomial curve fitting along the gradient direction profile. The grey-level ridge top or valley-bottom positions are found through profile derivative and curvature analysis. The length of a profile, L, is specified by the analyst and the curvature of the polynomial function must be greater than a specified value, T_k. An initial road network is extracted from an image by adjusting the three parameters, T_d, L and T_k.

Details on GDPA is found in Wang et al. (1992). When applied to Landsat Thematic Mapper (TM) and SPOT multispectral and panchromatic data, the GDPA algorithm produced best road network results using visible spectral bands such as the green or red bands (Wang et al., 1992; Wang and Liu, 1994). This algorithm has also been successfully applied to extract various linear features from scanned topographic maps (Wang and Zhang, 1996).

Image Classification and Clustering

To classify roads from the digital camera image, we employed a clustering algorithm, a supervised maximum likelihood classifier and a cover-frequency based contextual classifier. The clustering algorithm is ISODATA (iterative self organizing data analysis technique) (Duda and Hart, 1974). The cover-frequency based contextual classifier is found in Gong and Howarth (1992). It first converts the original image to a land-cover map using a regular per-pixel maximum likelihood classification (MLC) algorithm, or a grey-level vector reduced image with a grey-level vector-reduction algorithm, or a cluster map through clustering. The algorithm then extracts frequencies of land cover, cluster, or grey-level vector from a neighborhood of a pixel and uses the frequencies in classification of land uses or discrimination of road types for that pixel. The size of the pixel neighborhood is specified by the analyst. The contextual classification requires supervised training to determine the frequencies for each class. The same training set was used for both the maximum likelihood classification and the contextual classification.

We used 8 classes in the supervised classification. They include four types of road covers according to different road surface colors and materials, new asphalt, older asphalt, concrete and railroad. The other 4 cover types are residential, industrial, well irrigated grass land and dry grass land. We selected training areas for the 8 land-use classes and used the maximum likelihood classifier and the contextual classifier to classify the imaged area.

Image Resampling

All image resampling methods can be achieved through image convolution. To resample the 1.6 m resolution data to coarser resolutions, one may use image averaging, nearest neighbor, bilinear or cubic convolution resampling methods (Shlien, 1979). These different methods can be realized through different design of weighting factors. For example, with image averaging a convolution kernel with equal weights is usually employed (Benjamin and Gaydos, 1990; Pratt, 1991). With nearest neighbor a unit weight is applied to the pixel closest to the

newly transformed pixel location while a triangular function is used to determine the weights for the two pixels closest to the newly transformed pixel location along the row and column direction respectively.

In this study, we compared the four different image resampling methods. As expected image averaging had the greatest effects in blurring the edges or linear features as image resolution degrades. Edges and linear features were better preserved in the resampled images generated by the remaining three resampling methods and those resampled images looked very similar. Therefore, we tested the linear extraction algorithms using the images resampled to 3 m and 5 m resolutions with the nearest neighbor method. We did not apply classification methods to the resampled images because as discussed in the introduction images with coarser spatial resolution would be less useful for road network classification.

Accuracy Assessment

In order to compare the performances among the various road network extraction algorithms, we conducted field study and manually traced the road networks in the image. The manually traced road networks were used as the truth image (Figure 3). Three measures can be calculated for the quantitative comparison (Wang and Liu, 1994).

Let N_{ce} be the number of correctly extracted road pixels; N_{tr} be the number of true road pixels; and N_{te} be the total number of extracted road pixels:

overall accuracy = N_{ce} / N_{tr}
commission error = $(N_{te} - N_{ce}) / N_{tr}$
omission error = $1 - N_{ce} / N_{tr}$

The overall accuracy is the fraction of pixels correctly extracted as roads. The commission error is the number of pixels incorrectly extracted as road pixels divided by the number of true road pixels. The omission error is one minus the overall accuracy. Since it is not independent, we will only report the overall accuracy and commission error. Obviously a smaller number is better with the commission error and a larger number is better with the overall accuracy.

IV. RESULTS

We applied the GDPA algorithm to each band of the original image. Among the visible bands, the road is relatively brighter than the background vegetation. We used a ridge-searching algorithm. However, the sidewalks are readily observable and are the brightest in the scene. Therefore, the road networks extracted from the visible bands are mostly sidewalks and are far from perfect. In the near infrared band, road networks and house roofs are the darkest. Since the roads are a few pixels wide, we applied grey-level dilation filtering to the original image using a 3 X 3 structuring element for 5 times (Pratt, 1991; Sternberg, 1986). This filtering expanded bright component in the image such as vegetation while shrank the dark components such as the road and the house roofs. We then applied the GDPA algorithm to the filtered image by searching for valleys (Td=3.0, L=3, Tk=5.0). The road network result has been considerably improved. However, a lot of house roofs have been extracted as roads and they are too big to be removed using the noise removal method. The final road network extraction results after thinning and pruning are shown in Figure 4. Similar results have been achieved by applying the GDPA algorithm to the second principal component image obtained from the original 4 bands.

The best result with the GDPA algorithm has been achieved using an image that combines the ratio between the near infrared and red band with the

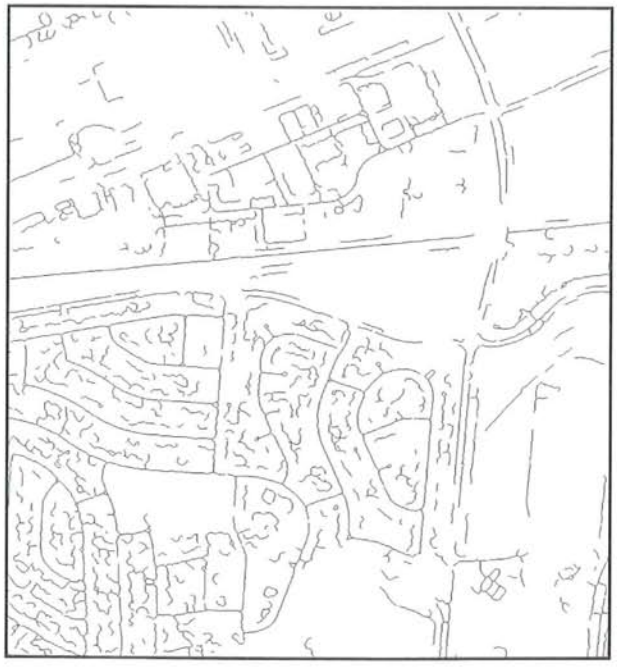


Figure 4. Final road networks obtained using GDPA from the NIR band.

Figure 5. The image obtained through NIR R+NIR.

Figure 6. The image after 3 iterations of dilation filtering from the image shown in figure 5.

near infrared band, [band 4/band 3] + band 4 (Figure 5). The newly generated image suppressed the shade and shadow in the original image through image ratioing, while the brightness of the vegetation portion of the image was enhanced. We then applied dilation filtering to the band combination image for 3 iterations (Figure 6). The initial road network was extracted from the dilated image (Td=1.0, L=5, Tk=3.0). The final road extraction result is shown in Figure 7.

With the image classification approaches, no preprocessing was applied to the original images. The four bands of image were clustered using ISODATA resulting in 49 clusters. After a cluster by cluster examination, we selected those clusters corresponding to road networks. All road network clusters were merged to form an initial road network map, which was further processed by noise removal, thinning and pruning, and resulted in a final road map (Figure 8). Supervised MLC classification was also performed on the original four bands. Similarly, noise removal, thinning and pruning were applied to produce a final road map. The final road network results obtained from MLC and from the clustering algorithm are similar. The clustering algorithm picked up more road details, particularly the rail roads, than the MLC method. It also included more non-road artifacts in the final result.

The cluster map was also used as the basis for the contextual classification. Classified road results were used as the initial road networks for subsequent noise removal, thinning and pruning. The final road network results obtained from the contextual classification are shown in Figure 9. We used a 9 by 9 pixel neighborhood size to generate cover frequencies in the contextual classification. We did not attempt to find an optimal window size for the classification. Generally, a large window size would remove more road details particularly when the road is relatively narrow. While a small window size would results in more unwanted details. At the relatively new residential area, the contextual classifier resulted in the cleanest road network. However, like the MLC results, the algorithm did not pick up roads on the upper left corner of the image. The roads are within the oldest residential area at that particular section of the image where roads are narrower than the newer residential sections. Some trees cover part of the road network causing the difficulties for the contextual classification methods. The supervised MLC method and the contextual classification methods may also be sensitive to training sample selection.

Table 1 lists some of the accuracy assessment results. With the 1.6 m resolution image, the best overall accuracy, 78.7%, is obtained by the GDPA

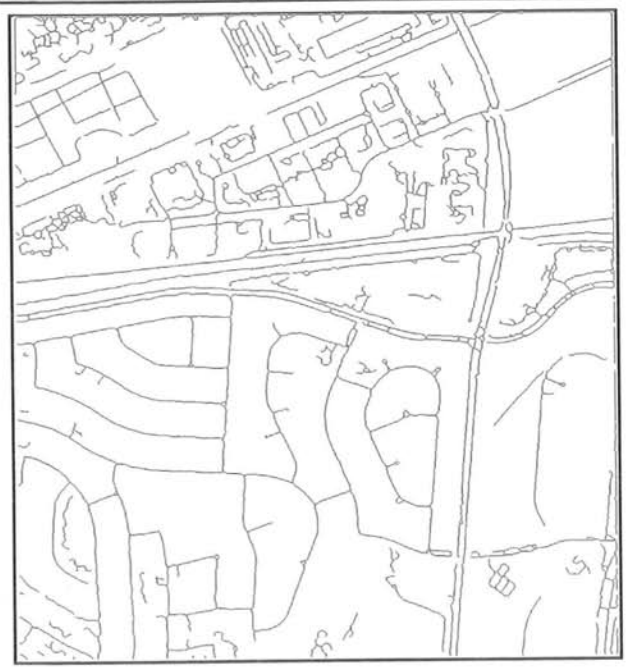


Figure 7. Final road networks obtained using GDPA from figure 5.

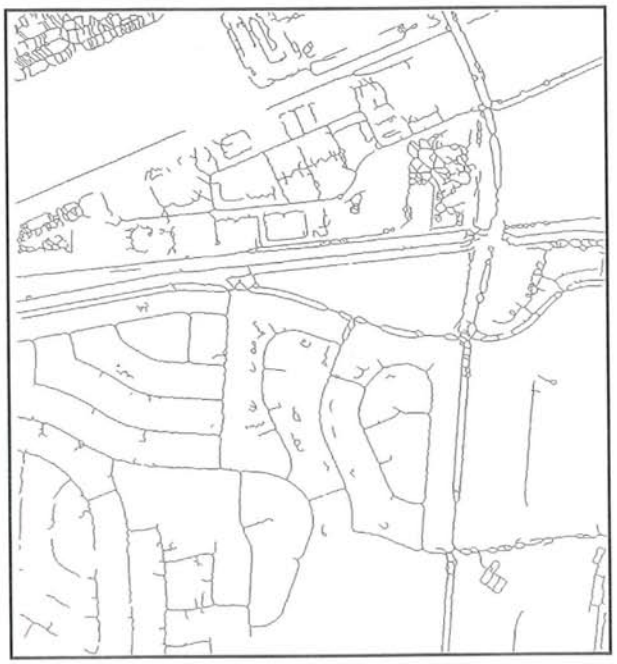


Figure 8. Road networks obtained from clustering-based method.

method from the image specially enhanced (b4 + b4/b3). The second best is achieved by the clustering method (74.5%). However, both methods have relatively high commission errors. The lowest overall accuracy is from the contextual classification results but the commission error is the lowest.

For the 3 m and 5 m resolution images, we conducted similar experiments as for the 1.6 m resolution data using the GDPA method. For each resolution, better accuracies have been achieved with the band 4 image and the specially enhanced image by taking the sum of band 4 and the ratio between band 4 and band 3. These were obtained from the preprocessed images with only 1 iteration of grey-level dilation. With the 3 m resolution, we obtained better overall accuracy and smaller commission error from the band 4 image in comparison with the 1.6 m resolution data. An overall accuracy of 74.6% has been achieved with the specially enhanced image. The commission error (0.832) from this image is smaller than that (0.984) obtained from the 1.6 m image with the same enhancement. When the resolution decreases to 5 m, the overall road network extraction accuracies are considerably lower.

V. DISCUSSION AND CONCLUSIONS

With the gradient direction profile analysis algorithm, previous studies using 10 - 30 m resolution Landsat TM or SPOT images indicated that visible spectral bands are more effective than the near infrared bands for the purpose of road network extraction . As the spatial resolution of digital imagery improves to 1 - 3 m, road networks are no longer lines of one pixel in width but narrow areas of several pixels wide. This makes it harder to directly apply line extraction algorithms to high spatial resolution images. In order to effectively use line extraction algorithms, it is necessary to preprocess the original image. Due to their spectral similarity with road networks, house roofs and drive ways are the primary disturbing factors to road network extraction. It would be desirable to completely remove the house roofs while keeping the road network at narrow widths. We found that a morphological filtering based on a dilation process applied to the near infrared band can accomplish this to a large extent. The visible spectral bands may not be the most effective bands for road network detection. The near infrared band exhibits the greatest contrast between road networks and their background. The best result using the GDPA method was achieved using an image produced by summing up the near infrared image and the ratio

Figure 9. Road networks obtained from the contextual classification method.

between the near infrared and the red band.

We had expected that better accuracies might be obtained by line extraction algorithms such as the GDPA method if the original high spatial resolution of the data was degraded. This expectation was not met in our experiments with 3 m and 5 m resolution images resampled from the 1.6 m resolution data. The best overall accuracy achieved with the 1.6 m data is only 4.1% higher than that from the 3 m data, while the commission error is also higher. We consider the best road network extraction results are comparable with the 1.6 m and the 3 m resolution data. However, the best overall accuracy obtained with the 5 m data drops considerably in comparison with those obtained from the 1.6 m and 3 m data. For similar types of landscapes, it seems to us that the procedure developed for the GDPA method would perform consistently well if image resolution is better than 1.6 m. If the resolution is finer than 1.6 m, all one needs to do is to apply the morphological dilation for a few more times. In conclusion, for similar types of landscapes the GDPA in combination with the preprocessing methods developed in this research works well for images with resolutions better than 3 m. For road network extraction 3 m resolution images may be more desirable from a computational point of view than finer resolutions because the image file is smaller.

Table 1. Performance evaluation of different road network extraction results

Images	Overall accuracy(%)	Commission error
band4	59.8	1.010
clustering	74.5	0.845
contextual	49.5	0.476
MLC	68.4	0.659
b4+b4/b3	78.7	0.984
band4 (3m)	66.0	0.951
b4+b4/b3 (3m)	74.6	0.832
band4 (5m)	53.5	0.882
b4+b4/b3 (5m)	61.6	0.693

With the classification based approaches, road networks can be obtained reasonably well, particularly with the clustering method. The variability among different road surfaces makes it impossible to treat road as a single class in image classification. Training must be carefully done if supervised classification is used. A large number of clusters is needed if clustering is used to extract road networks. With the cover-frequency based contextual classification algorithm, the most noise-free road networks are obtainable for relatively new residential areas. It works less effectively in older residential areas where trees block the visibility of road surfaces.

To help understand the performance of different road network extraction methods, the study area was purposely selected not to be as complicated as many old built-up areas could be. Old areas with high density of tree coverage present a problem to all road network extraction approaches. As an important component on large scale urban land use maps, road networks also have the potential in improving classification accuracies. It is worthwhile to explore the complementary roles that classification and traditional line extraction methods could play for improving both the road network extraction and image classification results.

The procedure described here may not work well for non-residential areas as can be seen from the industrial area in the image. The preprocessing method works well only when there are vegetation surrounding houses in relatively newer suburban areas. The morphological dilation procedure will not work well in areas where trees covering road surfaces and leaving the road networks to be imaged in different widths. For high resolution image data, image preprocessing is necessary and different

procedures are needed for different areas.

REFERENCES

[1] Benjamin, S., and L. Gaydos, 1990. Spatial resolution requirements for automated cartographic road extraction, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 1: 93-100.

[2] Destival, I., 1986. Mathematical morphology applied to remote sensing, *Acta Astronautica*, Vol. 13, No. 6/7, 371-385.

[3] Duda, R.O. and P.E. Hart, 1973. *Pattern Classification and Scene Analysis*, Wiley-Interscience, New York.

[4] Fritz, L.W., 1996. The era of commercial earth observation satellites, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 1: 39-45.

[5] Gong, P. and P. J. Howarth, 1992. Frequency-based contextual classification and grey-level vector reduction for land-use identification. *Photogrammetric Engineering and Remote Sensing*, 58(4):423-437.

[6] Gruen, A. and H.H. Li, 1995. Road extraction from aerial and satellite images by dynamic programming, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50, No. 4:11-20.

[7] Jensen, J.R., 1996. *Introductory Digital Image Processing, a Remote Sensing Perspective*, 2nd Ed., Prentice Hall, Upper Saddle River, New Jersey.

[8] King, D.J., 1995. Airborne multispectral digital camera and video sensors: a critical review of system designs and applications, *Canadian Journal of Remote Sensing*, Vol. 21, No. 3, 245-273.

[9] Lee, J.S., R.M. Haralick, and L.D. Shapiro, 1987. Morphologic edge detection, *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 2:142-156.

[10] Pratt, W., 1994, *Digital Image Processing*, John Wiley and Sons: Toronto.

[11] Schwarz, K-P., M.A. Chapman, M.E. Cannon, and P. Gong, 1993. An integrated INS/GPS approach to the georeferencing of remotely sensed data, *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 11, 1667-1674.

[12] Shlien, S., 1979. Geometric correction, registration, and resampling of Landsat Imagery, *Canadian Journal of Remote Sensing*, 5(1):74-87.

[13] Sternberg, S.R., 1986. Grayscale morphology, *Computer Vision, Graphics, and Image Processing*, 35:333-355.

[14] Wang, D, D.C. He, L. Wang, and D. Morin, 1996. Extraction of urban road networks from SPOT HRV images. *International Journal of Remote Sensing*, 17(4):827-833.

[15] Wang, J.F., P. M. Treitz and P.J. Howarth, 1992. Road Network detection from SPOT imagery for updating geographical information systems in the rural- urban fringe, *Int. J. Geographical Information Systems*, Vol. 6, No. 2:141-157.

[16] Wang, J.F. and W. Liu, 1994. Road detection from multispectral satellite imagery, *Canadian Journal of Remote Sensing*, Vol. 20, No. 2:180-189.

[17] Wang J.F. and K. Zhang, 1996. Linear cartographic data acquisition from scanned topographic maps. *Geographic Information Sciences*, Vol. 2(1):12-23.

【GIS Industrialization】

Metadata Strategy, Data Directory System and Emerging National Spatial Data Infrastructure in Australia

Wei Pei

GIS Branch, State Forests of NSW
423 Pennant Hills Road, Pennant Hills, NSW 2120 Australia

I. INTRODUCTION

Increasing public and political concern for proper environmental assessment and management, land-use conflicts, the issues surrounding ecologically sustainable development and biodiversity are resulting in a demand for high quality data on Australia's natural resources.

Due to the diversity of natural resources information, the metadata required to describe these resources are equally diverse. Over time an array of metadata formats have evolved which enable various organisations, agencies, academic, community or private industry to tailor metadata to specific needs. In most cases, the capture of metadata is easily and more readily achieved at the highest level, simply because summaries of datasets require less description and more general, while at lower levels, additional information is recorded. The development of any metadata standards must recognise this complex and multi-layer relationship and provide a facility to permit the capture of metadata at different levels and at different times.

In order to achieve a better understanding towards the complexity and multi-layer relationship associated with the capture of metadata, it's necessary to have a brief description of the spatial data directory systems in Australia. Using this as a starting point, metadata strategy for national, jurisdictional, theme-based and custodian agency data directory can be explored.

II. TERMINOLOGIES OFTEN USED IN AUSTRALIA

Metadata: refers generally to dataset descriptions and is equivalent to 'meta-information' (which is possibly a better term to use because 'metadata' is commonly used in data management to refer to data dictionaries).

> The popularly used metadata is normally *feature-based, cover-based, project-based* and *dataset-based*. There are also metadata about hardware and software.

Directory: refers to a collection of descriptions of datasets together with subject and spatial indexing. There are three types of directories most commonly seen in Australia:

> *National directory*: directory managed by Commonwealth government of Australia and functions as the repository of the master keyword thesaurus and gazetteer of spatial objects. An example of national data directory developed at an early stage is: National Directory of Australian Resources (**NDAR**), and **ASDD** at a later stage by National Resource Information Centre (**NRIC**)(1).
>
> *Jurisdictional directory*: directory managed at State/Territory level. Considerable efforts have been undertaken in Queensland, Western Australia, New South Wales, South Australia,

Declaration: The views expressed in the article are those of the author's. They don't necessarily reflect the views of any Australian government organisation or agency.

1082-4006/97/0301~2-60$3.00

Victoria to develop their spatial data directory. For example, in August 1995 Department of Land and Water Conservation(**DLWC**) in the State of New South Wales published the first **NSW Natural Resources Data Directory**. This directory was published in CD-ROM/Disk form for use on IBM PC computers and included an easy to use graphical search capability.

Theme-based directory: directory developed by theme-based organisations. For example, Australian Coastal and Marine Directory (**Blue Pages**), or Australian National Geoscience Information System (**NGIS**)(2).

Custodian agency: an organization which has the custodianships to take responsibility for collecting and maintaining of the dataset.

III. NATIONAL CONTEXT

It is estimated by the Economic Studies and Strategies Unit of Price Waterhouse[3] that for the period 1989-94 approximately $1 billion has been spent in Australia on investment in geospatial data. This investment produced benefits within the economy in the order of $4.5 billion. The study also found that this investment has saved users approximately $5 billion.

The US Government have long recognized the management of geospatial data as an important national issue and as early as April 1994 an Executive Order was made by President Clinton on "Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure".

In Australia there is a widespread support within Government and industry for the development of an Australian Spatial Data Infrastructure (**ASDI**)(3) in order to support decision making, facilitate timely access to and share reliable consistent geospatial data by government agencies and community, and maximize integration of datasets . Development of the ASDI is being managed by the Australia New Zealand Land Information Council (ANZLIC).

This emerging ADSI will comprise a distributed network of databases, linked by common policies, standards and protocols to ensure compatibility. It will eventually become an information window to Australia's physical, natural and administrative environment.

IV. EVOLVING '*directory*' CONCEPT

An important component in this emerging ASDI is the implementation of a "national directory system" with a distributed network structure to help access Commonwealth and State/Territory spatial data.

The concept of national directory is evolved from **'central databank'**, a facility first initiated by the National Resource Information Centre in 1988 (NRIC, a branch of Bureau of Resource Sciences, federal Department of Primary Industries and Energy (**DPIE**)(4)). By 1990, it is widely accepted that **'national directories/indexes'** for land information management are necessary to facilitate decision making. The original design for a national directory was for a single centralised system with NRIC carrying out all the metadata collection and maintenance, which later turned out to be a large task far beyond the resources of NRIC.

Under the recognition that a dataset description is best maintained by that dataset's custodian, this led to a concept of **'node model'**. Each State/Territory directory would form a node with State/Territory being responsible for the maintenance of the directory entries at that node. Ideally, these node directories would then be accessed from one or more user interfaces. The node provides periodic updates to the national directory, creating an accumulation of all relevant information held at the node level — thus forming a **'distributed directory system'**.

The optimum fully distributed system is currently being evaluated by the ANZLIC Metadata Working Group and is expected to be progressively implemented from February 1998.

Prior to full implementation of a fully distributed system, a process common software has been developed by NRIC to facilitate the transfer of dataset descriptions between State/Territory installation and the national centre. This is **FINDAR directory software the prototype Australian Spatial Data Directory (ASDD) system** (Facility for Interrogating the National Directory of Australia Resources)(5). FINDAR prototype ASDD allows metadata records from subordinate nodes, which have been flagged as suitable for transfer to the national node, to be uploaded to that node. The software comprises three main components: tables of attribute, keyword and spatial data for the direc-

tory entries; a gazetteer of geographic entities for spatial indexing and searching; and a thesaurus of standard terms for subject indexing and searching.

In addition to the distributed directory system model(NDAR ASDD), a number of high-level spatial data directories have been developed at national, jurisdictional level independent of each other, with some of them theme-based. An example is the Australia Land Information Group(**AUSLIG**) metadata directory(6).

V. ANZLIC GUIDELINE ON CORE METADATA ELEMENTS

A critical step towards the implementation of a national directory system for spatial data is the development of "**ANZLIC Metadata Guidelines on Core Metadata Elements**"(7). The Australia and New Zealand Land Information Council (ANZLIC), is the peak intergovernmental body responsible for spatial data management for Australia and New Zealand. It currently manages a National Strategy for land and geographic information. To support the adoption of the Guidelines ANZLIC also produced and freely distributes a Microsoft Access Metadata Entry Tool.

This is a guideline consistent with the Content Standards for Digital Geospatial Metadata developed by US Federal Geographic Data Committee(**FGDC**), and with the Australia New Zealand Standard on Spatial Data Transfer **AS/NZS 4270**. However, Australia's approach for the mandatory metadata items is deliberately less ambitious than what has been attempted in the US. The US counterpart consists of some 220 items (elements) which are intended to describe digital geospatial datasets adequately for all purposes, while in ANZLIC Guidelines, 31 core metadata items (grouped into 9 categories) are identified as generally common for all types of data. (Table 1).

The International Standards Organization (ISO) has a draft Metadata Standard developed through its Geomatics Technical Committee ISO/TC 211 which has been developed with input from FGDC, ANZLIC, Dublin Core and other metadata guidelines. This will have two levels, a "Cataloguing Metadata" level similar to ANZLIC and Dublin Core and a complete level which includes some 300 elements. It is expected that ANZLIC will adjust the ANZLIC core elements to the ISO Level 1 when finalized.

At the national level, one of the natural resource data directory which has adopted ANZLIC Metadata Guidelines is the Prototype Australian Spatial Data Directory (**ASDD**)(8) managed by National Resource Information Centre (NRIC) on behalf of ANZLIC, a prototype of the new version of national directory system which consists of updated description of Key Commonwealth and State/Territory datasets. It supersedes the former National Directory of Australian Resources (NDAR).

The optimum fully distributed system is currently being evaluated by the ANZLIC Metadata Working Group and is expected to be progressively implemented from February 1998.

The national programs provide the framework for any corresponding activities at State/Territory level. In State of New South Wales, NSW Metadata working group was formed in March 1997 to take advantage of national activities. The NSW Natural Resources Data Directory data collection and conversion to the **ANZLIC Metadata Entry Tool** is now close to completion and the ANZLIC Tool has been supplied for use in several key NSW natural resource agencies(9).

VI. ANZLIC 'PAGES' CONCEPT ADOPTED BY DIFFERENT LEVELS OF NATURAL RESOURCE AGENCIES AND INDUSTRY

In order to include additional meta-information which is not included in the National Metadata Directory System, ANZLIC metadata has adopted a '**Pages**' concept as the basis for a national metadata framework, where more general information is recorded at the highest level (Page 0) and additional information is recorded at lower levels (Page 1, Page 2). A conceptual indication of how the Pages Concept is the foundation of the national directory system is shown in Figure 1.

This 'Pages' concept has been adopted by the **RACAC** (Resource and Conservation Assessment Council) in the Regional Forest Agreements Project (**RFA**) of NSW for documenting additional metadata on datasets, software and hardware used by stakeholders of the project.

Subsequent pages (i.e. Page 1, Page 2, etc.) provide the opportunity for data custodian agencies at the national, state, local government, academic, community or private industry levels to include addi-

Table 1. Core Elements: A Summary

Category	Element	Comment
Dataset	Title	The ordinary name of the dataset.
	Custodian	The organisation responsible for the dataset.
	Jurisdiction	The state or country of the Custodian.
Description	Abstract	A short description of the contents of the dataset.
	Search Word(s)	Words likely to be used by a non expert to look for the dataset.
	Geographic Extent Name(s) OR	A picklist of pre defined geographic extents such as map sheets, local government areas, catchments, that reasonably indicate the spatial coverage of the dataset.
	Geographic Extent Polygon(s)	An alternate way of describing geographic extent if no pre-defined area is satisfactory.
Data Currency	Beginning date	Earliest date of data in the dataset.
	Ending date	Last date of information in the dataset.
Dataset Status	Progress	The status of the process of creation of the dataset.
	Maintenance and Update Frequency	Frequency of changes or additions made to the dataset.
Access	Stored Data Format	The format or formats in which the dataset is stored by the custodian.
	Available Format Type	The formats in which the dataset is available, showing at least, whether the dataset is available in digital or nondigital form.
	Access Constraint	Any restrictions or legal prerequisites applying to the use of the dataset, eg. licence.
Data Quality	Lineage	A brief history of the source and processing steps used to produce the dataset.
	Positional Accuracy	A brief assessment of the closeness of the location of spatial objects in the dataset in relation to their true position on the Earth.
	Attribute Accuracy	A brief assessment of the reliability assigned to features in the dataset in relation to their real world values.
	Logical Consistency	A brief assessment of the logical relationships between items in the dataset.
	Completeness	A brief assessment of the completeness of coverage, classification and verifi cation.
Contact Information	Contact Organisation	Ordinary name of the organisation from which the dataset may be obtained.
	Contact Position	The relevant position in the Contact Organisation.
	Mail Address 1	Postal address of the Contact Position.
	Mail Address 2	Aust and NZ: Optional extension of Mail Address 1.
	Suburb or Place or Locality	Suburb of the Mail Address.
	State or Locality 2	Aust: State of Mail Address. NZ: Optional extension for Locality.
	Country	Country of the Mail Address.
	Postcode	Aust:Postcode of the Mail Address. NZ: Optional postcode for mail sorting.
	Telephone	Telephone of the Contact Position.
	Facsimile	Facsimile of the Contact Position.
	Electronic Mail Address	Electronic Mail Address of the Contact Position.
Metadata Date	Metadata Date	Date that the metadata record for the dataset was created.
Additional Metadata	Additional Metadata	Reference to other directories or systems containing further information about the dataset.

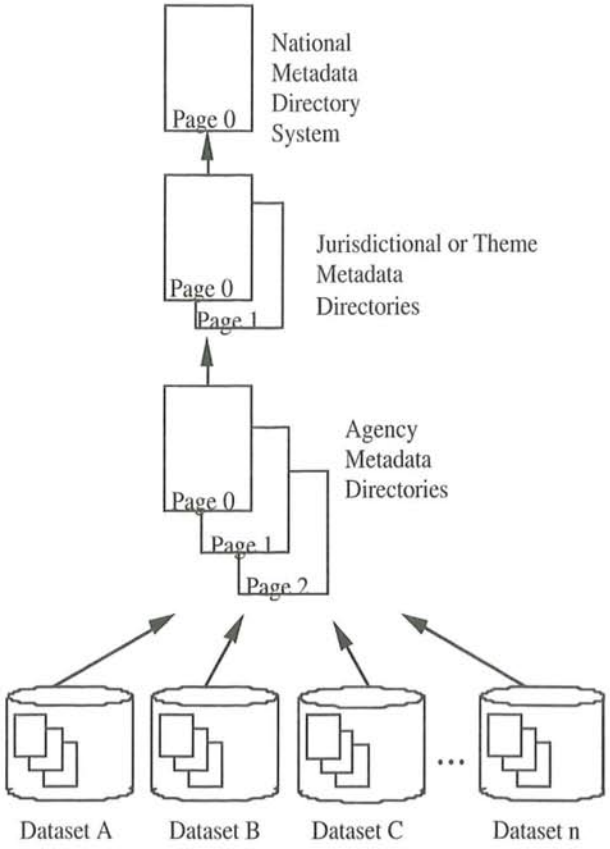


Figure 1. The Pages Concept

tional information not required in Page 0. This additional information may be in the form of sub-elements of specific Page 0 core metadata elements or entirely new and unrelated metadata elements. However, in order to ensure uniformity, it is suggested that any new metadata elements should be consistent where possible with corresponding metadata elements in the FGDC Content Standards.

VII. FEATURE-BASED METADATA SYSTEM AS BASIC META-INFORMATION

In order to retain corporate knowledge of the characteristics and processing history of datasets that an agency uses, the most detailed level of metadata is needed to ensure the efficient management and effective utilization of data within a custodian agency.

A **feature-based metadata system** is developed by the GIS Branch of the **State Forests of NSW**, with ESRI Australia Pty Ltd providing consulting services. The system is built on the top of Arc/Info GIS software and maintained by State Forests of NSW. The design of the metadata system is to tag every arc in an Arc/Info coverage with a metadata-id, which is stored in metadata table(.MDA). This .MDA table is linked to arc attribute table(.AAT) with metadata-id as common item. Therefore this allows the attributes in the .MDA to appear to be part of the .AAT. The tagging of arc with metadata-id is carried out during the process of data capture and literally built in the process of data capture, making this a compulsory step in data collection.

With this metadata system to provide the detailed, comprehensive information on data characteristics and data processing history, metadata required for higher level directory systems can be summarized or rolled up from the feature-based system.

VIII. CONCLUSION

There is a close link between metadata standards and the structure of spatial data directory system. Metadata standards and directory management software should be designed to provide flexibility in describing datasets which exist in a wide range of formats and in meeting the needs of different users. It should also facilitate the capture of metadata at different levels and different times.

It is widely recognized that metadata is best collected and maintained by custodian agency. Data custodian is also at the best position to summarize low level metadata (such as the feature-based metadata) into higher levels, for example, cover-based, project-based and dataset-based. These higher level metadata can then be submitted to nodal data directory. Here the items that are suitable for parent directory be tagged and uploaded into data directory above. Through the transfer of information towards higher levels, a summarisation of information has been achieved. At the peak, national data directory holds an accumulation of metadata stored at its subordinate directories and duplication of data storage is avoided. It also leaves the metadata maintenance to node/custodian, thus facilitate timely update of national data directory.

The distributed data directory system endorsed by ANZLIC is the directory model that could facilitate the flow and accumulation of metadata information. Combining with directory management software as interface, data summarisation, keyword indexing and spatial searching can be automated.

From ANZLIC experience, it is advisable to keep compulsory metadata items to the generally common, basic items that are common for all types of datasets. With the "Page" structure, additional metadata information can be accommodated by adding "pages" when more detailed information is required. The 'Pages' approach endorsed by ANZLIC is an important step towards the forming of metadata standards with flexibility as well as uniformity and consistency.

In order to ensure the speedy implementation of metadata standards, metadata entry tools should be produced by peak government organisations and freely distributed to all the major government agencies.

As stated in draft ASDI, the future trend is to develop a distributed network of databases, linked by common policies, standards and protocols to ensure compatibility and timely update of spatial information.

IX. HOW TO GET MORE INFORMATION ABOUT METADATA ACTIVITIES IN AUSTRALIA ?

Here is a List of Useful Web Sites and Metadata On-Line Directories (by the sequence they appear in the text). The Environmental Resources Information Network (**ERIN**) internet site supports a collection of pointers to "standards" servers and is a good starting point for surfing the internet for data standards. Please check it up in reference (10).

(1) National Directory of Australian Resources (NDAR) developed by National Resource Information Centre(**NRIC**):
http://www.nric.gov.au/nric/data/ndar_overview.html

(2) ANZLIC www sites of Related Interest supports a collection of pointers to various jurisdictional, and theme-based natural resources data directories:
http://www.anzlic.org.au/anz_site.htm

(3) Commonwealth Spatial Data Committee(**CSDC**), Draft Commonwealth Position Paper on the Australia Spatial Data Infrastructure(**ASDI**), Draft 4, 17 November 1997.
http://www.auslig.gov.au/pipc/csdc/sdi4.htm

(4) Bureau of Resource Science (**BRS**):
http://www.brs.gov.au

(5) Johnson, B.D., Shelley, E.P., Taylor, M.M., & Callahan, S., 1991 — The FINDAR directory system: a meta-model for metadata. In Medyckyj-Scott, D, Newman, I, Ruggles, C, & Walker, D, 1991 (Eds) — **Metadata in the Geosciences**. Group D Publications Ltd, Loughborough, UK.

(6) Australia Land Information Group(**AUSLIG**) Metadata: http://www.auslig.gov.au/meta/meta.htm

(7) **ANZLIC** Metadata Guidelines on Core Metadata Elements from ANZLIC (Australia New Zealand Land Information Council):
http://www.anzlic.org.au/metaelem.htm
http://www.anzlic.org.au/anz_sdts.htm
http://www.anzlic.org.au/anzstrdd.htm

(8) Prototype Australian Spatial Data Directory (**ASDD**), supersedes NDAR, currently being populated with records which follow the ANZLIC Core Metadata Guidelines, managed by National Resource Information Centre(NRIC), supercedes NDAR in reference (1):
http://www.nric.gov.au:80/nric/data/data.html
http://purl.nla.gov.au/net/asdd

(9) Draft New South Wales Natural Resources Metadata Management Strategy, 1997, NSW Natural Resources Information Management Steering Group (**NRIMS**) NSW Metadata Working Group.

(10) A collection of pointers to "standards" servers from Environmental Resources Information Network(**ERIN**):
http://www.environment.gov.au/gis/gis_standards.html

(11) A discussion forum on ANZLIC Guideline: Core Metadata Elements supported by ERIN:
Subscribe by sending an email containing the message, "subscribe ozmeta-l", to **majordomo@erin.gov.au**

【GIS 产业化】

美国环境保护署环境及空间数据库系统
USEPA Environmental and Spatial Databases

毛宛红* 李 刚†

*AnteonCorp., USEPA Region9, 75 Hawthorne Street
San Francisco, CA 94105
†Department of Geography & Environmental Studies
California State University Hayward, CA 94542

Abstract

In order to better manage large quantity of environmental and spatial information, US EPA has built up Spatial Data Library Systems (ESDLS), Envirofacts and Spatial Data Clearinghouse. ESDLS is a repository for the Agency's new and legacy geospatial data holdings. It enables Agency access to its geospatial data in ARC/INFO format, integrates these data in a standardized, framework. Environfacts was developed as an enterprise data repository for Agency data systems. It is a relational database implemented in the Oracle Relational Database Management System and contains seven EPA's program system databases. Geospatial Data Clearinghouse provides a pathway to find information about geospatial data used at EPA. Although ESDLS can be accessed through several GIS applications, it is mainly for EPA internal uses. Meanwhile, Environfacts can be accessed by the public through Internet. Spatial Data Clearinghouse currently provides public access to its metadata. However, it is promised that public can access its geospatial data via the Internet in the future.

美国环境保护署(US Environmental Protection Agency, EPA)是美国政府民用机构中空间(地理)数据的最大用户之一。为了最大限度地利用他们在空间信息资源上的投资,EPA于1987年展开了自己的地理信息系统(Geographic Information Systems, GIS)计划。这个EPA的GIS计划主要由三部分组成，分别是:(1)基础组织机构;(2)空间数据系统;(3)结合GIS和已有的环境数据系统进行环境分析于评价。基础组织机构包括分布于十个EPA大区的GIS中心，位于北卡罗来纳州“研究三角区”内的技术服务，环境监测和项目评价部，以及在弗吉尼亚州总部的一些研究办公室。空间数据系统储存了EPA的环境污染监测，处理及控制设施点的分布，以及一些生态与自然资源的数据。另外，它还结合了美国国家气象站的气象数据，美国人口普查局的普查区，道路及人口数据,和美国地质调查所的地形及其它地理底图。GIS应用分析评价到目前已展开了一些生态系统典型区的环境监测和评估，如在五大湖地区，墨西哥湾及切萨皮克(CHESAPEAKE)湾等进行的将各种时空尺度的环境监测数据转换成生态条件评估和未来自然资源可利用性的预测等。下面将重点介绍EPA的空间数据系统和环境数据系统的数据库结构，数据储存格式，数据内容，数据更新状况，数据提取方式和数据的应用。这两个数据系统是EPA数据管理三大系统中的两个主要组成部分。第三个部分是空间信息交换所。

一. EPA 空间数据系统 (EPA Spatial Data Library Systems, ESDLS)

EPA空间数据系统(ESDLS)是由EPA信息资源管理办公室和EPA地理信息系统计划组负责建立的.ESDLS3.0版是目前最新的EPA国家空间数据系统,它储存了EPA所有的空间数据，可以用来支持各种GIS的应用。

EPA建立ESDLS的基本目标是：建成一个国家范围内统一的空间数据库;将空间数据整体化于一个标准的框架内；应用恰当的标准和指南来管理和使用数据；提供ARC/INFO格式的空间数据给EPA的用户；逐步发展成为环境数据库的一个基本部分。

1. ESDLS 的数据来源

ESDLS的数据来源包括：1990年美国人口普查数据；1992TIGER (Topologically IntegratedGeographic Encodingand Referencing)；美国地质调查局地理名称信息系统2；水文单元代码；1：2,000,000数字线化图DLG(包括街道，水系，行政界限等)；EPA环境数据；环境污染优先治理区界线；主要河流分布；1994年鱼消耗量；邮政编码界限和清单等(请参阅http://www.epa.gov/nsdi/pages/title_browse.html)

1082-4006/97/0301~2-66$3.00

在建立过程中，ESDLS逐渐纳入EPA以外的数据。由于这些数据来源繁杂，ESDLS采用了分步实施的方法。每一步都规划了所有归并数据的地理范围，专题和内容，以及用户服务范围和完成日期。详见表1。

3).第三级生态区界线：此生态区界线取自于应用美国大陆的生态区界线和改编的Omernik'sframework生态区界线。这些界线包括：修改过的生态区，定义的亚生态区和一些参考点位。它提供了一个生态空间构架，

表1。EPA空间数据系统实施步骤

版本	地理范围	专题内容	用户范围	完成日期
1.1	美国大陆	1992，1990年TIGER街区， 街区组界线和中心点； EPA基础设施点； 地理名称点位(GNIS2)数据(医院，学校， 人类活动区)； EPA超级基金项目(Superfund)区界线； 1:100,000县，州界线； 1:2,000,000DLG道路，水文，县州界线	EPA	1996,3,14
1.2	美国大陆	新增加： 主要河流； 1:250,000水文单元	EPA	1996,3,31
1.3	美国大陆	新增加： Omernik生态区域； 大气环境信息及提取项目数据	EPA和其它 联邦机构	1996,8,30
2.0	美国大陆， 阿拉斯加，夏威 夷，波多黎各， 以及维尔京群岛	新增加： TIGER92非美国大陆地区； GIRAS土地利用； 1994年鱼消耗情报； 气象站的稳定性数列数据	EPA和其它 联邦机构	1996,9,30
3.0	美国大陆，阿拉 斯加，夏威夷， 波多黎各，以及 维尔京群岛	新增加： 非美国大陆上的县，州边界； 1994年的联邦土地DLG数据； 5位数字的邮政编码边界(美国大陆， 阿拉斯加，夏威夷)	EPA和其它 联邦机构	1997,9,30

2. ESDLS的数据

空间数据系统的空间数据采用ARC/INFO格式，有标准的和统一的数据使用和管理规则。这些空间数据按地区等级分为国家级，州级和县级。比例尺分别是：1：10万(县级)，1：25万(州级)，1：200万(国家级)。地理区域包括：美国大陆，阿拉斯加，夏威夷，波多黎各，以及维尔京群岛。数据涉及的主要专题内容包括：

1). EPA常规设施点位分布：反映了由EPA掌握的主要环境监测，处理，控制设施点的分布。这些点位数据主要取自于EPA的环境媒体及其它研究项目，详见本章第二部分。这些数据可用于多种分析，例如：人口和设施的比邻关系等。目前，这些数据用于EPA生态方面的研究和万维网上的各种GIS应用。

2)．主要河流分布：它包括了美国大陆上将近700,000公里的水系。这个数据来源于美国地质调查局1：250,000地形图上的所有水文要素，并在其基础上加工更正充实。它完成于1982年，已经多年被用于EPA和国家的鱼类，野生动物以及气象服务。

可用于环境资源的研究。

4).1994年鱼类受污染信息：它包括了化学污染情况，受污水域，受污染鱼种和数量，以及污染的开始与结止日期，持续时间等。它向公众揭示了一系列水域和鱼种的潜在污染危机。

5).美国气象站的气象数据：它总结了国家气象站的气象数据，其中包括：六个风速的综合频率，十六个风向，以及六个稳定性等级。这些要素取自于各气象站年报中每小时的记录。本数据用于EPA的图形研究模型系统。

6).1992年美国人口普查局的TIGER空间数据：包括道路，铁路，水文，点，线，面等。这些线划数据是根据1990年人口普查的表格数据和公开发表的边界由县或同等区级发布的。EPA的国家地理信息服务中心就这些数据处理成ARC/INFO格式。此类数据可用于作图，路线分析，人文地理分析，以及EPA设施布局与人口分布的关系分析。

7).1995年美国地质调查所地理名称点位:此数据所

含有的地理名称是以实际地物的经纬坐标来标识的。本数据通常用作地理底图。

8).1990年美国人口普查局的街区界限，街区中心点，街区组界限等：这些数据取自于1990年美国人口普查局的TIGER线划数据。此线划数据提供了1990年人口普查所用的普查区边界。

9).1:2,000,000比例尺的州际公路，水文，州界及县界：这些数据取自于1991年美国地质调查局1：200万数字线划图DLG数据。

10).美国超级基金项目点位：反映了1,223个国家优先考虑的综合环境污染点的分布区。

11).美国和墨西哥边界100公里范围内的大气污染监测点：美国和墨西哥边界100公里范围内从1983年起被列入环境关注区。本数据被美国－墨西哥边界项目所用。

12).1:250,000 GIRAS (Geographic Information Retrievaland Analysis System) 土地利用和覆盖信息:此数据是由美国地质调查所收集，由EPA转换为ARC/INFO数据。由于数据的采集经历了从70年代到80年代10多年的时间，数据详细程度差异比较大，土地利用分类采用Anderson的分类方法。本数据所反映的土地利用组合形态曾应用于与水质有关的环境评价,农业生产管理,以及其它环境影响评价分析。

此外还有各种关于大气，水文，土地环境状况的信息，这里不一一介绍。除了现存的数据以外，将来准备增加新的专题数据,其中包括:水文单元类型索引系统;国家公园边界；国家野生动物保护区；濒临灭绝物种分布等等。

二．环境数据系统(EnviroFacts)

环境数据系统储存了EPA的环境设施，项目及观测结果数据。环境数据系统储存的数据取自如下EPA的七个信息系统和数据库:

1). 综合环境治理，补偿，责任追究信息系统：这个项目是针对美国多年来积累的大量废弃物无人处理的问题由参议院于1980年设立的，是EPA超级基金(Superfund)的一部分。此项目的目的是彻底清除大型灾害性废弃物。这个项目积累了从1983年至今的大型灾害性废弃物评价，清除计划，基金投入，和所用技术等等信息。

2). 排污许可检控项目：美国所有的污水排放，不论是城市污水还是工业污水排放,都需要事先获得排放许可。排污许可检控项目是根据环境数据来检查和调整排放许可的上线以及排放许可的多少。

3). 资源保护信息系统：这个系统提供了灾害性废弃物信息，用来支持资源保护的措施。这个项目要求灾害性废弃物的制造者，运输者，储存者，处理者向州级环保部门提供他们的活动数据。州环保部门再将信息汇总到EPA大区和总部。所提供的数据内容有：废弃物的物主资料；废弃物的处理方法，废弃物处理的监视程度，废弃物处理的跟踪调查等等。

4). 有毒物质排放清单系统：这个系统所提供的数据包含了由制造厂每年向空气，水，土地释放的大约300种有毒物质。工业设施所提供的信息包括：工厂生产,处理,使用化学品的设施的位置；化学品的储备量；化学品的大概排放量;化学品就地再加工和再循环的能力；化学品运往它处进行处理，再加工，再处理的大概数量等。

5). 空气信息提取系统及大气点污染源信息系统：这个系统记录了几乎所有的美国点源污染点的资料,包括工厂类型，位置，污染物排放点和排放量，以及工厂内部生产过程和可能造成的污染。这些资料由空气信息提取系统和大气污染设施子系统(Aerometric Information RetrievalSystem/AIRS Facility Subsystem, AIRS/AFS)来储存，管理和提取。

6). 安全饮用水信息系统：这个系统所储存的安全饮用水信息包括公用水系统和违反EPA安全饮用水规定的记录。EPA以法律形式规定了安全饮用水的标准，水处理技术，水质监测等等以确保人民的健康。

7). 研究与治理基金的管理信息系统：这个系统存有所有EPA项目的基金申请，批准和管理信息。

此外，环境数据系统还支持三个横向关系数据库:(1)EPA设施检索系统,(2)主要化学物质综合数据库，以及(3)点位数据库。其中EPA设施检索系统从横向联系了多个数据库中的EPA设施。主要化学物质综合数据库为排污许可检控系统,综合环境响应及补偿系统和有毒物质排放清单系统提供了化学物质横向联系的标准。点位数据库以经纬度坐标形式描述了EPA设施的地理位置。

1. 数据库构成及数据储存

环境数据系统是应用ORACLE关系数据库建立的。它将上述EPA的七个各自独立的数据库和三个横向关联的数据库联在一起。只要进入这个系统就可以提取到上述十个数据库的任何信息。图1表示了各数据库间的关系。

2. 数据查询

EPA已经在万维网上建立了多种环境事实数据查询和制图功能。用户可以有选择地查询某个EPA项目数据库，也可以同时查询所有的EPA项目数据。查询方式有：按

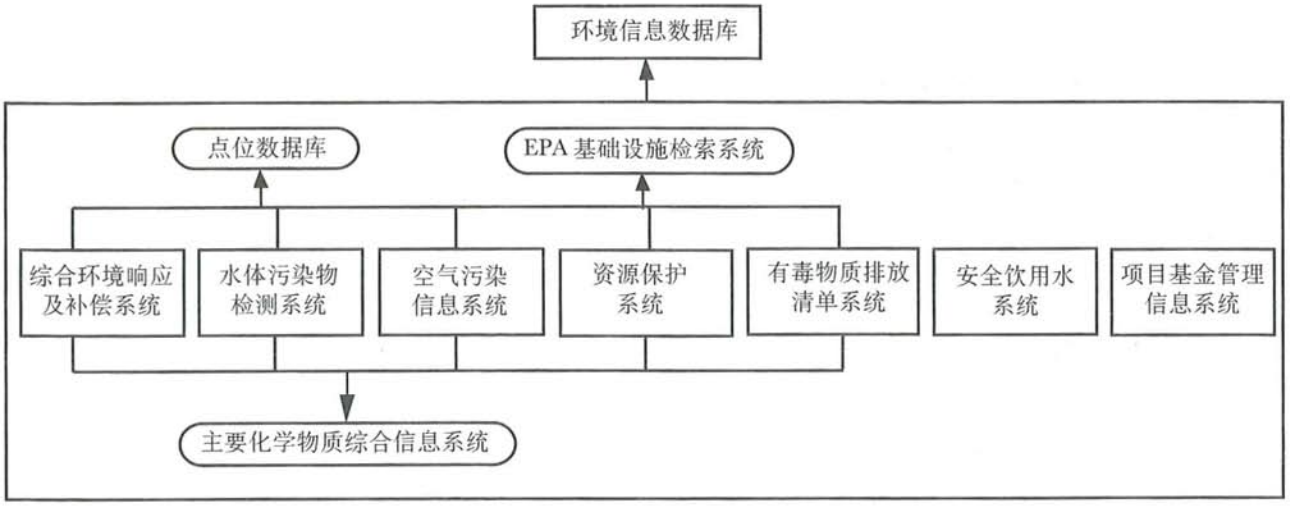


图 1. 环境信息数据系统数据结构

EPA 设施名称或代号，地理区域，标准工业分类代码，或按照数据库中的某个数据元素。数据查询的结果是一个列有符合用户要求的 EPA 设施清单。

3. 数据更新

环境数据系统基本上是按月更新。其中，研究与治理基金的管理信息系统是每两周更新一次。安全饮用水信息系统是每三个月更新一次。对于按月更新的 EPA 信息系统，今后环境数据系统将增加新的EPA信息子系统，其中包括：紧急事件通报系统和综合环境响应及补偿系统。

这个数据系统对 EPA 内部用户和公众免费开放，其目的是一方面使得公众可以接触到各种 EPA 的数据，另一方面是为EPA用户提供了万维网和其它ORACLE工具用以检索数据。

三． EPA 地理空间数据交换所(Spatial Data Clearinghouse)

国家地理空间数据交换所 EPA 交换点是国家空间数据基础构架的一部分。EPA交换点的运行是在技术服务中心(Enterprise Technology Services Division, ETSD)的支持下通过信息资源管理办公室(Office of Information Resources Management, OIRM) 的国家地理信息系统来完成的。这个交换点为公众开辟了一条通往 EPA 空间数据的渠道。

国家地理空间数据交换所的目的是让用户了解现有的数据，找到所需的数据，评价数据的可应用性，以及定购数据。EPA交换点存有50多种元数据，面向公众开放。它的万维网地址是：http://www.epa.gov/nsdi。

EPA 的元数据是通过 EPA 总部及其分部的 GIS 活动获取来的。这些元数据需要参照联邦地理数据委员会(Federal Geographic Data Committee, FGDC)标准进行检查和补充。由于这个查证过程比较复杂，EPA 的公关技术服务部 (Enterprise Technical Services Division)专门在北卡罗来纳的“研究三角园”设立了一个中心以从事元数据的检索，上载及传播．

EPA交换点根据需要每个月增加一些新的元数据．它们大多数按照EPA设施的编号进行检索，查询，和提取。这个交换点同时保存了没有进行检索处理的(non-indexed)的 EPA 空间数据源和其它地理空间数据源，并且建立了连接联邦，各州，大学，其它国家，商业，和其它国家空间数据基础设施(National Spatial Data Infrastructure, NSDI)信息站的通道。

四．空间信息系统与环境信息系统的结合—按需制图(Map on Demand)

按需制图功能应用于将环境数据库的数据按各种用户需求作图。这个制图系统采用了所有的 EPA 数据，其中包括 EPA 环境信息系统的数据，EPA 空间信息系统的数据和人口统计数据。因而用户可以有多种多样的制图选择。目前，按需制图的功能主要有：定点制图，流域制图，县域制图，邮电编码制图，EPA 设施密度制图，以及资料查询结果制图。

1). 定点制图功能是将用户按经纬度选择的地点上的人口统计信息显示出来。制图区域是由用户设定的半径大小所决定。用户可以按照环境数据系统中的查询表来选择经纬度。这个功能最初被 EPA 的十个大区分部用来帮助环境决策。表 2 和图 2 是这个功能的实例。

2). 流域制图功能是按用户要求显示人口统计和EPA信息系统的数据在流域中的分布。同时还以文字报告的形式给出 EPA 常规设施在水文单元中的信息。

表 2. 按需制图－定点制图申请表

地图主标题 [ ]
地图副标题 [ ]
制图区域

	度	分	秒
纬度	[ ]	[ ]	[ ]
经度	[ ]	[ ]	[ ]

制图区域半径或比例尺

图幅 14x11 平方英寸，制图比例尺取决于所选择的半径大小。
可选择的半径有:

[2 公里]
[5 公里]
[10 公里]
[10 公里和 2 公里（两张图）]
[5 公里和 2 公里（两张图）]

图幅 24x24 平方英寸，制图比例尺为 1:24000

选择所需的文字报告

EPA 常规基础设施信息

[水中污染物排放量的监测设施及数据]
[有毒物质排放的监测设施及数据]
[综合环境响应及补偿系统设施]
[资源保护系统设施]

人口，种族统计信息
公用水供给信息
敏感环境条件信息

选择背景图信息

[人口密度] [有色人口百分比] [贫困人口百分比]

选择一个或多个专题信息

公路和铁路
水网
湿地
国家公园及消闲地
EPA 授权的水流量检测设施
EPA 超级基金项目设施
资源保护及恢复项目的设施
EPA 有毒物质释放的检测设施
公用地下水供给
经纬网

地图输出格式

用于万维网上看的图形数据格式:

GIF 低分辨率（72dpi）
GIF 中分辨率（144dpi）
GIF 高分辨率（261dpi）
AcrobatReaderPDFPlotfile

用于非万维网的地图及打印文件格式:

Arc/InfoArcplot.gra.zipfile
GeneralColorPostscriptPlotFile.ps.zip
AdobeIllustratorPlotFile.ai.zip
HPGL2-8PenDraftmasterPlotFile.hpgl.zip
WindowsBitmapFiles.bit.zip
...

你的 e-mail 地址: [ ]

[发出作图要求] 或 [重新选择参数]

3). 县域制图功能是制作以县为单位的有关EPA管理和健康问题的图件及报告。例如，EPA常规设施，人口统计，和安全饮用水在用户指定县中的分布。
4). 邮电编码制图功能和县域制图功能相似。所不同的是，前者是以邮电编码为制图单元。

5). EPA 设施密度制图是编制以国家，州，县，及水文流域为单元的 EPA 常规设施图。用户可以从这类图上了解到 EPA 常规设施的集中分布，同时，也从所提供的文字报告中掌握这些设施的数量情况。

6). 查询结果制图是将环境事实数据系统查询结果制图。图上除了环境信息以外还显示了制图区域的人口统计信息，土地利用，土地覆盖，自然和文化特征。图上面状符合可表现为土地利用状况，人口密度，少数民族百分比，贫困人口百分比，人均收入。查询结果制图容许多层次地理数据制图。图型文件的格式有：GIF，TIFF，JPEG 等的压缩格式。

五．小结

本文对 EPA 环境与空间数据系统进行了概述。目前这些系统仍不够完善，EPA对这些系统及它们内部的各个部分仍在进行改进。比如，空间数据系统(ESDLS)和地理空间数据交换所(SpatialDataClearinghouse)还不能对公众完全开放。空间数据系统现只对EPA内部用户开放。公众可以查询地理空间数据交换所的元数据(Metadata)，但还不能通过它直接调用空间数据。另外，为了提高设施点位数据的精度，EPA 从 1990 年开始进行点位数据更正项目。这个项目的目的是为了确保数据的精度，格式的一致性和数据的完整性。它规定将精度保证在(25 米。因此，采集 EPA 设施位置的精确经纬度是这个项目的目标。数据采集的对象主要包括：所有 EPA 常规设施，EPA项目活动点和环境监测观测点。除了经纬度坐标以外，其它需要收集的相关数据有：经纬度坐标的测量方法，测量精度，地物实体，元数据比例尺等等。为了更加规范地记录所需数据，EPA特别制定了“方法，精度和描述”参照表来指导数据的正确表达。

参考文献

US Environmental Protection Agency, Envirofacts Database Description Document Version 2.0, SDC-0055-051-MA-5031, September 30, 1996.

US Environmental Protection Agency, ESDLS Version 2.0 Operations and Maintenance Manual, SDC-0055-065-DY-5039B, August 22, 1996.

US Environmental Protection Agency, Design Document for the Geospatial Data Component of ESDLS Version 2.0, SDC-0055-065-KF-5034, September 30, 1996.

US Environmental Protection Agency, Design Document for Locational Reference Tables in Envirofacts Version 2.1, SDC-0055-091-RP-6013, January 21, 1997.

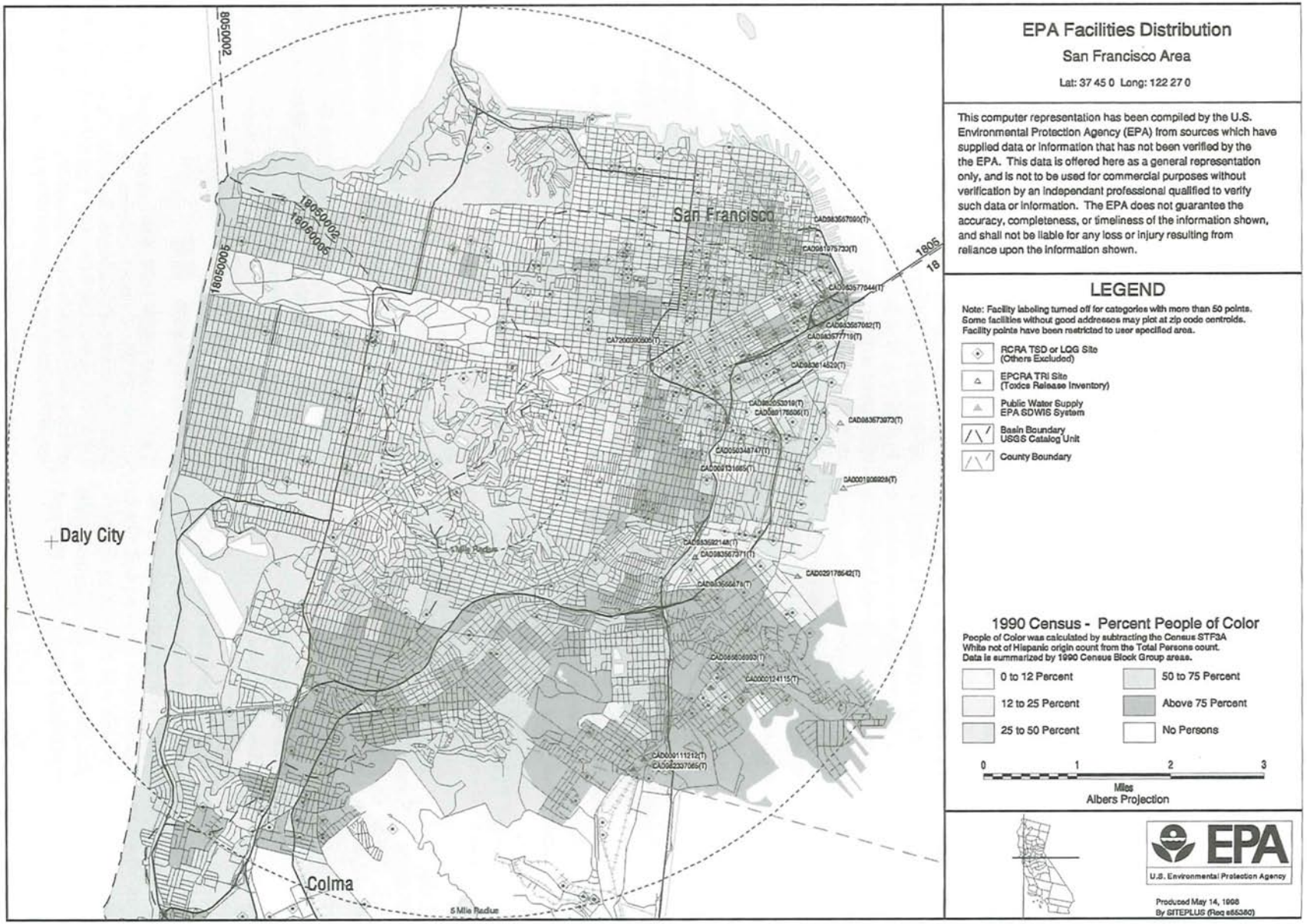


图 2. 旧金山市 EPA 设施分布图－按需制图功能的一个实例

Instruction to Contributors

1. Scope: The journal will publish original research on the theory, methods, development and applications of geographic information systems, remote sensing, global positioning systems, and cartography. Brief papers on hardware and software design and associated codes are welcome. GIS-related news, highlight articles and commercial advertisements will also be published.

2. Language: Either Chinese or English. Abstracts will be published in both Languages.

3. Manuscript Format: Use American Letter Sized or A4 white papers. Manuscripts must be double spaced and should normally not exceed 5000 words per manuscript in length. An abstract is limited in less than 150 words. Submission in digital form is welcome. List figure captions on a separate page and put figure number and author(s)' name(s) on the back of the figures. Color illustrations may be published with a charge to the author(s) depending on fund availability of CPGIS. SI units must be used. Citation of references in the text can be made either by a number index such as [2] or by using the author(s) followed by a number index, e.g., Chen and Wang [2]. A reference should include items in the following order:
Number Index, Author Name, Year of Publication, Title of Article or Book, Journal Name or Book Publisher, Volume, Issue, Starting and Ending Pages.

4. Manuscript Submission: 4 copies must be submitted to

Dr. P. Gong
Department of Environmental
Science, Policy, and Management
151 Hilgard Hall
University of California
Berkeley, Ca 94720-3110 USA
Fax: 510 643-5438 Tel: 510 642-5170
Email: gong@nature.berkeley.edu

5. Review, Proofs and Offprints: Technical papers will be submitted to referees whose names will be kept confidential. The Editors reserve the right on whether a paper should be accepted. Once a paper is accepted, proofs will be sent to the corresponding author for checking. 25 offprints will be sent to the corresponding author when the paper is published.

6. Advertisements: Please contact the editor.

征稿启事

《地理信息科学》是由中国海外地理信息系统协会(CPGIS)主办的一个专业性刊物。其宗旨为发表有关地理信息系统、遥感、全球定位系统及制图学方面的理论、方法、发展与应用等方面的开拓性研究论文，以作为联系海内外中国地理信息系统学者的纽带，作为了望世界高科技和交流地理信息系统等国际重要科研成果的窗口。诚挚地希望广大学者给予支持，欢迎一切与本刊宗旨相附的论文与图片的来稿。亦欢迎登载与上述领域有关的消息，产品及商业广告。

本刊文章用中文或英文两种语言发表，摘要将用中英文两种文字出版。文稿长度通常不超过5,000单字。摘要限在150 字左右。原稿请用标准信笺或方格稿纸隔行打印或书写工整，并欢迎软盘投稿。

希望作者将论文中的图题和标题另页列出。另外在图的背后注上作者的姓名及图号。彩图亦可发表，但须根据 CPGIS 的经济状况向作者收取一定工本费。文中计量单位采用国际标准单位制。正文中的参考引文可用编号如[2] 或用作者姓名跟编号如Chen 和Wang [2] 等形式。参考文献要求按照如下顺序列出:
编号，作者姓名（西文姓在前，加逗号），发表年份，引文标题或书名，期刊名称（如是图书，则列出出版社），卷、期及起止页码等。

来稿请详细写明作者姓名、单位、职务、职称、地址、邮编、电话或传真。

学术性论文将会送给审稿人评审。审稿人姓名将予以保密。稿件一经采用，排版后的清样将会寄给作者本人校对。出版后即奉送作者 25 份论文单行本。

稿件要求一式四份投给如下地址:

Dr. P. Gong
Department of Environmental
Science, Policy, and Management
151 Hilgard Hall
University of California
Berkeley, Ca 94720-3110 USA
Fax: 510 643-5438 Tel: 510 642-5170
Email: gong@nature.berkeley.edu

需刊登商业广告者，请与上述编辑联系。

Subscription 征订

This journal (ISSN 1082-4006) is published semiannually by CPGIS. Contact address: CPGIS, 151 Hilgard Hall, University of California, Berkeley, Ca 94720-3110, USA, Telephone: 510 642-1351, Fax: 510 643-5098. The cost for each issue is $10. Postage and handling is $2.5 for North America and $5 for other regions.

本杂志(ISSN 1082-4006)由CPGIS出版。联络地址: CPGIS, 151 Hilgard Hall, University of California, Berkeley, Ca 94720-3110, USA, 电话:510 642-1351,传真:510 643-5098 暂定每半年一期，六月和十二月出刊。每期暂定价10美元。每期邮寄费北美地区为2.5美元，其它地区 5 美元。中国大陆定价与邮费请与 CPGIS 联系。

GEOGRAPHIC INFORMATION SCIENCES《地 理 信 息 科 学》

Table of Contents, 1997　　1997 年总目录

Number 1-2　第一至二期