School of
Civil Engineering and Geosciences
University of Newcastle upon Tyne

FROM GEOSPATIAL DATA CAPTURE TO THE DELIVERY OF GIS-READY INFORMATION: IMPROVED MANAGEMENT WITHIN A GIS ENVIRONMENT

Zainal A. Majeed

Thesis Submitted for the Degree of

Doctor of Philosophy

UNIVERSITY OF
NEWCASTLE

School of Civil Engineering and Geosciences
University of Newcastle upon Tyne

2006

ABSTRACT

This thesis presents the research undertaken to investigate how geospatial data handling techniques and technology can be potentially used to enhance the existing management of entire survey datasets from their captured stage to a GIS-ready state and the delivery of this to the user. Discovery of the current systems for managing survey data and information in the Survey and Mapping Department Malaysia (JUPEM) has been presented. In addition, the surveying practice and processes carried out have been examined, especially the different type of data and information existed within the raw data capture right through to the production of GIS-ready information. The current GIS technology and techniques for managing geospatial data have been inspected to gain an in-depth understanding of them. Geospatial object as an approach to model reality of the world has been discovered and used to model the raw, processed, the GIS-ready information. To implement the management, a prototype Database Management System (DBMS) has been implemented, and a trial data population and processing steps have been carried out. An enhancement of the management of the datasets from geospatial data capture to the GIS-ready information has been demonstrated. To deliver online the final product, demonstration of available methods were illustrated, and then contrasted. A range of datasets around Malaysian context were used in the research.

The investigation revealed that raw, processed and GIS-ready information can be successfully modelled as object in an object-relational spatial database. Using inherent GIS tools, survey datasets management and processing steps within the same system are evidently achieved in a prototype implemented DBMS. An improved management showing the capability of 'drill-down search' and 'two-way traceability' to access and search spatial and non-spatial information in the system is effectively illustrated. Demonstration of the vendor specific and open source technology for the GIS-ready information delivery leads to the comparison between them. The thesis concludes by recognising that a management for raw captured data, processed set of data and GIS-ready information, and the delivery of this, within GIS environment is possible. The inherent GIS tools and DBMS have presented a single-view system for geospatial data management providing superior interfaces that are easy to learn and use, and users are able to specify and perform the desired tasks efficiently. Delivery of

data has some constraints that need to be considered before embarking into either vendor specific application or open source technology. In JUPEM, time and cost can be reduced by applying and implementing the suggested GIS application for cadastral and topographic surveys right up to the creation of GIS-ready information, as detailed in the thesis. The research also finds that the in-depth understanding and experience, practically and theoretically, of all aspects of current GIS technologies and techniques gained through this research has achieved an overarching inspiration: equalisation of a high level of awareness and ability of staff in handling GIS project development within currently developing countries with those in the developed countries, and within the national survey and mapping department with those of other government departments and commercial GIS contractors.

ACKNOWLEDGEMENTS

Firstly, I would like to thank God, Allah the Almighty for giving me courage, patience and ability to be able to pursue this research study and write this thesis. I would also like to express my gratitude to the Public Service Department of Malaysia and the Department of Survey and Mapping Malaysia as well as the University of Newcastle upon Tyne who provided financial support, international research scholarship, study leave, facilities, transportation and assistance during my stay and study in the UK.

I would especially like to express my greatest appreciation to my supervisor, Professor David Parker for his patience, thoughtful guidance and support, assistance and editing my manuscript. He has been my mentor and understanding teacher in the most truthful sense. My thankfulness is also devoted to Dr David Fairbairn, Phil James and Martin Robertson for the expert advice, technical advice and enthusiasm while trying to overcome the hurdles towards the final deal. I would like to thank Mr Sanphet Chunithipaisan for giving the opportunity to cooperate in his research thesis which was partially applicable to my thesis.

Many people from the Geomatics Department in the School of Civil Engineering, where I worked my PhD, lecturers, postgraduate students and the staff from the school office especially Lindsey Early and Lynn Patterson, all gave generously and freely of their time in assistance, friendly atmosphere, advice and supportive environment. The aid given has given encouragement to work on the thesis and left me with many memories.

Most of all to my dearest wife Hayaton Haji Shaari, my three children, Syazwanis, Syafiq and Ameera, I am truly indebted to all of you for being tolerant, supportive and patient throughout my studies. They all sacrificed a lot of valuable times without me in the family scene, so that I can accomplish my research. I am very grateful to my lovely wife for her patience, support and understanding. We have overcome many ordeals with tolerance and endurance while living in the UK. Without her love and support, I would have never had the power to keep on the confidence of finishing my PhD study. Your unconditional devotion and enliven spirit gave me the enthusiastic persistence to strive ahead successfully. To my mother, your prayer has done it all and for so long without having your son near you while you are in difficulties back home in Malaysia, I am truly indebted and will always pray for your health and well-being.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Different field data capture techniques for the development of geospatial-quality end products such as survey plans, maps, air photographs, satellite images and GIS-ready information require the collection and processing of large volumes of data. Efficient management and maintenance of the full life cycle of the datasets from raw data to the processing stage right through to the applicable end products may provide complete traceability of the processing steps and the origin of the data. An integrated system for managing a survey package from raw survey data to processed dataset and all the way through to the final product can allow efficient planning, data acquisition and processing. It would provide the ability to concurrently collect large volumes of data and to fully track the processing steps, all the way through to the final data products. Most data collecting organisation may have a few management systems in place for captured data and processed data, but they are in different locations, not accessible to all who need this information and do not allow a single view of all the datasets and processing steps. A national survey and mapping department is a geospatial data collecting organisation which is largely involved in collecting raw data, processing it and producing an end product from it. Almost all end products of this organisation are in a GIS-ready information state. To deal with these kinds of management issues, this research has developed an integrated management system within the Geographic Information System (GIS) environment to manage the level of datasets and processing steps from raw and processed data up to the point where they become GIS-ready.

A typical geospatial data collecting organisation is the Survey and Mapping Department of Malaysia (JUPEM). This organisation is the sole government organisation that produces and handles geospatial data and information for land administration and mapping purposes, and at the same time supplies them to the public and other government organisations. There is a vast amount of geospatial data and information, raw and processed, being held in JUPEM in hardcopy and digital forms, situated in various sections and units which deal with different types of geospatial data, such as air survey data, topographic data and cadastral data. There are three types of surveys carried out in JUPEM: topographic survey, cadastral survey and engineering survey. Engineering surveys are less significant and rare because they are carried out on a demand basis by other government organisations. Topographic and cadastral surveys form the core business of JUPEM for land administration, as well as localised and national mapping purposes.

GIS technology has existed in the modern world for decades and has provided many solutions to management issues for spatial and non-spatial information. It is a combination of a computer application and database system used for storing, manipulating and managing spatial and non-spatial data. Its technology combines normal database operation such as query and statistical analysis, manoeuvring a unique graphical presentation in the form of a map and/or spatial object model. In this way, the user gets a clearer visualisation of the information and not a numerical expression. It has the power of managing geospatial datasets, creating maps, manipulating information, viewing current scenarios, solving complex problems, promoting new and useful ideas and developing effective solutions that cannot be realised using other techniques. GIS technology has been used by multi-disciplinary professions as quite a superior data and information management tool in such areas as engineering (Yunus, 1997), medical (Casper et al., 2000), project organisation (Huxhold and Levinsohn, 1995; McConnell and White, 1999) and surveying (Jericke, 2002). With the incorporation of a spatial dimension as an element for deriving decision support management, GIS technology has provided management techniques for decision makers (Kennan, 1998). Truly, GIS has become known as an excellent tool, application and storage for a variety of usages. Essentially GIS and databases cannot be separated in accomplishing a powerful system for geospatial data and information management.

There are two main classes of geographic data representation: objects and fields, which demonstrate discrete and continuous representation of reality respectively. These representations have been increasingly recognised as two different approaches for conceptualising and modelling geographical phenomena (Worboys and Duckham, 2004; Burrough and McDonnell, 2000; Zeiler, 1999). Many phenomena in geographic reality are readily professed as objects. Land parcels, buildings, lakes, boundaries and roads are a few examples. In this viewpoint, these identifiable objects can be seen as occupying a space in a container; fundamentally each has identity, spatial embedding and attributes (Cova and Goodchild, 2002). Object-based data representation contributes to the world of GIS by its direct insight and perception of the human view of the Earth's features or phenomena being modelled. The field model can be demonstrated by a continuous surface with a finite number of variables each measurable at any point on the Earth's surface and considered to be position dependent (Longley et al., 2002). Since the field-based view of data recognises surfaces with continuity as the principal element, modelling of entities by an object-based view is used for this research because object entity is not everywhere defined in surface. It represents phenomena at specific associated positional information and with descriptive single or multiple attributes (Laurini and Thompson, 1995). Therefore the field-based view is not explored extensively: instead this research focuses on the application of object theory and technology to handle geographic phenomena regarding surveys, processed datasets and GIS-ready information. Further explanation for this will be stated in the relevant chapter.

After either the object-based or field-based model is chosen for how data can be modelled and utilised for GIS purposes, the choices of traditional file processing or the database model can be designed, which eventually allows the data to be queried and displayed in maps, tables and reports. In this case, the database approach is explicitly the choice because of its many advantages. Importantly, this approach allows maintenance of a single repository of data that is defined once and then permits accessibility by various users (Rodriguez, 2004). Detail of the significances of this approach is discussed in Chapter 4. In most GIS, there are generally four types of database designs being implemented; namely hybrid model, relational model, object model and object-relational model. The relational model is contradictory to the hybrid model for the reason that the former implements an integrated system which stores all data and attributes within the same database with separate relational tables, whereas the latter has a separate set of files and separated relational database system (RDBMS) (Demers, 2000; Worboys and Duckham, 2004; Adam and Gangopadhyay, 1997). The object-oriented approach mainly has the ability to abstract or summarise, which makes it appealing for software engineering (Leung et al., 1999; Blaha and Premerlani, 1998; Oosterom and Vandebos, 1990). The power of incorporating both features of the object model and relational model enables the implementation of objects in a database and at the same time maintains the full functionality of the relational model (Elmasri and Navathe, 2004; Kriegel et al., 2004). The huge amount of geospatial data in different formats that exists nowadays forces management strategies to be planned incorporating both database models. This strategy becomes part of the objective of the research, which mainly focuses on modelling them into various objects and linking or relating them to each other.

GIS is an important organisational and decision tool that should be most beneficial and fully utilised together with the Internet. The Internet has introduced World Wide Web (WWW) as the road to transport the load, no matter what the backend systems are (Stoimenov and Djordjevic-Kajan, 2005). With the existence of geospatial interoperability standards, the movement and transfer of data within GIS and among geospatial datasets, especially when the geospatial data has to be transformed into fully GIS-ready information, can be achieved. Geospatial interoperability has been a matter of debate among many academics and provides influential issues among people of the government in particular (Jerome, 2005). The distribution of geospatial data through Internet services can benefit many enterprises and communities in the current information and communication age. The resultant GIS-ready information in the geospatial management in the research is presented in a way that can be disseminated and shared amongst the geospatial users and providers.

This thesis describes research, which investigates and addresses the use of GIS technology and the current design of database models that are feasible for management of various geospatial data captured and processed, up to the point where they are ready to perform as GIS information in a single portal accessibility. As widely as possible, existing spatial data representation and models, current appropriate database technology, relevant GIS tools and technologies, and Internet GIS capability are used to develop solutions to achieve this aim.

The sections that follow discuss in more detail the background and motivations, aim and objectives for carrying out this project. The research methodology is described and the thesis structure is set out.

1.2 Background and Motivations

Recently, it has been proclaimed that information is the most important commodity in the world economy (Gregorius, 2005). Almost 80% of all information has a geospatial component. Geospatial data is an indispensable 'raw material' for the production of maps, land parcel and engineering survey plans, and business management (Longley et al., 2002). There is a crucial need to evaluate and create strategies for geospatial data storage and development (Crompvoets et al., 2004). The modern geospatial data capture (surveying) for the production of parcel plans and maps requires the collection and processing of large volumes of data while maintaining complete traceability of the processing steps and the origin of the captured and processed data.

Progress in geospatial data capture and processing technologies, such as remote sensing, GPS, field-to-finish total station survey, light detection and ranging (LIDAR), and laser scanning, raises a sequence of data management challenges for a multi-disciplinary profession (Thurston et al., 2003). Decades ago, geospatial data was point data and in small quantity but this has been substituted by many gigabytes of data. Such data may be captured and generated daily or even hourly during survey or digitisation and scanning of hardcopy map document. Due to the vast amounts of data now in reality, the issue of data storage, management and access must be tackled. Moreover, data collection and data conversion from hardcopy to digital form are costly. Information about data sources, collection and conversion methods used are nowadays significant for historical evaluation and analysis. This data and information is actually one of the exclusive sources for many types of analysis, decision making, development, organisational geospatial management. However, such information is sometime located in various 'islands' of sources and everywhere in a traditional file processing system (Figure 1.1). It would seem logical that there is a growing need for efficient and well-defined data management systems to store, manage and distribute tremendously large and multiple-located datasets, in a single view by multiple users.

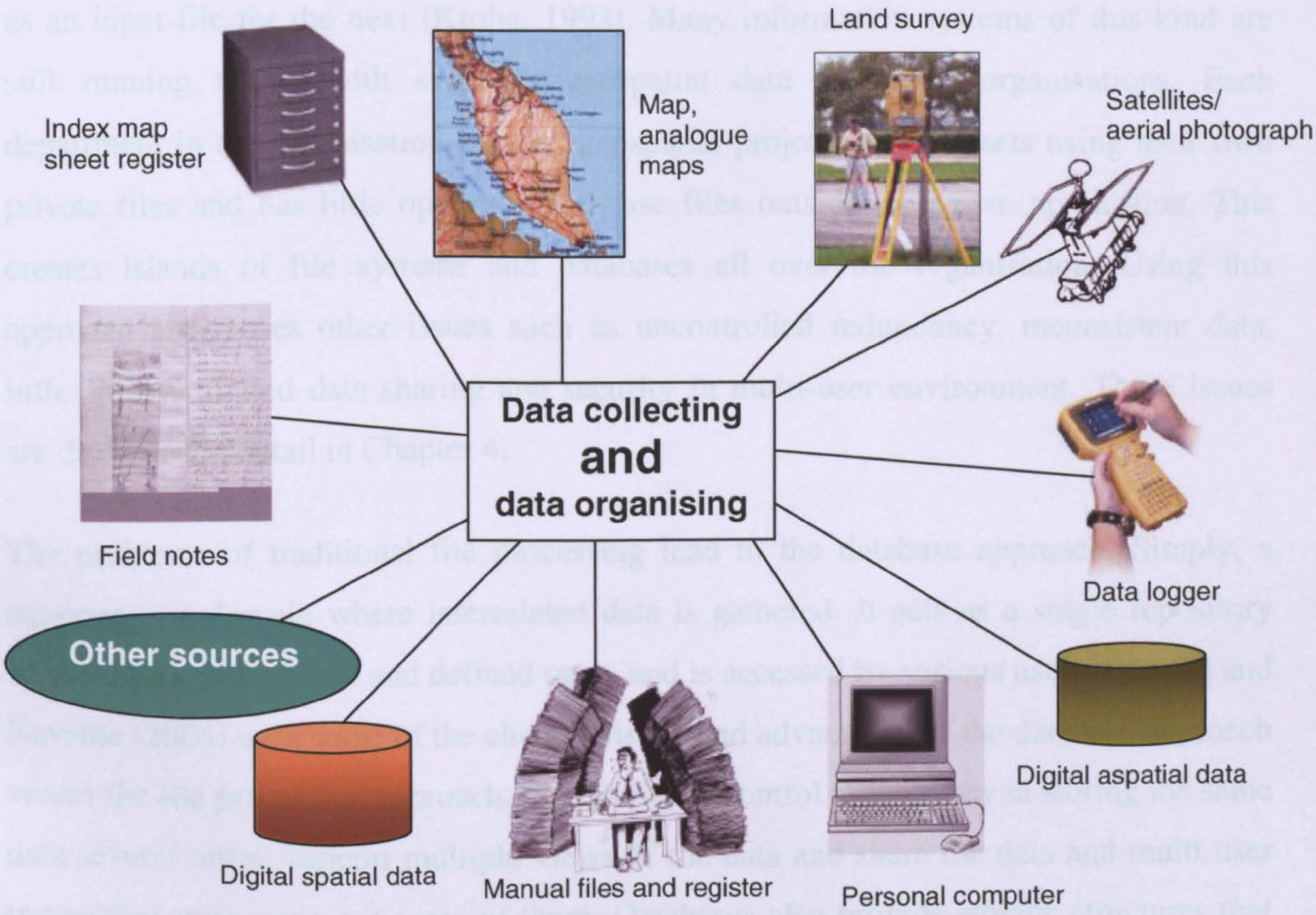


Figure 1.1: Numerous and various sources of data

The need for concurrent geospatial visualisation for viewing all activities and items of geospatial capture, processed data and GIS-ready information is an important issue. It could be crucial to maintaining historical information about data capture or survey especially when needing to validate or resurvey a new site, or reference a survey status. A portal that allows query of previous geospatial data capture can save time and laborious exploration. Similarly, paperless storage, adequate order and proper digital filing of information regarding raw survey data can be very informative and beneficial during the processing and creating of GIS-ready information. Eventually the management and linking of all geospatial data capture, the processed datasets and GIS-ready information will become critical, especially for the sake of data integrity and usefulness.

A typical approach in data collection and storage used in some computing environments is a traditional file processing system. This kind of information system is built from a set of files and application programs in such a way that the output file of one process is used as an input file for the next (Kroha, 1993). Many information systems of this kind are still running today, with some in geospatial data collecting organisations. Each department in the organisation handles geospatial projects and datasets using their own private files and has little opportunity to use files outside their own application. This creates islands of file systems and databases all over the organisation. Using this approach also poses other issues such as uncontrolled redundancy, inconsistent data, inflexibility, limited data sharing and security in multi-user environment. These issues are discussed in detail in Chapter 4.

The problems of traditional file processing lead to the database approach. Simply, a database is a domain where interrelated data is gathered. It acts as a single repository where data is maintained and defined once, and is accessed by various users. Elmasri and Navathe (2004) state some of the characteristics and advantages of the database approach versus the file processing approach. The ability to control redundancy in storing the same data several times, support multiple views of the data and share the data and multi user transaction processing are some of them. Databases also provide storage structures that enable queries and updates to be executed efficiently. Databases include the capability to represent a variety of complex relationships among the data as well as to retrieve and update related data easily and efficiently. The design also allows the database to be viewed in its entirety so that interaction and linkages between elements can be defined and evaluated. Further aspects of database approach are addressed in a Chapter 4.

A well-known and popular database model is the relational model in which the database contains data and relations organised in tables. These tables are controlled by RDBMS. The relational model is based on a well-established mathematical foundation. The user makes relations between various tables by identifying a common field in two tables. An advantage of the relational structure is its greater flexibility and thus the model is used by nearly all GIS systems when compared to the other two models. However, as the content of the world becomes more complex, real world objects, operation and attributes become more difficult to model. In addition, the relational model's mathematical foundation has limitations in most cases (Raza and Kainz, 1999; Egenhofer and Frank, 1992; Hughes, 1991). In terms of data quantity and complexity, the relational model offers a lack of modelling power and performance (Kroha, 1993). To model objects and behaviour, object-oriented approaches have been applied for data modelling and implementations. Object-oriented database management systems (OODBMS) provide higher performance management of objects, and enable better management of the complex interrelationships between objects (Loomis, 1995). Object technology is of great interest in the software and application world, because it promises to help solve the real complex problems (Clementini and Di Felice, 2001).

Relational models have traditionally been used in GIS systems for data storage. Although this model has a limitation in capturing object entity together with behaviour and attributes, its technology is mathematically reputable and more commercially used in the database realm than the object-oriented model. Since they are both equally superb models, their concepts have been combined in recent developments into what is called the object-relational model. This approach extends the relational model and integrates it with object-oriented concepts (Worboys and Duckham, 2004; Stonebraker et al., 1999; Premerlani et al., 1990). The combined approach provides the full power of the relational technology and at the same time provides the ability to capture an object and its behaviour within the object-oriented concept.

The need to present and distribute much heterogeneous geospatial data to a wide variety of users is a difficult challenge faced by many data providers. Government- oriented data providers especially are coping with the call to supply information and services to a wide range of data users (Wilmersdorf, 2003; Greenwood, 2002; Gauna and Sozza, 1999). One way to achieve this is by using a Web browser. The implementation of e-Government has been a significant issue for a timely development of the economy and infrastructure. Using the WWW to make e-Government a reality is a challenging task, but it helps various users' communities to understand the quality and history of the data they are receiving. There are many benefits of geospatial objects being active in a repository for internal and external access. Geospatial data management and distribution using graphic and non-graphic interfaces enables the existence of intelligent data objects for variety of purposes. At least the other government organisations specialising in agriculture, transport, drainage and rivers, public work or local authorities can access the spatial and aspatial data objects developed by the Survey and Mapping Department of Malaysia for example, as the main geospatial data provider.

JUPEM is a classic organisation that needs a data storage and management techniques to make its products more useful and valuable, and accessible by multiple users. The multiple data and information that exists in JUPEM can be pictured as in Figure 1.1. JUPEM is mainly focused on administration and operation of survey and mapping for the production of geodetic data, cadastral data and mapping data. JUPEM currently archives about 2.5 Terabytes of digital data, which is available for government, business, public and individuals' consumption. To meet the era of e-Government, the process of converting existing hardcopy maps, aerial photography, cadastral plans and index map sheets is currently the most important core business of JUPEM. Survey data and processed data are located in 13 states in the country and some are in separate locations in the same building in the JUPEM headquarters. JUPEM needs a single portal management that can be accessed by the staff who need them (Figure 1.2).

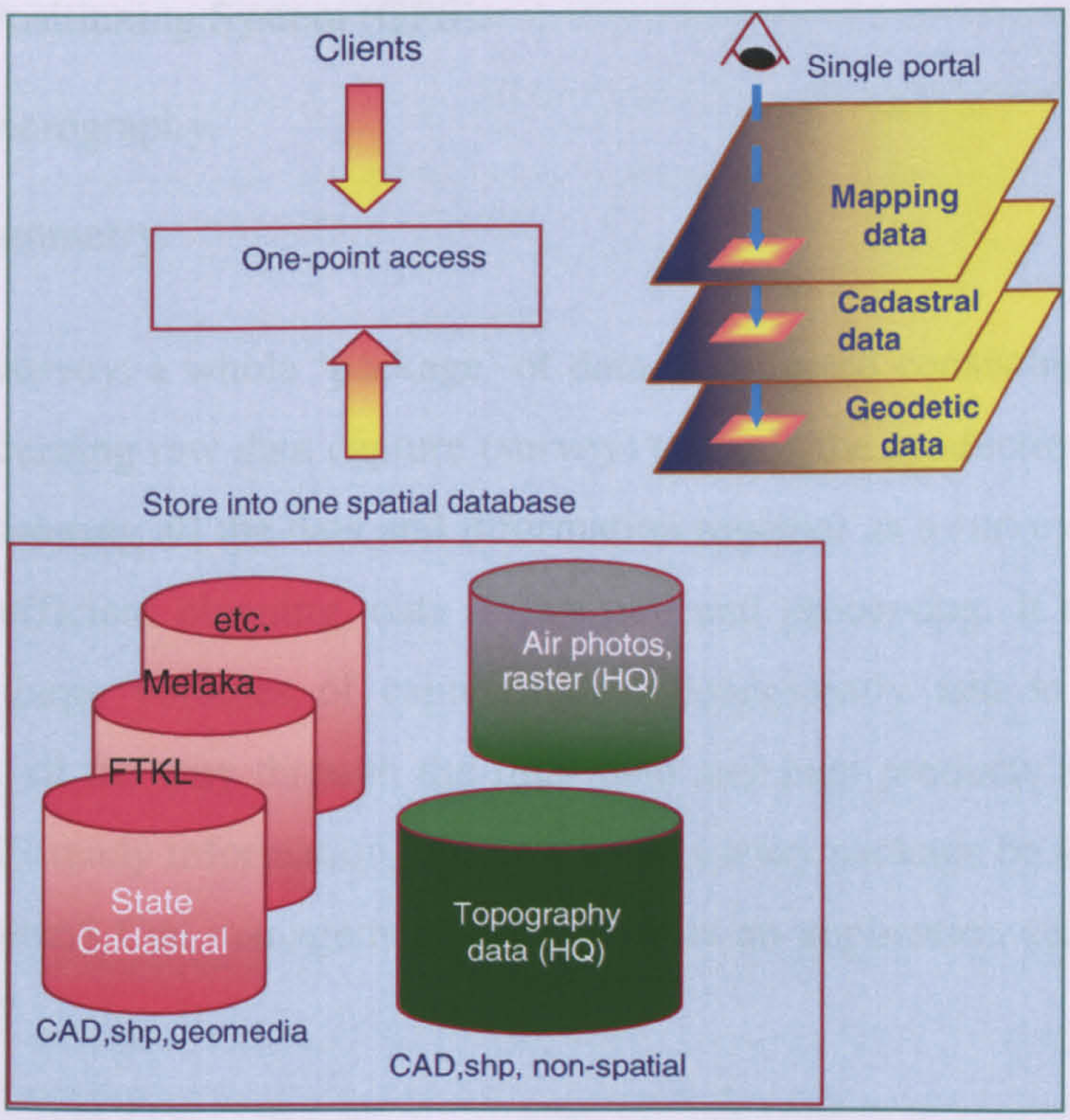


Figure 1.2: The vision of JUPEM and this research

JUPEM is the solitary organisation in Malaysia that provides high quality survey and mapping products and services to the government, business, public and individuals for the purpose of national development, security and defence. JUPEM has been acknowledged to be a major 'building block' for the National Spatial Data Infrastructure

(NSDI) of Malaysia (Tong, 2002). Therefore it plays a vital role in providing huge amounts of geospatial data to the Clearinghouse of the NSDI as well as other government organisations and the public.

As mentioned earlier, topographic survey and cadastral survey are the main type of surveys carried out in JUPEM. These surveys are accomplished using various techniques:

- Conventional theodolite and electronic distance measurement devices (EDM).
- Total station which is a combination of electronic transit and EDM.
- Geodetic survey comprising of gravimetric measurement, precise levelling and tide gauge reading.
- Global Positioning System (GPS).
- Aerial photography.
- Photogrammetry.

For a complete survey, a whole 'package' of data is collected consisting of all data and information concerning raw data capture (survey) towards the production of a processed set of data. Considering all the data and information together as a survey package can be very useful for efficient planning, data acquisition and processing. It may provide the ability to trace large volumes of captured data concurrently and to fully track the processing steps all the way through the final plan and map products as well as to the production of GIS-ready information. So how can a survey package be implemented in a management system? Can management using GIS be an application suitable to address the above issues?

JUPEM has several management systems for captured and processed geospatial data but they are in different locations and can only be accessed by several people from individual sections and units. In order to develop and maintain efficient and better storage and develop this collected and processed data, it is essential to have one management system. It would be ideal to integrate them into one single view that can be used by multiple users. A system can possibly be created that allows all the information to become more accessible to all who need it. It is assumed that by having this ideal system, better management of the operation in JUPEM could be achieved on a day-to-day basis. Therefore an adaptable way of using the management system that can be adopted by many users at different levels of operation should be achieved.

JUPEM implements surveys and data processing whereby all raw capture data needs to be passed through many stages to create maps and GIS-ready data. A perfect method through application systems may therefore be potentially created that makes all the additional processed data that was produced along the processing steps manageable in the same system as the raw data. At the same time, this application may permit some actual processing steps to be performed within the same management system.

The raw and processed datasets need to be accessed for potential usage by clients who are internal to the organisation, other government departments and the general public. In today's Internet era, information can be transported in a various way. Data sharing and access is an important part of geospatial data management. Therefore it is the intention that this research be carried out to find how the GIS-ready data can best be delivered to a wide range of users.

In a national survey and mapping department, it is vital that new technologies and techniques of managing geospatial data be in the same level of understanding with other government departments, other private licensed surveyors, and the commercial IT and GIS contractors. With this leverage, as an important geospatial data collecting organisation, JUPEM will be able to drive towards the bright future of the raw and processed dataset and may become an acceptable geospatial data and GIS-ready information providers in the whole nation and probably in the world. With the designed system, can this be accomplished? Will the system researched be able to equalise the ability of staff in handling GIS project development, operation and data delivery within the currently developing countries with those in the currently developed countries?

The background and motivations explained above thus pose research questions such as:

- Is it ideal to have one system for the management of geospatial data capture and processed data towards the production of GIS-ready information?

- Would this system be within a GIS environment?
- Can the same system be used to manage all the additional processed data together with the raw data?
- Can any of the actual processes be performed within the geospatial data management system?
- How best can GIS-ready information be delivered to clients?
- Can we get into a position whereby there is an equal level of understanding of GIS techniques and technologies within the JUPEM compared to that of other National Survey and Mapping Departments, other government departments and the commercial GIS contractors?

1.3 Research Aim and Objectives

The background and motivations explained in the previous section led to this research addressing the following research aims and objectives.

The aim of this research is to investigate how geospatial data handling techniques and technologies may potentially be used to better manage complete survey datasets from their raw stage to a GIS-ready state and the delivery of this to the users. Throughout the research the main objectives can be stated as follows:

1. To discover, in national survey and mapping departments, current systems for managing survey datasets from their raw stage to their output stage.
2. To gain an in-depth understanding of current GIS technologies and techniques for managing geospatial data right through to the production of GIS-ready information.
3. To model as geospatial objects the raw, processed and GIS-ready information.
4. To implement a prototype database management system (DBMS) to the specified data model.
5. To demonstrate, with trial data, population of the implementation of the DBMS.

6. To demonstrate improved management of the survey datasets using the developed GIS capability.
7. To demonstrate the use of inherent GIS tools to assist with the data processing.
8. To demonstrate and contrast the use of vendor specific GIS and open source technologies Internet solution for the delivery of GIS-ready information to the clients.
9. To report on the achievement of the research describing the potential benefits and pitfalls of implementing a GIS application for the full life cycle management of the survey datasets.
10. Propose future investigation to further the research.

1.4 Research Methodology

This research is conducted with the principle that observable spatial fact, detail and resulting 'artefact' such as a series of data and information produced through the geospatial data capture to the GIS-ready information can be modelled and managed as objects in the real world. To achieve this, the core concept of object orientation and database modelling will be examined and utilised for the development of geospatial data models. In this respect, the research approach is therefore to use object-orientation concepts to logically identify the various classes (objects) at various stages of the modelling, attributes and operations of these objects, relations between these objects and consistency constraints (e.g. topology) on these objects. In the case of database modelling the structuring and storing of data for efficient and effective management, query and access of the framework are affected. Possibly, structuring these objects using the power of relational modelling techniques could be carried out using geospatial data handling technologies available.

A sample of a wide range of original raw data was processed through to a GIS-ready state so as to gain detailed experience of the processes involved and the wide range of intermediate information and data generated. These datasets were acquired from various sections in JUPEM. These sections include a topographic map production section, a cadastral survey section, an aerial photographic section and a geodetic survey section of the JUPEM.

A current geospatial data handling and modelling technology is chosen to create an intelligent and organised management of the geospatial data. This includes the storage, production and access of a GIS-ready feature object. Steps for the design of conceptual, logical and physical data models are introduced in this case. A Unified Modelling Language (UML) approach is used to create the database model, which is then converted to the schema for the database design. ArcGIS Desktop, which uses object-relational software technology from Environmental System and Research Institute (ESRI), is currently one of the popular geospatial data handling technologies available. It is used here as an example for storage, access and query, and to envision its capability and potential to address the aim and objectives of the research. Processing steps and population of datasets are carried out for the raw, processed and GIS-ready information set of data.

A vendor specific application for geospatial data delivery over the Internet is tried in the research. In addition, Geographic Markup Language (GML) technology and Open Source Technology have been used to enable online data delivery, sharing and query. A prototype of GML feature schema was implemented to describe a simple feature used to represent the real world phenomena. This was carried out in conjunction with Chunithipaisan et al. (2003), in which 50% of the work was invested by the author of this research. In this project, he created and organised the GIS-ready data out of JUPEM's raw and processed data, assisted in the creation of Spatial Relational Database (SRDB) and tested the data dissemination. The two data delivery applications are contrasted in the research.

The research is not intended to develop a fully-functional geospatial data management and access system; rather it focuses on issues and possibilities of structuring all the geospatial objects within geospatial data capture, processed data and GIS-ready information.

1.5 Thesis Structure

This thesis is organised into seven chapters including this introduction. Chapter One has explained an overview of the thesis with some background and motivations of the research, the aim and objectives as well as the research methodology. The rest of the chapters are as follows:

- **Chapter Two** presents the literature review entailing some issues concerning geospatial data management from raw to processed stage right through to the GIS-ready information. Geospatial data and technologies are discussed concerning the collection and processing of geospatial data. Modelling by object as a significant core matter in GIS technology is described. Surveying as part of geospatial data capture is reviewed in broad terms. JUPEM, as a geospatial data collecting organisation, is examined in term of surveys carried out, how it is done and the types of survey package and information within it. The gist and development of GIS is presented, encompassing concepts, theoretical views of models, technologies, research and development with particular relation to geospatial data capture and processed data, and the benefits of management by GIS. Reviews of management practice of geospatial data to the creation of GIS-ready information in three different government organisations, namely Ordnance Survey (OS) of Great Britain, United States Geological Survey (USGS) and Survey and Mapping of Department Malaysia (JUPEM) are outlined. GIS interoperability is reviewed here.
- **Chapter Three** covers the technological concepts of object modelling, object-oriented modelling and UML. These technologies serve as a background to design the model and database.
- **Chapter Four** details the existing working practices and context within JUPEM and its impact to this research. The data modelling and data management approaches is then discussed. The geographic database, a type of DBMS, and its elements are discussed. Datasets existing and used in JUPEM are examined and discussed to suit with the data models and geographic database. How GIS-ready

information could be delivered with particular focuses on vendor specific application and Open Source Technology are investigated in this chapter.

- **Chapter Five** focuses on the implementation of a physical model to create a database system. DBMS is implemented for the application to the specified data model. The database schema is produced and the database is designed. Data processing and trial data population are demonstrated to create GIS-ready information using inherent GIS tools. An improved management of datasets from geospatial data capture right through to the GIS-ready information stage is shown. In this respect, a prototype is tested to address the achievement.
- **Chapter Six** demonstrated a trial online delivery of GIS-ready information via the Web using inherent vendor specific application and Open Source Technology combined with GML. Discussion is made to compare the two data delivery applications and the impact of these two technologies towards the development of the geospatial data handling.
- **Chapter Seven** concludes the thesis with a review of the research undertaken, highlights achievements and recommendations for further work.

Chapter 2

Literature Review

2.1 Introduction

The previous chapter introduced the thesis by presenting an overview of the research, a short background and motivations to the research, the aim, the objectives, a brief description of research methodology and the thesis structure. In this chapter, an account of a literature review discusses some issues concerning geospatial data management from raw to processed stage right through to the GIS-ready information. Geospatial technologies, such as traditional surveying, are explored as the basis for the production of any geospatial data. Object-based GIS is discussed briefly. A perception of crucial practices in surveying for geospatial data capture is discussed. A review of JUPEM as a geospatial data collecting organisation is made, leading to the issues for the research. Broad aspects of GIS which include science, knowledge, research and development with some issues affecting data management are also considered. Reviews of the geospatial data management practices in three geospatial data collecting organisations are drawn upon. Aspects of the delivery of GIS-ready information that is necessary for data distribution conclude the literature review. The last section summarises the chapter.

2.2 Management of Geospatial Data

Geospatial data and information are essential and increasingly demanded in this world of consumer-guided markets. In this respect, and to meet this demand, there are three requirements, which have been acknowledged to be central in today's geographic information age; more well-organised and diverse methods of geospatial data capture (Pieraccini et al., 2005; Khalsa et al., 2004; Holland and Tompkinson, 2003; Aschenwald et al., 2001; Gordon et al., 2001; Teng, 2000; Smith and West, 1999; Reed et al., 1996; Parker et al., 1996), improved geospatial data management techniques (Gogu et al., 2006; Balley et al., 2004; Hoult, 1995; Taylor, 1991; Raafat et al., 1991), and high-speed and more user-targeted methods of data dissemination (Yang et al., 2005; Dragicevic, 2004; Gong and Wang, 2004; Peng and Zhang, 2004; Chunithipaisan et al., 2003; Huang and

Lin, 2002; Ramroop and Pascoe, 2001). However, these developments have not focussed on taking consideration to combine management from geospatial data capture to the provision and dissemination of GIS-ready information in a single portal application.

It has been acknowledged that a lot of geospatial data and other types of data are being exploited for development and modernisation purposes (Kingsbury, 2005). For storage, management and use of geospatial data, various technologies are used and they can be aimed at just creating cartographic map products or for a full geographic information system. Geospatial data has played an important role for the accomplishment of area-wide mapping, inventory, analysis project and business management (Longley et al., 2002). These assertions imply that there is a requirement for these data to be collected, processed, managed and distributed in a manner that they can be accessed, checked and used efficiently. Historically and also now, most geospatial data were captured and placed in different locations, systems and formats, digitally and in hardcopy. They were deposited, filed and stored in islands of storage and systems within the geospatial data collecting agency or organisation. There has been a growing need for these sets of data to be integrated and shared together so that a single portal for access and usability can be realised. The research on handling of these varieties of geospatial datasets from data capture until the provision of GIS-ready information has not been recognised so far.

Many parts of the world have already developed an information-based society which is described by global information infrastructures, pervasive computing and information at our finger tips (Strobl, 2002). However, without taking care of locational aspects in information, the ultimate task to live and interact in this world may not be achieved. Thus data with location is mainly needed to manage geospatial data that we live with. Geospatial data capture is collecting the detail on land and putting it on map or plan with references system against the earth surface. Bringing a current position ‘onto the map’, putting it into the rich context of place and therefore facilitating joining up all the pieces of information pertaining to that location is at the heart of what geospatial information and GIS can contribute. Managing these data towards the production of GIS-ready information in one single application and view is a problem yet to solve.

In the information and communication age, the way digital data can be accessed and delivered has been acknowledged to be extensively and intensively researched and examined as they increase in volume and complexity (MacDonald, 2001). Successful data management depends on many issues affecting technology, resources, storage, networking and versatility. Geospatial data, the most common type of data, requires unique techniques and storage due to its format and heterogeneity. The facts that geospatial data have longitudinal and latitudinal aspects, and are actually global and large scale indicate how tremendous they exist in magnitude. GIS technologies include the global location as the key player for implementation.

In an organisation supplying geospatial data, data collecting and input is by far the single most time consuming exercise (UNESCO, 1999). It forms the major 'bottleneck' and consumes over 80% of available funds. Figure 2.1 demonstrates this graphically. In the activity of geospatial data capture and processing, users either develop their own formats for storing geospatial information or employ vendor-defined applications and their proprietary software. Most formats are often specific to the needs of particular projects and not meeting the needs of a broader range of users. It is then often very difficult to share and transfer data. Users recognise that much time and money is wasted when datasets could not be shared among a broader clientele. Organisation of data collection seems critical so that recapture and reprocessing of the same geospatial data can be eliminated.

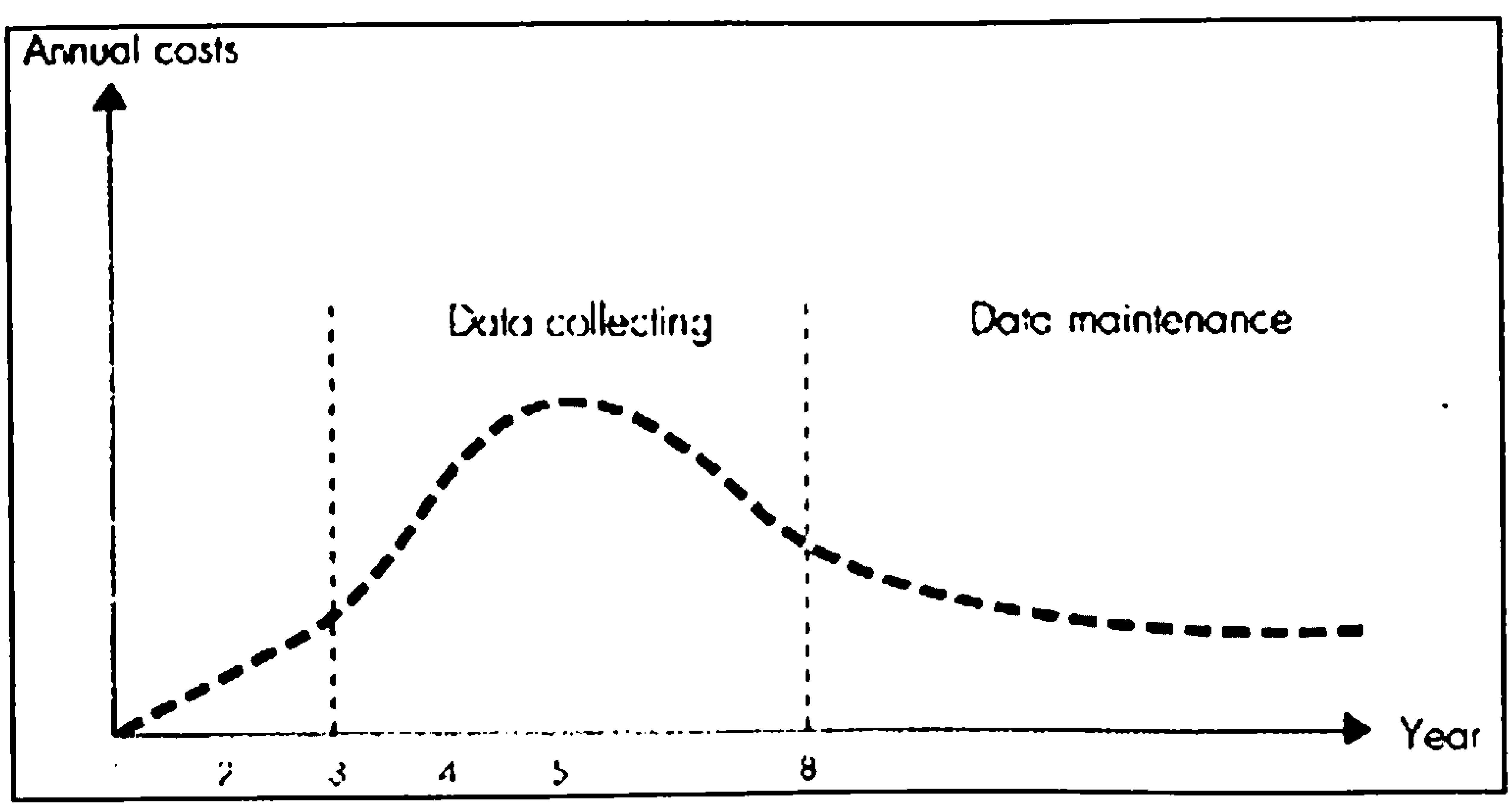


Figure 2.1: Data collection demanding high costs and long time periods (UNESCO, 1999)

As the world objects and their interaction become denser and more complicated, the issue of availability, production and development of GIS data and information presents a challenge to geospatial data professionals and geospatial data infrastructure (Maguire and Longley, 2005). Without the right mechanism for the management of this huge amount of data, the production of this information might face many hurdles. Using a single view of the data in raw capture, processed and GIS-ready state may help in reducing time, energy and budget.

Most organisations require data for their core business. In most developed countries, the existence of government-oriented organisations such as public authorities has been made dependent upon selling data (Su et al., 2000). The need for data in most data-dependent organisations requires solutions to the availability of data. Geospatial data capture and processed data that are available for the production of GIS-ready information in one virtual management is logically an issue. This is because islands of information files and databases can be integrated into a single management and access system which saves time, laborious search and query. Data sharing among users is another issue to be solved. GIS data is wasteful if not shared by data provider and user public alike. Sharing can mean many things to people, either from simple logical notion (Fulghum, 1987) to abstract matters of semantics and perception (Frank, 1992). In the geospatial community, data sharing is about the exchange and digital transfer of geospatial data between stakeholders and data users (Crompvoets et al., 2004). With the concept of sharing, data needs to be recorded only once by the data collector and then it can be accessed by multi-users electronically. Sharing data has huge benefits in spatial information systems. Importantly, it creates greater efficiency as higher cost can be avoided. Repeated data collection, storage and registration result in the redundancy of the data. Wider sets of needs can be achieved with less utilisation of resources.

Handling and provision of geospatial data efficiently from data collection to its dissemination had caused many issues. One of them is that all the geospatial data and information that were captured and maintained are seen as islands of files and databases in a typical geospatial data collecting organisations. In many of the organisations, administration and management are carried out based on divisional departments or units that produce different kinds of geospatial data. This is what happens in JUPEM where several sections implement different ways of capturing and processing them, various ways of storing and managing them. JUPEM produces topographic survey, cadastral surveys, air surveys and GPS surveys using different survey techniques producing different types and formats of data. There is no single portal accessibility of the various datasets collected. There is also the problem of data redundancy, semantics and data formats (Xu, 2003). GIS-ready information can be uncertain and inaccurate when produced from several surveys carried out over a period of time on the same site (Zhang and Goodchild, 2002). An improved management with 'birds-eye view' of the whole datasets from raw to the processed data may minimise these uncertainties by 'drilling down' and reviewing the origin of the data and how they are captured and processed. A way to manage the datasets by way of a single portal is needed using database concepts and geospatial tools through sophisticated graphics software. One problem the portal may help solve is the lack of awareness of what geospatial data already exists or has already been transformed to GIS-ready data within the geospatial data collecting organisation.

Geospatial data needs to be evaluated within two components: representational and communicative (Raper et al., 2002). The representational component is concerned with how objects situated in space and time come to be represented in geographic information, while the communicative component deals with how representations of geospatial data are understood by the users. GIS is a common system that handles geospatial data by using georeferencing techniques to tie information to the Earth, and geometric techniques to represent the phenomena. Communication of geospatial data is a combination process involving the creator and the user. In the communicative way, geospatial data can be made understandable by effective management which requires strategy, integration, information seeking and retrieval process, highly interactive graphics, and project management. Goodchild (1998) suggests that a geolibrary should consist of browser, basemap, gazetteer and collection components to handle geospatial data and information. A methodology for the management and access by GIS of geospatial data from capture to the GIS-ready information enables maintenance of a 'library' of information and datasets and eventually leads to departmental access. In this way, geospatial information can be seen as a benefit, appreciated and evaluated in a communicative way.

To date, GIS has great impact on a wider scopes and professional areas such as information communication technology (Kingsbury, 2005; Wiesel, 2002), development in education and academia (McKee, 2002; Longley, 2000; Wikle, 1998), business strategy (Ahern, 2004), geospatial infrastructure (Rajabifard and Williamson, 2004; Wheatley, 1998; Haines-Young and Watkins, 1996), local community (Greenwood, 2002; Kohsaka, 2000), applications for construction and property industry (Rollins, 2005; Lawrence, 2004; Daud, 1999) environmental treatment (Khalsa et al., 2004; Diepenbroek et al., 2002; Pundt and Bishr, 2002), geological changes (Lu and Ping, 1999), temporal and higher dimensional geospatial data (Nash et al., 2005; Parsley, 2001), natural and biological conservation (McBride, 2004; Vanderpoorten et al., 2004), military (Weers, 2004), medical geography (Casper et al., 2000) recreation park management (Wei-Ning, 1996), and land management (Karikari et al., 2005; Jeffress, 2003). Management of geospatial data from raw to processed and GIS-ready state using GIS has not been clearly presented in research specifically linking to raw survey field data, though there has been research related to developing scanned and digitised data into GIS-ready data. Various techniques of data capture include geodetic survey, cadastral survey, topographic survey, astronomy, air survey, photogrammetry, gravity measurement, hydrographic survey and tidal measurement. For management of these surveys, a one-view handling system for geospatial raw works and information, processed data right through to the creation of GIS-ready information is not yet available so far. In one-view handling system, a specific survey data and its end product should be accessed and managed entirely in one piece. Integration of geospatial handling solutions represents maturity in how we view, think about and communicate geospatial related issues.

From geospatial data capture to processed datasets and towards GIS-ready information, there are a lot of data types (eg. survey files, CAD files, remote sensing datasets, TIN model etc) involved that require managing. Modelling, storing and management of these data types in a single portal within GIS environment has not been achieved so far. Initially, these spatial and aspatial data are captured and processed in the office within several departments that have separate systems of storing and access. For management in one system with multi-user access, it is sensible to create the management package consisting of three different features namely, capturing geospatial data, processing the dataset and converting into GIS-ready.

2.3 Geographic Information Systems

The GIS field began around 1960, with the discovery that maps could be 'programmed' using simple code and then stored in a computer allowing for future modification when necessary (GIS Development Pvt Ltd, 2005). This was a welcome change from the era of hand cartography when maps had to be painstakingly created by hand; even small changes required the creation of a new map. The earliest version of a GIS was known as 'computer cartography' and involved simple line work to represent land features. From that evolved the concept of overlaying different mapped features on top of each other to determine patterns and causes of spatial phenomenon. The information and communication age has thus witnessed the expansion of GIS across the globe depicting research topics from the technological detail to the social context.

GIS has many definitions dependent on its context of use in community (Chrisman, 1999). Longley et al. (2002) give these definitions:

> *"...a spatial decision support system"* and *"...a mechanised inventory of geographically distributed features, and facilities" (page 10)*

The first definition is quite relevant for a geospatial data manager when needing to decide and propose survey by viewing the properties of spatial survey limits, history and scope of the field project. The second definition explains the value of GIS to keep track of such entities as surveys, aerial surveys, processed data (CAD) and eventually the management of the GIS-ready information.

The purpose of GIS as stated by Zeiler (1999) is;

> *"...to provide a spatial framework to support decisions for the intelligent use of Earth's resource and to manage man-made environment" (page 2)*

Management of a task setting concerning survey package and its applicable attributes for the development of GIS-ready information can be categorised as man-made environment or situation. Specifically, GIS can be utilised to manage project tasks from its initial administration aspect until the point of geospatial data and information output itself within a single spatial database. A spatial database is the storage of geographic data in a prescribed format, including the location, shape, and description of geographical features as well as the relationships between different features (Rigaux et al., 2002). A spatial database usually includes co-ordinates and topological information. Then using GIS tools and Web technology they can be converted into GIS-ready information that can be accessed throughout an Intranet within the enterprise and public domain. Geographic information systems tools thus can improve the performance of an organisation's business operations considerably.

What is the relationship between GIS and management? There are two scenarios involved in relating GIS and management. Firstly, the GIS project created needs to be managed for maintenance and future development. This includes all components of the actual GIS project from specifying need for the GIS project, the manpower and skills, the computing hardware and software. The second relationship covers the use of GIS to manage many projects which is not merely spatial related but will have spatial properties when well-managed by GIS. The purpose of using GIS to manage projects is to produce more effective, more efficient, more equitable, or more predictable outcomes. By using GIS, people also enjoy advances in technology, organisational development and modification of working environment. Good geospatial data managers are able to turn these circumstances to their advantage.

A more general and broad idea of GIS is that it consists of physical components such as trained staff, computer, digitiser (to produce digital map), scanner (to produce digital aerial photograph), large format plotters (to print hardcopy map), digital spatial and attribute data (consist of survey and processed data), databases (management), network communication (data sharing and map services) and much more (Figure 2.2). These can be categorised into GIS main 'attributes' namely hardware, methods, software, data and people as illustrated in Figure 2.3 below. GIS may be represented as object consisting of these attributes.

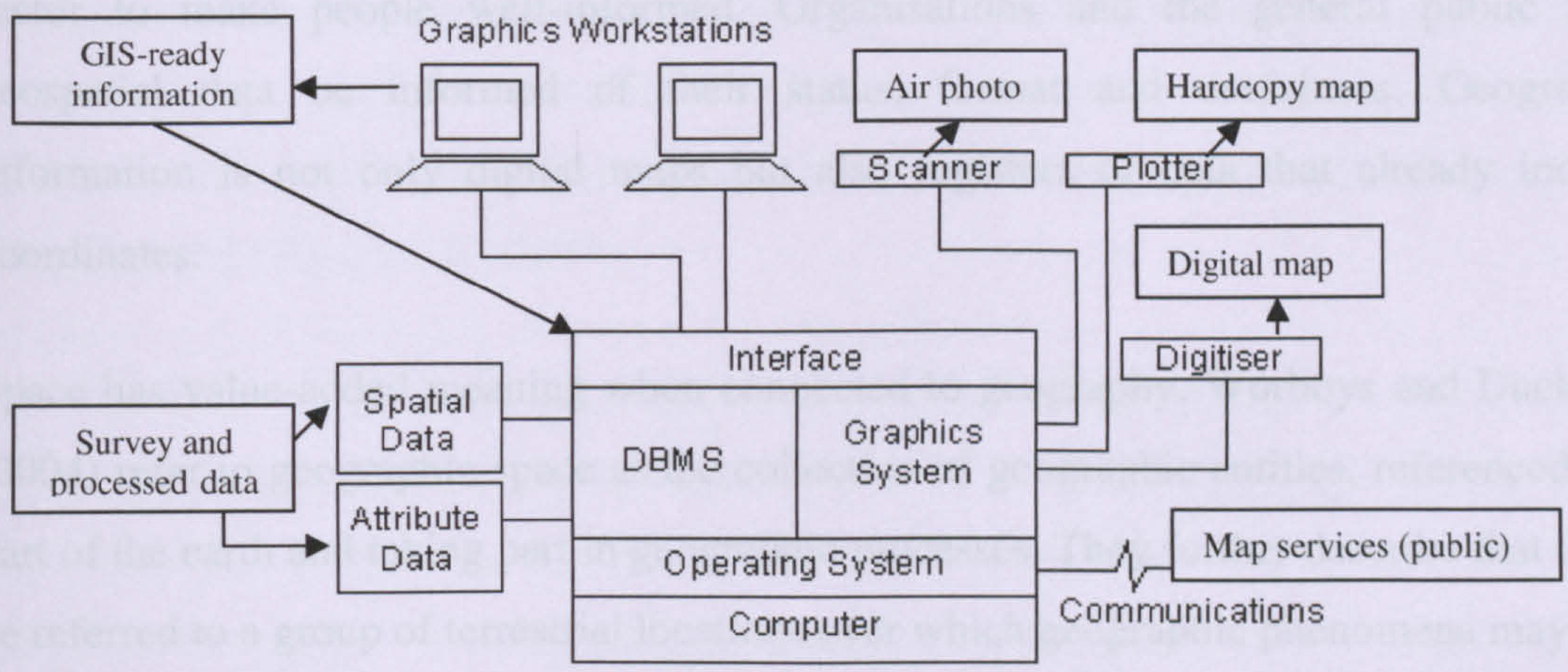


Figure 2.2: Physical components of GIS

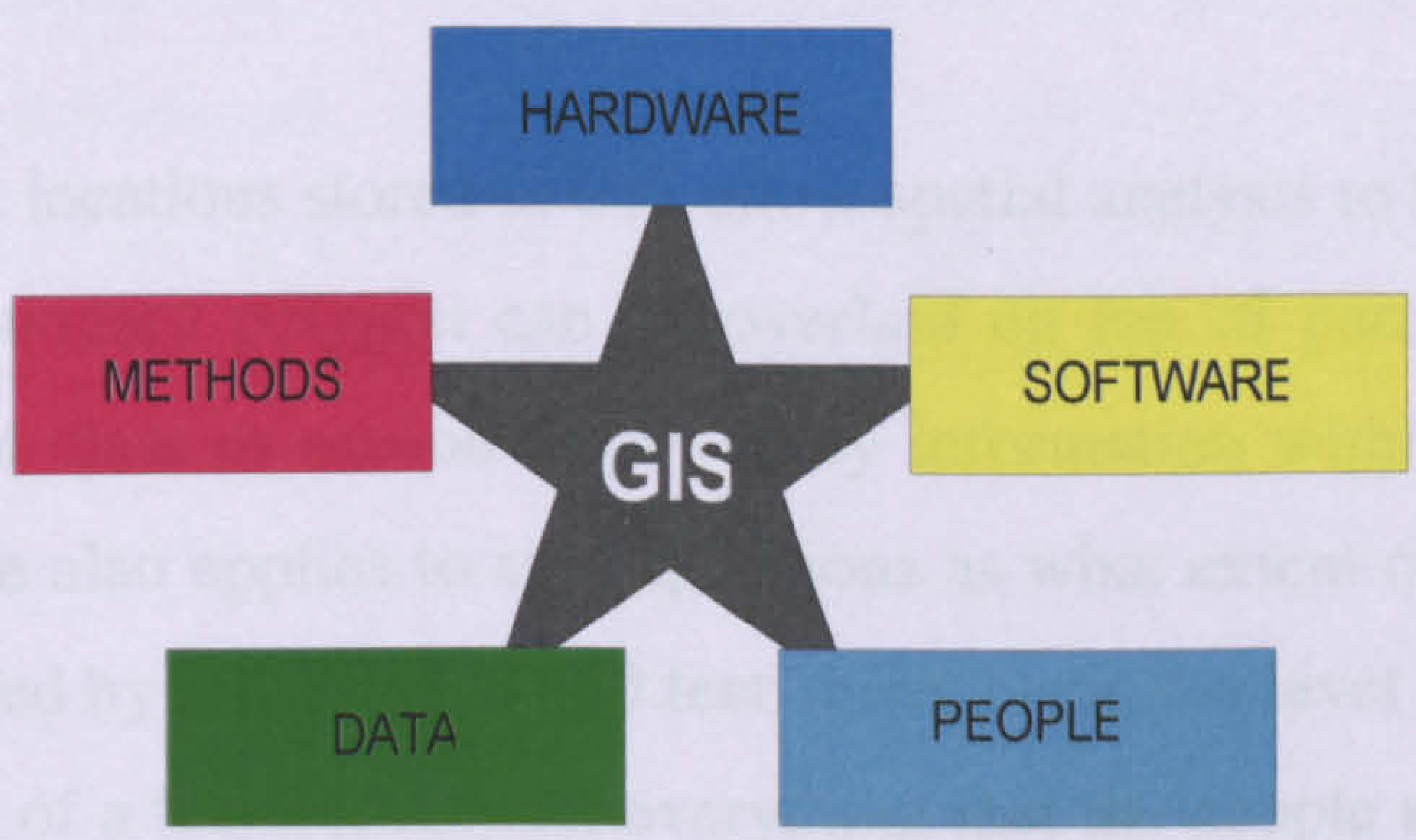


Figure 2.3: GIS as an 'object' and attributes

The abbreviation GIS is no longer fresh to many people. A GIS stores and manages all the data features that were found on a map such as roads, buildings, water tanks, and river and district boundaries. Survey marks, survey lines and polygon parcel lots are data features found in survey plan. These features are stored as points, lines, polygons (areas), and cells (image pixels) in a computer. GIS users put all the similar features into categories called themes, which are stacked on top of each other to form the map that is displayed and printed. The GIS also stores information about each feature such as road name, building type, parcel number, owner name, and survey network information in a database. The digital features and database are tied together so that a person can perform analysis on the map features. GIS allows us to answer questions such as Where is it? What is it? What spatial patterns exist? What has changed since...? What if...? How many have…? Our information society has realised that geographic information is an important factor to make people well-informed. Organisations and the general public need geospatial data be informed of their status, format and usefulness. Geographic information is not only digital maps but also registers of data that already include coordinates.

Space has value-added meaning when connected to geography. Worboys and Duckham (2004) refer to geographic space as the collection of geographic entities, referenced to a part of the earth and taking part in geographic processes. They further describe that it can be referred to a group of terrestrial locations over which geographic phenomena may take place. In this context, GIS have been used as an increasingly useful tool that integrates the complex information from different data sets in space, location and with geographic phenomena.

Features with actual locations stored in GIS allow spatial analysis to be performed, which means a district boundary polygon can be overlaid on top of parcels and a list of all parcel lots with first class or second class survey information within the district can be generated. The same also applies to such questions as what extent or portion of areas of topography is covered by hilly land at 500 feet above mean sea level or how many people live within 500 feet of a feature. Almost everything that the people and the world do has some measure of 'place' and can be represented in a GIS.

2.4 Object-based GIS

An object-based GIS, which employed spatial units as the spatial object, has been used as a project management tool to track status and as a reporting mechanism (McConnell and White, 1999). This was possible because the work involved was associated with spatial areas. Young (2004) describes the possibility of object-oriented software in managing surveying companies which was very much involved in the handling of surveys and geospatial data collection information. He mentioned that with surveyors' data being always geo-referenced and related to spatial issues, GIS tools would allow a day by day basis management of the surveying business. Mitasova et al. (2002) explain, that huge amount of one-purpose spatial data or isolated 'information islands' may not be efficiently and economically used if geospatial data and processes are not well-organised. Voisard and David (2002) emphasised that geospatial modelling with database conception for this purpose should be object-based because it allows efficiently defining collection of geographic objects and relationships amongst these objects. Object-based GIS have conceptual values that lead to the full understanding and intuitive perception of the real world (Wang et al., 2005).

An object-oriented GIS allows encapsulation of geospatial entities so that all of its geometry, data, and behaviours are contained in a single object (Worboys, 1994). Traditional GIS represents the geometry of spatial entities as simple raster or vector data with, at best, a simple relational table connected to it. Any data put into this table must be normalised and converted from its natural form, and any linking to modelling is a difficult disconnected process. Object-oriented GIS, which embedded geometry, data (in natural form), and behaviours in a single object, offers the capability of remodelling and reuse of data (Ghandeharizadeh, 2003). By capitalising on the benefits of object-oriented design, an application using GIS can present a single coherent view of a database of arbitrary objects to both the user and the developer (Worboys et al., 1990).

A technical committee (TC) formed to develop standard in GIS under the International Organisation for Standardisation (ISO) has proposed the use of UML as the data modelling language in the *ISO/TC211 Geographic Information/Geomatics,* a committee responsible for the ISO geographic information series of standards (ISO, 2005). For this reason UML is utilised for this research. Moreover, most recently, most object-oriented model developers have implemented UML as a standard notation for deriving an object model and it is approved by leading software and database companies (Pooley and Stevens, 1999).

2.5 Surveying as a Geospatial Data Capture

Surveying, an aspect of geospatial data capture, is a discipline within geographic information science which is associated with the capture of measurable information and data on earth and above. On the whole, surveyors are concerned with accurate representation of the Earth, while GIS professionals care more about data and their relation to a particular geometry. In actual fact, both are using spatial information and map or sketch the Earth. This resemblance therefore should facilitate the treatment or management of the information and data in an appropriate, chosen single system. GIS has become imperative to surveyors because GIS can allow management of surveys due to the surveyor's geo-referenced data and also GIS has the management capability to know the status of a survey and who is the surveyor working on it (Weir, 2004). Preliminary geospatial data such as surveys, secondary one such as processed dataset and primary type of geospatial data such as GIS-ready information can be managed within the GIS environment. Data with spatial dimension and coordinated using a reference model to the Earth's surface is what GIS can serve particularly admirably. With advances in communication and Internet, surveys can be directly linked to computer which enables efficient integration of current geospatial data handling technology with survey technologies (Figure 2.4).

A few years ago, there was a growing trend of influencing GIS within the space of surveying work and data capture practices (Granger, 2005; Majeed and Parker, 2004; Utler, 2002; Adair et al, 1998). GIS systems and technology have also advanced in a way that raw and accurate survey data can be incorporated directly into the system. Surveyors perform classical work that can be integrated within the modern GIS framework. Surveyors are skilled in cartography as they have to map out the detail of the measurement in an informative way. Their synopsis of a particular survey can be told in whole in the form of a map or hardcopy plan. Survey data are always accurately spatially referenced and have superior advantages for GIS usage, but the survey process and project of geodetic survey, topographic survey, air survey, and other type of surveys are lacking a management infrastructure to ensure efficiency in time, economics and effortless application throughout the creation of GIS ready-information.

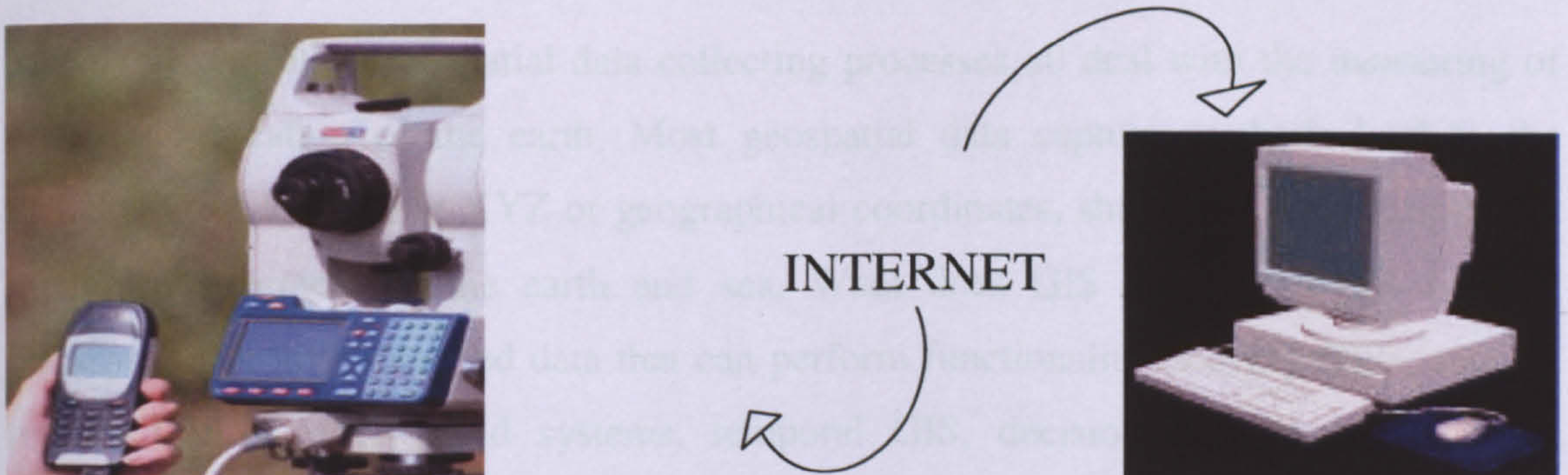


Figure 2.4: An Internet connection using mobile phone made to the survey office enables data be downloaded directly into computer and vice versa if surveyor needs previous survey information from the office

Historically, the surveyor was once also explorer, resource appraiser, town planner, delineator of route ways and the moulder of rural cadasters and landscapes (Kain, 2002). But can surveyors search features of the survey done temporally? Can we get historical snapshots of the data? No, only current versions of the data are supplied. Historical feature can be created when applying change-only update by storing outdated features. In today's available informative computing, there is a need that surveyors and GIS people should be able to access 'features' and 'attributes' of survey package to get information on the chronological sequence, the state, location and status of the survey. They should be able to access survey datasets, processed datasets and processing steps for future spatial data assessment, spatial data creation and decision support as well as GIS-ready information development. In essence, surveyors' inputs for the solution of the real world problems are not minimal.

Figure 2.5: Traditional surveying emphasizes on distance and space (PSM, 2004)

Surveying and other geospatial data collecting processes all deal with the measuring of space and location of the earth. Most geospatial data capture methods lead to the production of location in XYZ or geographical coordinates, shape and parameter of the dynamic movement of the earth and sea. What does GIS need from this setting? Surveyors can give surveyed data that can perform functionality in every GISs, such as mobile GIS, location-based systems, temporal GIS, decision support system, road network analysis, and 3-dimensional GIS because they are location-based and spatial. Geospatial data captured from beginning till the processing stage are rich with information that is friendly to GIS.

GIS-ready information development is dynamically evolving, and its linking and benefits to surveying and geospatial data capture are essentially important. In ESRI's User Conference from 2001 until 2004 the topic of survey and GIS has been a popular issue. The number of relevant and related papers presented increased from 5 in 2001 to 40 in 2004. A special conference called 'Survey and GIS Summit Bridging the Gap' has begun and has been held since 2004. This shows a great increase in research and awareness in the issues of surveys and GIS. Survey data as essential geospatial data, their integrity, surveyor professionals, and collaboration between surveyors and GIS specialists are certainly matters that may be integrated and debated for efficient production and dissemination of geospatial data and GIS-ready information. Approaches to efficiently manage raw geospatial data and processed datasets and eventually perform as useful GIS-ready information and disseminate them over the Internet are certainly becoming main topics in the issues of GIS data development.

Traditionally, a gap has existed between data capture (surveying) standards and practices and those in GIS. GIS specialists had expressed that the surveying community are slow in producing map and data products, not to mention expensive and supplying unfinished artefacts. They instead produce their own data which do not conform to the standard of formality. Surveyors responded that clients are confused and given incorrect norms of data along with the map (Jericke, 2002). This creates misunderstanding and surveyors' esteem may be challenged in term of their public mission and work. This mental disagreement at least can be overcome with the provision of a handy management system via GIS technologies by which surveyors may incorporate their geospatial data capture, processed data, leading to the creation of GIS-ready information. Geospatial data capture project management and the related products should therefore be stored and managed in a GIS domain. An improved quality of surveys and manageable datasets may increase the reliability of maps and plan layers stored and used in a GIS. It is therefore believed that the link between surveyors and GIS specialist can be refreshed.

2.6 Survey and Mapping Department Malaysia (JUPEM)

JUPEM is the only government organisation in Malaysia that produces spatial and non-spatial data for national mapping and land administration. It is the custodian, accountable for production and maintenance of a complete national geospatial data within the cadastral and mapping databases. In addition, it has responsibilities to coordinate planning and development of cadastral and mapping systems based on ICT, thus supporting the e-Government concept. JUPEM's official vision and role is:

"to provide efficient and high quality services in land surveying and mapping and to manage and disseminate geospatial information in accordance to the National Vision" (JUPEM, 2004)

The main tasks of JUPEM are to produce geodetic, cadastral, topographic and air survey data. Geodetic survey is mainly concentrated on the establishment of GPS stations, measurement of gravity, levelling for height control, maintaining of Malaysian Active GPS System (MASS) station and tidal observation. Cadastral surveys are important for land parcel mapping and administration as well as cadastral base mapping. Topographic survey and air survey are carried out for national mapping purposes. Basically the production of these survey datasets is done by various techniques: conventional level, theodolite and EDM, total station which is a combination of electronic transit and EDM, GPS survey, air survey, remote sensing and photogrammetry.

Figure 2.4 shows the JUPEM organisation chart which illustrates the existence of two divisions comprising of several sections and units. Several sections and units in JUPEM which are not interconnected and related administratively reveal that surveys, survey datasets and geospatial data are everywhere in different systems. As in Figure 2.6, it can be noticed that several sections and JUPEM in the states within both Cadastral and Map

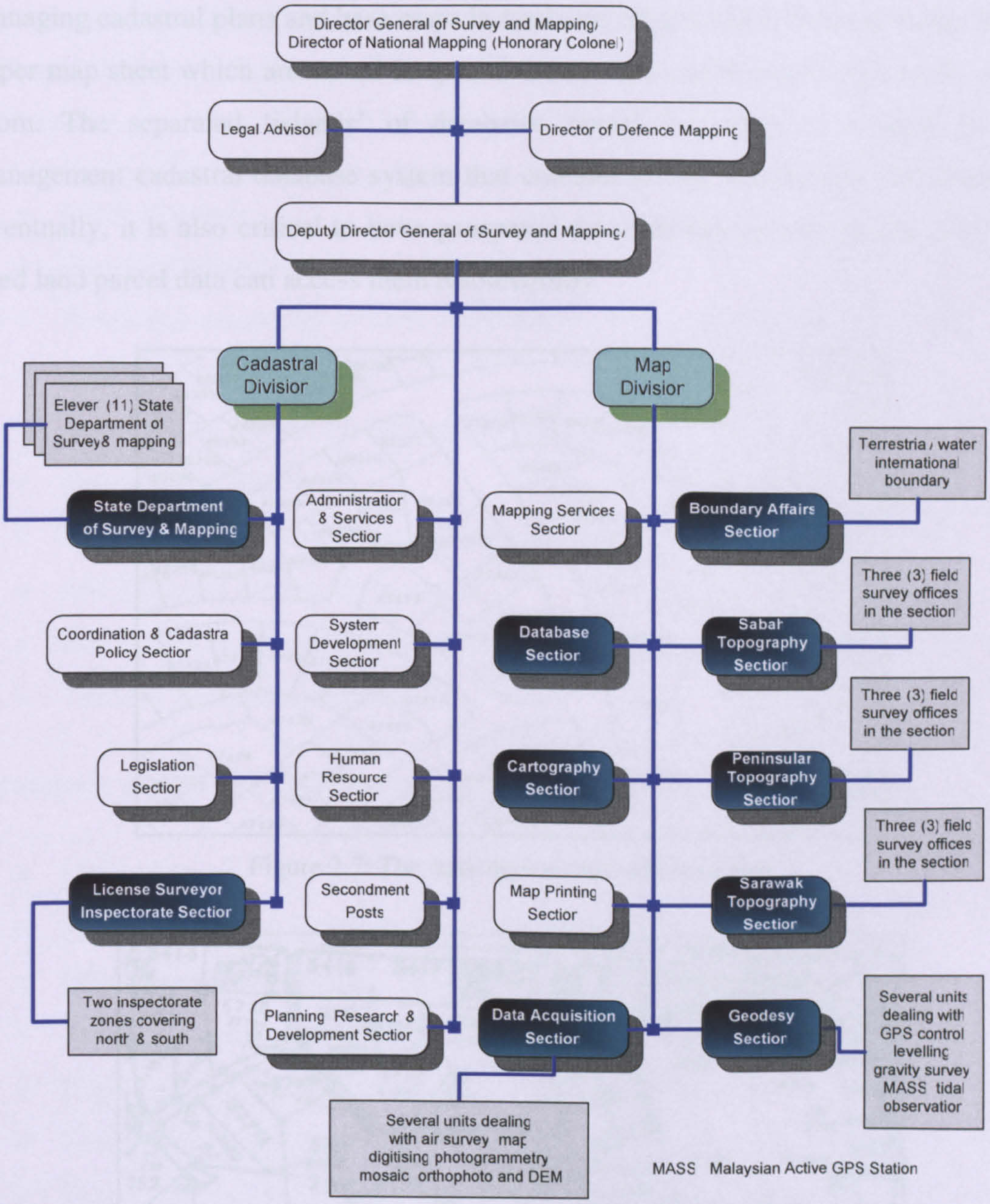


Figure 2.6: The organisational chart of JUPEM. The shaded sections have several units that implement few techniques of surveys and store data in various locations

Division, have different sub-units in different administrative zones and scopes of work, dealing with different kinds of data. There is no Database Section in the Cadastral Division, centrally at the JUPEM headquarter that organises in one system all the cadastral data in all states. State departments only implement cadastral surveys and some states are currently implementing their own digital database. Figure 2.7 shows sample of a digital cadastral database in a state department. Some state departments are still managing cadastral plans and base maps in hardcopy (Figure 2.8). These are large format paper map sheet which are stored in special drawer that can be easily used in the strong room. The separated 'islands' of databases reveal that there is a need for one management cadastral database system that contains all the nation-wide cadastral data. Eventually, it is also critical to have geospatial data delivery system so that users who need land parcel data can access them resourcefully.

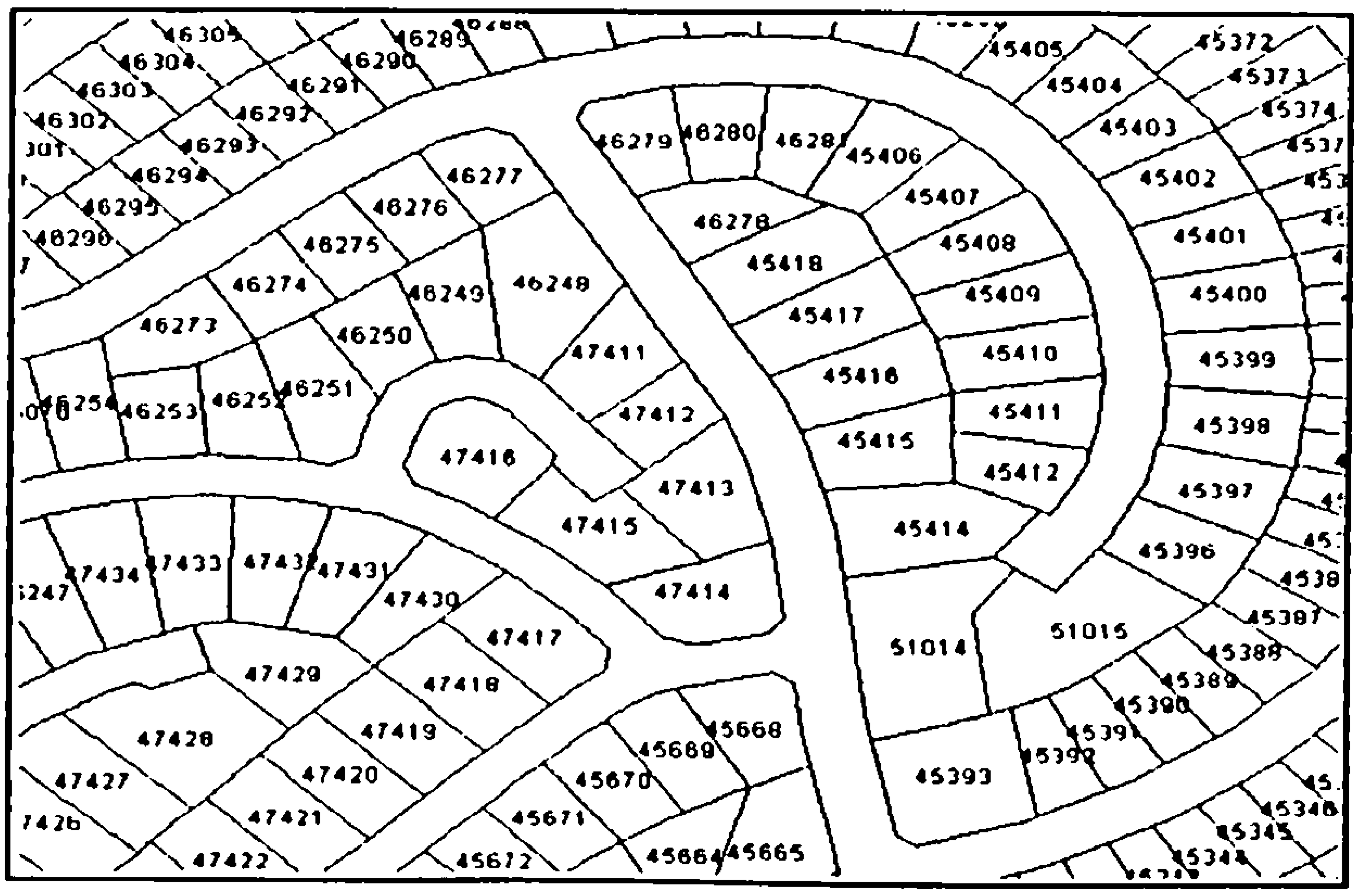


Figure 2.7: The state digital cadastral base map

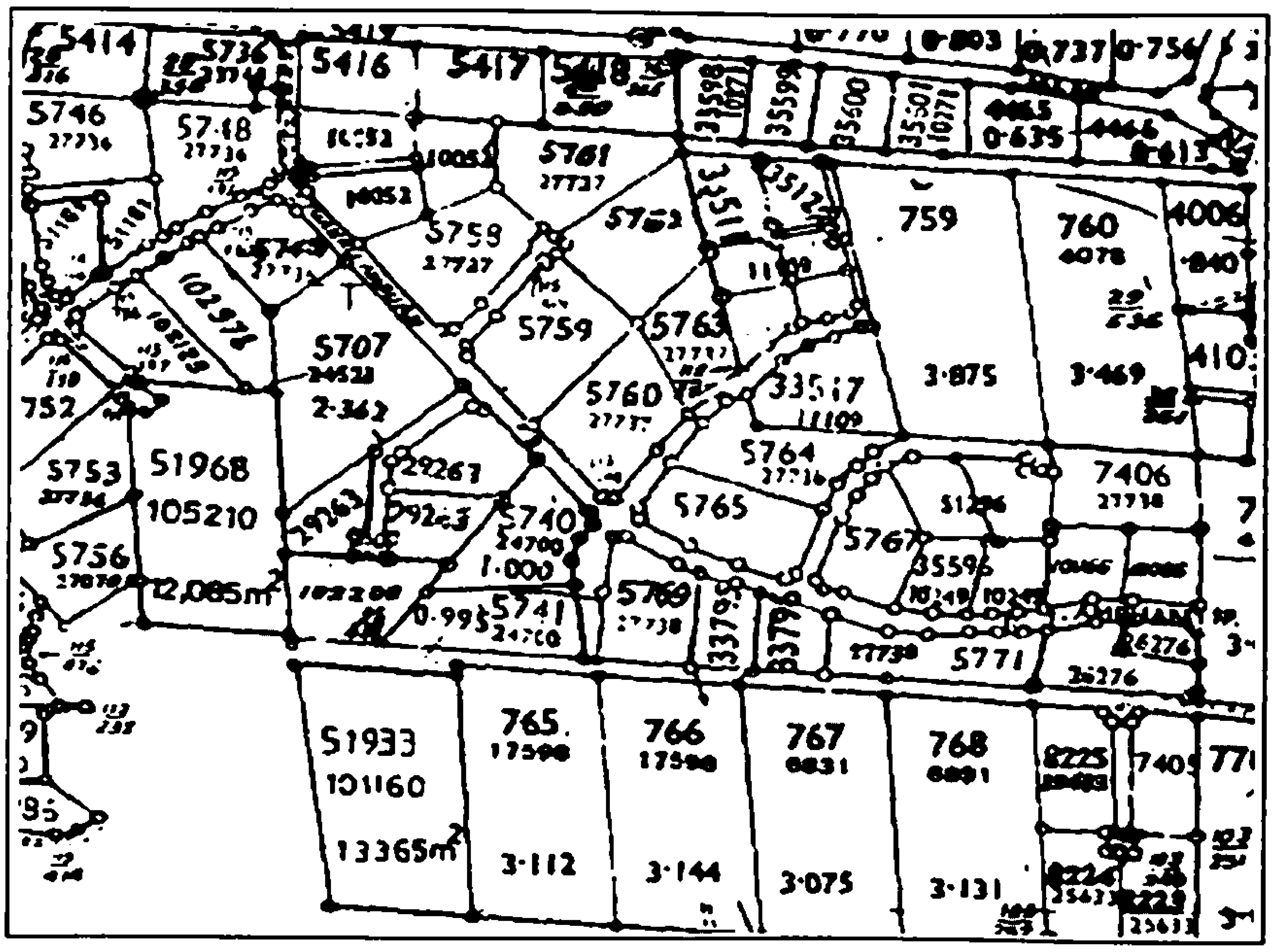


Figure 2.8: The state hardcopy cadastral base map

A survey package is a collection of surveys, their associated information and their collected datasets. A cadastral or topographic survey may consist of the following information:

- Limit of survey.
- Sketch of survey site.
- Field notes consisting of survey network and survey stations (conventional method).
- Local control framework data (traverse and GPS station).
- Surveyor and date.
- Field tracing or witness diagram of existing surveyed stations.
- ASCII files of control observation.
- ASCII files of observation from survey station 1, 2, 3 and so on.
- Traverse computation files.

Information and data for an air survey package may consist of the followings:

- Limit of survey.
- Flight information (pilot, photographer, aeroplane etc).
- GPS ground control data.
- Aerial photographs produced.
- Flight line information (course, distance and coverage).
- Camera detail and instrumentation.
- Camera calibration information.

A survey package thus has a lot of information and data to be managed and maintained. They actually have historical value and are useful as source information for resurvey and land development. In fact, they can be very valuable as origin data points in the processing steps and for the development of GIS-ready information. But the datasets and processing steps need to be managed in one system that can be accessed concurrently by all users involved in the raw data capture, processing and the production of GIS-ready information. Processed data are mainly CAD files, ortho-photographs, mosaiced raster images (coverage) and DEM. Aerial photographs are processed to produce coverages and also used in photogrammetry to produce CAD files and DEM. Field topographic survey package is processed to produce CAD files for national basemap. Field cadastral surveys are carried out to produce a certified plan and this is input in a cadastral sheet that serves as a state-wise cadastral basemap.

According to information provided from JUPEM, in the year 2002, JUPEM has about 2.0 TB of digital geospatial data which are available for government, business, public and individuals consumption, consisting of:

- Cadastral data (780 GB).
- Mapping data (1100 GB), and
- Geodetic data (100 GB).

The above information has become the strategic driver for the implementation of this research. As JUPEM is an important data provider to serve e-Government purposes, this amount will increase tremendously. Outsourcing work has been carried out to fast-track the national mapping activities. Cadastral survey has been largely assisted by private surveyors who are to deliver digital plans instead of hardcopy land parcel plan for land title preparation.

Cadastral data are handled by each Malaysian state. The data are produced and stored in the form of a seamless database and raster file data (scanned title plan). There are no proper GIS that allow single view and multi-user management. Mapping data are handled by the various sections at JUPEM, namely Mapping Division, Cartography and National Map Library. The data are produced and stored in the form of raster file data and vector file data in each section. They are sold in hardcopy form. Geodetic data are handled by the Geodesy Section at JUPEM headquarters. The data are produced and stored in the form of text file data. These data can be sold in hardcopy and digital form at the counter. The scattered data and information across the department actually need single portal management so that staff and public are able to access data resourcefully. A one-stop portal (front counter customer service) to search for land detail and map for development purposes could be a solution for the public to find out about data and information available in the department. Currently, most data can only be accessed according to sections and unit and by certain staff within the sections and units in the department.

2.7 Geospatial Data and Technologies

What are geospatial data? Firstly, spatial refers to any surface, not confined to space of the Earth's surface. Information or abstracts of the reality shown on maps or those organised in a digital database that are tied to the referencing system by coordinates, address, or other means are collectively called geospatial data (Bossler et al., 2002). Therefore, geospatial data are a subset of spatial data. A great majority of all data in this world are geospatial data.

Geospatial data are therefore data existing within space and at the same time are specifically applied to the surface and near-surface of the earth. As geospatial data are captured in the field and processed in separate environments, they are also handled in digital filing systems, file cabinets and in a standalone database. Essentially, they should be stored and managed properly because they can become information that possesses the power of GIS. Previous developments show that they can be stored intelligently, analysed effectively and queried effortlessly using graphical interface and database system (Su et al., 2000; Worboys, 1994; Medeiros and Pires, 1994). The information then can be delivered across the network that makes it fully GIS-ready information. From geospatial data capture to GIS-ready information, GIS tools can be seen to provide considerable service.

Geospatial data is generally handled in GIS by way of several stages. Molenaar (1998) described the following, which are relevant to the research. The first stage is related to data acquisition via field surveys, digitising, photogrammetry and photo interpretation etc. The collected data are normally kept in a database where they can be accessed for query and further processing. The output of the second stage data model will then be presented to the user in some readable format. Each stage has its own needs with respect to the data management, meaning that the data models that sustain the activities in them are different as shown in Figure 2.9. The arrows indicate some data processing to convert these data to the data model of the next stage. This model is intended as a basis for the geospatial data to GIS-ready information management.

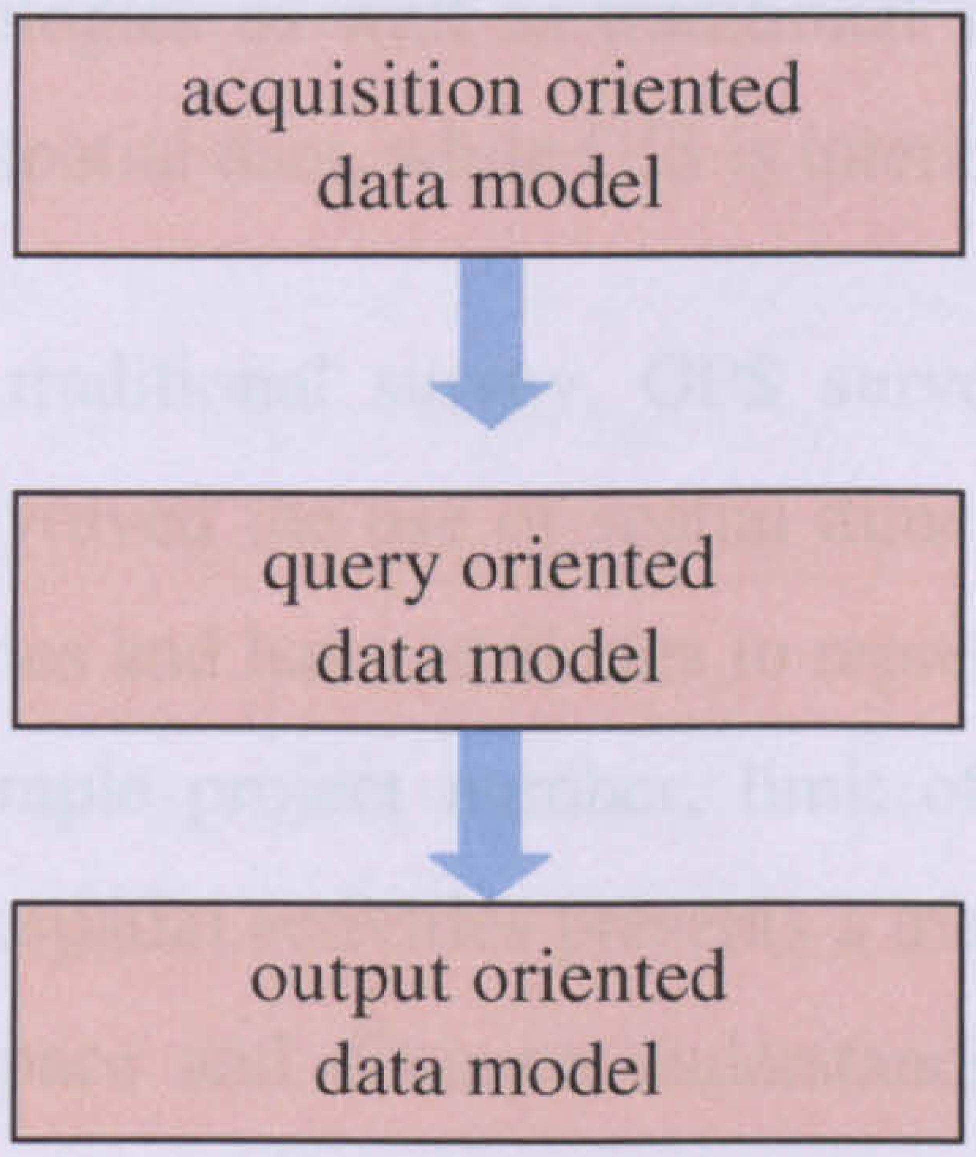


Figure 2.9: Data models should be defined per processing stage (Molenaar, 1998)

The nature of any natural or economic activity with a spatial dimension cannot be properly understood without reference to its geospatial qualities (Gielsdorf et al., 2004). Geospatial data have two essential parts: location and attributes.

- Location: typical locational references are latitude and longitude and national grid references such as the National Grid. However, other geospatial codes can also be used to identify location, such as postcodes.
- Attributes: any locality will have a number of characteristics or properties associated with it. These attributes are usually kept in tables, containing such information as building types, population, elevations, etc.

Most geospatial data from field are captured by:

- Digitising maps.
- Traditional survey.
- Modern positioning methods using GPS.
- Remote sensing techniques that is, satellite or aerial survey.

Geospatial technologies as mentioned by Thurston et al. (2003) are GPS, GIS and remote sensing including photogrammetry, LIDAR, air survey and satellite imagery. GPS survey and remote sensing technologies as well as traditional surveying are the main way of capturing or surveying geospatial data, while GIS is intended as the management tool.

Digitising existing maps, traditional survey, GPS survey and remote sensing are all geospatial activities that involved the use of spatial dimension, location (geo-referenced coordinates), have boundaries and have attributes to represent the real world phenomena. The attributes are for example project number, limit of surveyed area, land title and status. The result of the geospatial activities presents a map, data and object visualisation that are embedded with space and distance understanding. GIS core matter is about representing spatial phenomena. GIS tools can therefore perform excellently in storing and managing these geospatial data capture and technologies.

2.8 What is GIS-ready Information?

GIS-ready information can be defined as an object, abstracted from the real world, perceived as an entity with attribute or behaviour that can be linked semantically to the human's own view and action, organised in a database that has query capability and online accessibility through a sophisticated graphical interface. This informative object is the vision of the output of the geospatial data management addressed in this research. GIS-ready data are essentially what one wants to share over the Internet.

2.9 Developments of GIS

GIS has seen a tremendous speed of evolution especially as it can be multi-use in many multidisciplinary areas. Mitchell (1998) describes its significance in neighbourhood, towns and cities and how it is poised to reach even greater levels when the technology becomes accessible to more people in many new ways. GIS is not only about supplying tools for specialists but is assisting citizens with services within the enterprise, inter-organisations and between public and the GIS implementer.

Developed countries have contributed extremely well in GIS that map and information can be handled at the fingertips. There are achievements in GIS in the developing countries which indicate strong awareness such as in utility management (Igbokwe and Emengini, 2005), land administration (Karikari et al, 2005), local authority (Van der Vegt, 2001) and enterprise GIS towards e-government (Baumann, 2004). Malaysia as a developing nation is currently exploring into infrastructure GIS (Shariff and Krishna, 2003; Rainis and Sulaiman, 2001) and interoperability of land and survey data (Chunithipaisan et al., 2003). Malaysia is a country undergoing transformation into an advanced computer-driven and knowledge-based society with ICT technology being largely used for economic development (Awang, 2004).

GIS is a dynamic and growing trend evolving rapidly indicated by researchers and implementers promoting increased numbers of related publication and conferences. Popular research media such as the International Journal of Geographic Information Science began in 1987 and the decision to form the Association for GI (AGI) a UK national organisation for GI was made in 1988. Journals such as Institute of Electrical and Electronics Engineers (IEEE), Computers and Geosciences, International Journal of Applied Earth Observation and Geoinformation, Journal of Geographical Systems have included and recognised GIS as part of the crucial technology and science for developments of human, earth and its resources.

GIS has served many professions efficiently. Chainey and Ratcliffe (2005) described in their book how the power of geographical mapping can offer important clues that aid detection and improve our understanding of crime so that we can more effectively reduce it. In estate management, GIS inspires a more specialised image through accurate, consistent, professional-looking maps that carry company's or government agency's brand and that create real virtue to public requirement (Rollins, 2005). In conservation biology, GIS has been recently applied to examine the impact of a series of factors related to ecological conditions and land use on species diversity and rarity patterns at the landscape scale (Vanderpoorten et al., 2004; Neldner et al., 1995). Social scientists have used GIS to evaluate gender development (Bosak and Schroeder, 2004) and social behaviour in the ICT community (Nyerges et al., 2002).

Surveying, an activity of geospatial data capture, is a practice to measure an area of land, and to record the details of it, especially on a map. Surveying had shown great impacts and growing development of influencing geospatial science and GIS trends (Jiang, 2005; Jericke, 2002; Nix and Hill, 2001). Survey datasets, as part of geospatial data, which are spatially and aspatially related information are significant and essential to spatial science.

GIS therefore can play an important role in influencing raw surveys and processed surveys right through to the production of GIS-ready information, in a single management.

2.10 Review of Current Geospatial Data Management Efforts

This section reviews geospatial data management efforts by national data collecting organisation that begin from data capture to the production of GIS-ready information in three countries, United State of America (USA), United Kingdom (UK) and Malaysia.

The United States Geological Survey(USGS) of the US was instituted on March 3, 1879 by President Rutherford B. Hayes who signed the bill appropriating money for sundry civil expenses of the Federal Government for the fiscal year beginning July 1, 1879. The USGS serves by providing reliable scientific information to describe and understand the Earth. USGS has roles to minimise loss of life and property from natural disasters, manage water, biological, energy, and mineral resources, and enhance and protect quality of life. Its main core business is not centralised on capturing ground raw data and producing GIS data but more on archiving the ready data and using them to serve the organisational purpose. Raw survey data for the production of processed geospatial data and GIS-ready data has not been handled using single GIS application. It has a National Clearinghouse for geological maps, datasets, and related geosciences information, links to major USGS geosciences databases and programs as well as resources for creating digital geologic maps. Geospatial data resources are not their own but are obtained by contracting out the data collection to surveyors and GIS data supplier.

Land Survey Information System (LSIS) of USA is a database enterprise that keeps land detail, cadastral parcel, map sheet division, district and states administrative boundaries. Geospatial data in Shapefile format in the LSIS can be accessed from state to state through the Internet (Arctur and Zeiler, 2004). The Public Land Survey System (PLSS) institutes the land parcel record of the nation in a rectangular survey system. It is called a rectangular system because wherever practicable the units are in rectangular form. The Geographic Coordinate Data Base (GCDB) is the digital coordinate-based representation of the Public Land Survey. GCDB is a collection of geographic information in coordinate form representing the PLSS of the USA. The GCDB grid is computed from

Bureau of Land Management (BLM) survey records (official plats and field notes), local survey records and geodetic control information. Many of the GIS applications used in the USGS's various systems and bureaus are vendor-specific software such as PLSS and LSIS which use ESRI's product.

So, as we can see, there are various bureaus and systems being set up that cover from the geospatial data capture (BLM) towards the creation of GIS-ready information (LSIS) and they is no establishment that allow the access of all the information in one single view by many users.

In the UK case, Tony Gray from Ordnance Survey (OS), Southampton has responded on an email sent to him enquiring how survey and datasets were managed until the data is ready for GIS purposes and online sale.

In the UK case, DDMS (Digital Data Management System) is the data management system that provides storage and retrieval system for the digital data. DDMS is the source for the Topographic Basic Scales data from which both the Landline and MasterMap Topo Layer products are derived, together with other product files. DDMS uses 191.1 Gegabytes of storage. The topographic data is held in individual 'flat files' according to an Ordnance Survey specific data format known as OS 96. The application that manages the archive allows internal users to extract data with write permission, to be updated by an edit and returned to the archive, or as read only data which is not returned to the archive. When map data is edited and returned to the archive a new copy of the map data file is saved 'overwriting' the file as opposed to just the changes being made within the archive. DDMS provides a distribution service to transfer the data to the various production systems for them to edit in their own environments - i.e. the data is not edited directly on the archive. Before any data is committed to the archive it is validated to ensure that it is to specification. Access to the archive, and certain metadata, is controlled by an ORACLE relational database which tracks the current status of the data files. Only one user at any time is allowed write permission on any specific map data file. The DDMS software has been written in-house by a dedicated development team using C++, FORTRAN, Oracle and UNIX.

Data sources that are populated into DDMS can be from remote sensing, private land surveyors and in-house land surveyors but information about data capture are not kept. Private land surveyors only give resultant datasets. Remote sensing data and information about its development are kept in separate system in a different division and can be accessed in-house only. DDMS is therefore an archive of geospatial data and not a management system of geospatial data that provides information about raw capture, processed data and the production of GIS-ready data. There is no other system existing in OS to get raw, processed data and GIS-ready information into one management.

In Malaysia, the move to create a Digital Cadastral Database (DCDB) and the National Topographic Database, initiated by JUPEM in the mid 1980's, had put in place the foundation for the development of GIS in the country. Since then, various federal and state lands related agencies have undertaken their own initiatives to harness the power of GIS and have developed computerised systems to expedite the processing of land related matters. However, these standalone systems, which contain valuable land information, exist as 'islands of information system' thereby negating the vast potential offered by modern networking technology and making it difficult for users of land information to get access to them.

JUPEM is comprised of many sections that capture and provide different type of geospatial data. The responsibilities of these sections can be categorised as follows:

- Computer Assisted Mapping System (CAMS) deals with the digital mapping of all topographic survey carried out in the field and the digitising of hardcopy maps and aerial photographs for photogrammetrical operation. Digital topographical map stored in CAMS are derived from sources of aerial photography, hardcopy maps and topographical survey accomplished by three responsibilities as explained consecutively below.
- Topographic Field Data Capture deals with the survey of all details and elevation to implement digital topographical maps. The responsible sections supply field survey notes and digital data to the CAMS after correcting and final checking for quality control.

- Aerial Photography deals with the exercise of planning and carrying out flights and aerial photography. This division develops the photography and hands them over to CAMS for scanning, ortho-rectification and triangulation via control point set up. This section implements a traditional file processing system for data and flight photography management.
- Geodesy Section is involved in the production and establishment of GPS station around the nation, maintenance of Malaysian Active GPS System (MASS), gravitational measurement for GPS correction, precise levelling and tidal observation, as well as providing dual frequency GPS data from various locations to users for post processing applications. This section supports CAMS in term of adjusting and supplying X, Y and Z real world coordinate.
- Cartography Section deals with the production of atlas maps and e-map for public use over and above what CAMS has covered.
- Cadastral Field Data Capture deals with the production of cadastral certified plan and cadastral base map in each state. Each state manages its own Digital Cadastral Database System. There are 11 states that carry out cadastral survey over the nation and each has its own local coordinate system which if superimposed with CAMS's product will show great discrepancy.

All sections in the Cadastral Division and Mapping Division have their own system and database which are not linked and networked over the Intranet. Access to data is restricted to in-house divisional staff with security username and password. As GIS is just about to embark fully into JUPEM, there has been a futuristic project to provide a single portal of all survey and geospatial datasets. This has yet to be realised because of other significant initial investments and political stress to leverage proprietary data through this technology. The other problem is an initiative to implement coordinated cadastral system (CCS) which is yet to be tabled. States' reference coordinate systems are localised. Without the realisation and implementation of CCS, the national base map and cadastral base map retain two separate different spatial database systems due to their different spatial orientation and location. In reality, a single portal management of all these important geospatial data with the same spatial reference within the department is a critical issue so that the huge amount of information about geospatial data capture, processed data and eventually GIS-ready information can be handled to the maximum benefits.

2.11 GIS Interoperability

GIS-ready information is not completed in its real sense if it is not possible to access the spatial object online. Advances in technology concerning Internet particularly in multimedia and visualisation approaches generally pose great potential for the delivery of the GIS data in real time or near real time (Zhang et al., 2005). This potential brings with it both conceptual and operational issues that require consideration and collective, long-range planning in order to effectively implement the developing technologies. Delivery online provides the sharing of data among staff, organisation and the public.

Geospatial data and information interoperability emerges increasingly in the GIS community over the last decade (ESRI, 2003a). The move and transfer of data within GIS needs standards for interoperability especially when the geospatial data has been transformed to be fully GIS-ready. GIS interoperability revolves around the distribution of geospatial data through Internet services which benefits many communities in the spatial information and communication age. Interoperability has advantages to companies and government organisations in term of reduction in data duplication, definitive data and costs savings.

In various GIS applications when coupled with Internet, GIS interoperability is a new concept and is an on-going research. GIS interoperability is sought by users as important but it has been difficult for system developers (Laurini, 1998). It is a long term ambition to be achieved and federating geographic databases can be seen as a preliminary step towards full interoperability (Kraak, 2004). By geospatial data interoperability, inter-departmental benefits of sharing standardised information via the Spatial Data Infrastructure (SDI) website can be appreciated. The dynamic of the whole SDI concept is principally reflected by referring to all important issues regarding interoperability (Rajabifard et al., 2002).

Worboys and Duckham (2004) refer to GIS systems as interoperable when they have the ability to share data, information and processing. To share spatial data means to transfer them. Initiatives such as SDI are based on particular transfer formats to ease data sharing. There are so many different standards for geospatial data that transferring and converting them pose barriers to interoperability (Bishr, 1998). Standard organisations such as Open Geospatial Consortium (OGC) are coordinating their efforts in an attempt to minimise such barrier. In OGC, specifications are tabled with core intention that all data model and features are to be defined and manipulated so that they could provide a reference point that will allow data exchange among heterogeneous systems (OGIS, 1998). However standardisation does not address the problem of how to convert existing data into a standard format and how to integrate data from various sources. So, a software system needs to be developed to support data interoperability. This has been an existing research issue related to Internet GIS.

Another approach to data interoperability is database integration. This has been a popular, most sophisticated approach in the late 20th century with much current research activity. Database integration is a process of merging existing data sources into a single, uniform, non-redundant data framework which serves a GIS application (Devogele et al., 1998). An initial part of database integration can be carried out by splitting new applications into pieces which are tailored to access one data store and to pass the local data for global application use. A simpler alternative way is to extract the data of interest from the local sources and copy them into a new single database which is linked to a new application.

Geography Markup Language (GML) has become important in the Open GIS concept and is used in major GIS systems. Geospatial information exchange is concentrated on the use of the simple XML-based encoding standard of GML. GML basically uses XML tags to describe geographic feature properties and the geometry (OGC, 2003). GML is widely used because it provides both a vendor neutral as well as implementation neutral format that are optimally suited for distribution over the Internet (Lake, 2001). Plain XML files or compressed ones can be streamed and the user is not troubled with downloading an entire file before opening. This greatly enhances usability in a networked or Internet environment.

2.12 Chapter Summary

This chapter has discussed and reviewed several subjects and issues regarding geospatial data management, the geospatial technologies used to produce geospatial data, object-based GIS as the significant modelling technique and surveying aspects related to geospatial data. Some knowledge and developments of GIS are described. A brief examination of JUPEM is put forward which mainly drove this research. Types of surveys and datasets existed in JUPEM are examined as the significant matter to be dealt in the research. Reviews of the geospatial data management practices in three geospatial data collecting organisations are drawn upon. An aspect of GIS interoperability that is necessary for data distribution and sharing is examined.

Discussion content has shown the need for geospatial data management from geospatial data capture to the creation of GIS-ready information due to huge data quantities, scattered storage and old filing systems. Geospatial data capture is about surveying the information on the earth and is considered very close to GIS because of the spatial dimension, location and geo-referenced system. A review of object-based GIS and its relevance to the geospatial management was made. Object-based GIS technology seems very suitable because it can describe the actual object, its attributes, operation and properties that can be simply visualised. Geospatial data capture towards the production of GIS-ready information can be divided into three object entities namely:

- Geospatial data capture.
- Processed data.
- GIS-ready information.

These object features are expected to be used as models in the GIS environment. Coupled with relational model for database, these objects can be layered in relation to each other for query and management. In these so-called feature datasets, there are sub-features containing surveys, survey files, CAD files, TIN model, geo-referenced raster coverage, GIS objects (building, road etc), and more features within raw survey, processed datasets right through to the production of GIS-ready information.

The texts on geospatial data and technologies have also shown that geospatial data, surveying and geospatial technologies are in Earth space, having boundaries and with attributes and operation. This shows the relevance for GIS technology and object-based modelling.

The description of knowledge and development of GIS showed the significance of GIS tools in handling all kind of reality of the world including ability to manage projects and illustrate graphic of object and relation between them. A GIS environment provides management tools for various geospatial data as long as they are related to the surface of the Earth and have location for understanding of space.

Review of the practices of geospatial data management in three countries offered the situation where these organisations only centralised on the production and existence of geospatial data for sale and access without taking much consideration on archiving the information for raw data acquisition, capturing, surveying and processing towards the resultant GIS-ready information. There is no management of data and information from the geospatial data capture to the provision of GIS-ready information in any of the two developed countries and one developing country.

Interoperability in GIS is a must for GIS-ready information to be well informed to the community. Geospatial information is not only for internal use but for the external part of the organisation and the public. The emergent WWW and GML allow cheap and vendor-independent Internet GIS system be developed.

In general, there is a requirement that a solution to the linking of the various geospatial data within the domain of geospatial data capture right through to the production of GIS-ready information be realised. These geospatial data capture, processed data and GIS-ready information entities and their 'attributes' which are all data and information within them, spatial and non spatial, should be harmonised into one single portal management system which is manageable for access.

The next chapter discusses the basic concepts of object technology and the object-oriented database model as well as UML. They serve as the background to the design model.

Chapter 3

Object Models, Database Modelling and Unified Modelling Language

3.1 Introduction

Chapter Two has discussed geospatial data management and technologies, knowledge and development of GIS, review of geospatial data organisations and geospatial data interoperability. Real world objects are becoming increasingly complex though are rich in location information, space and distance considerations. Due to their richness and complexity, technology and techniques on how to handle geospatial data has been an central in the GIS researches nowadays.

The basic concept of object representation in GIS, database modelling and UML is described briefly in this chapter. This is crucial for the implementation of the management system using GIS for the datasets from raw capture right through to the GIS-ready information. The last part summarises the chapter.

3.2 Object, a Real World Package of Information

What is an object? Generally, an object is an entity with a localisation represented by attributes and by a group of operations. However a more formal definition as described by Longley et al. (2002):

> *"An object is a self-contained package of information describing the characteristics and capabilities of entity under study"(page 197)*

Characteristic, a synonym to attribute, means a typical quality of something, and capability means able to do things effectively and skilfully, and so modelling by object excellently presents the whole concept of world entities and their operations. Basically, an object has attributes and operations. Object is defined by its entity.

The real world is like a container filled with a collection of objects and the relationships between the objects. An object is an entity in the real world that is included in GIS. An integrated, organised unit of data stored in a project file or a database including vector, raster, surface model and text are forms of objects (Bernhardsen, 2002). A set of objects of a similar type is known as a class. In reality, classes are a more vital perception than objects from the implementation point of view (Crowther and Hartnett, 2001). When creating an object data model, the data model creator specifies classes and the relationship between classes. Objects, that are instances or examples of classes, are actually created when the data model is used to create a database.

To constitute an object, an entity must have identity, relevance or be of concern to the application realm and describable i.e. have attributes or characteristics. In object modelling, an object is referred to a digital representation of a discrete spatial entity of the space of the real world. By means of the modelling process, each entity relevant to an application domain is represented by a corresponding object in the data model. The object in the model should have properties that describe the characteristics of the corresponding entity in the universe of communication. Each object is defined by three properties (Figure 3.1): its identity (which distinguishes it from other objects), its attribute (that describes the fundamental characteristics of the phenomena involved) and its operation (that the object performs in the application domain). These are described below.

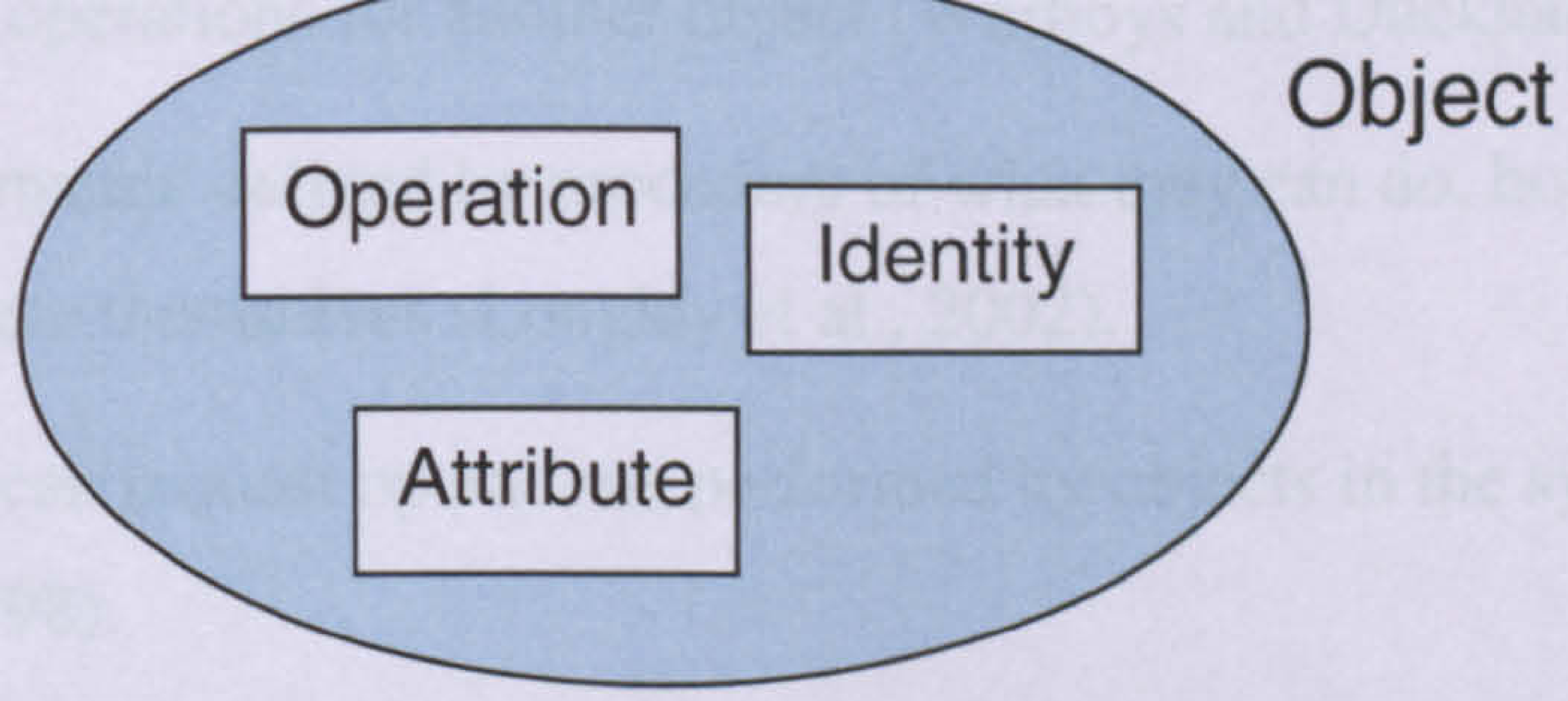


Figure 3.1: An object has identity, attribute and operation

3.2.1. Object Identity

A system oriented to object representation has objects which are unique. This uniqueness of an object is reached by means of the object identity. Object identity belongs to an object and identity uniquely distinguishes it from all other objects (Egenhofer and Frank, 1992). By introducing a unique identity for each object, different objects can be distinguished from each other without the need to compare their attributes and operation (Kim, 1990). The object identity is usually system generated, unique to that object, and never changes for the object lifetime (Loomis, 1995).

3.2.2. Object Attribute

The state or attribute of an object is described by the values of its properties at one moment in time. An object attribute is actually a named property of the object class that describes a value held by each object of that class. An object can have a single state throughout its lifetime, or it can go through many state transitions. When an object moves from one state to another, we talk of a new version of the object.

3.2.3. Object Operation

Operation is led by the behavioural trends of an object as the result of dynamic force and interactions with other objects. In object-oriented systems, the term operation or dynamic behaviour has been utilised to illustrate several characteristics of object trends. These characteristics are as below:

- Objects interact by sending messages to each other that activate certain behaviours or operations for another object (Worboys and Duckham, 2004).
- Objects are dynamic defined by procedure of what they can do, how they can create and delete themselves (Longley et al., 2002).
- Other objects can request operations performed by objects in the system (Molenaar, 1998).
- Objects are characterised by methods that refers to an operation on the data, a procedure which can be applied to a class of objects (Laurini and Thompson, 1995).
- An object possesses dimensions which are semantically valid operations that are applicable to the individual objects and these are highly application specific (Kemper and Moerkotte, 1994).

- An object acts and reacts, in term of its state changes and message passing (Pooley and Stevens, 1999).
- An operation in an object model is a function or process that may be applied to or by objects in a class (Blaha and Premerlani, 1998).

The linking of identity, operation and attribute constitutes the approach of an object-oriented system which is discussed in the following sections. The knowledge and concepts dealt with throughout the content of this chapter are from the standpoint of an object-oriented system in a broad term.

3.3 Object Orientation

The foundation of object orientation is the concept of an object. Object orientation can be portrayed as an approach to depict interacting objects with its attributes and behaviour (operation) within an organised system (Rigaux et al., 2002). These objects have identities that can be differentiated between each other. For the object-oriented approach, the key notion is that:

object = identity + attribute + operation

Object-oriented technology is based on the assumption that the real world can be modelled as distinguishable entities or objects that can be grouped together into classes. As discussed earlier these objects are used as a model of reality with the help of identity, attributes and operation. The object oriented paradigm has been applied to many technology areas including object-oriented programming, object oriented analysis and design, object oriented databases and object oriented graphical user interface design. These areas bring considerable advantages to the successful implementation of a GIS, especially object-oriented programming languages which provide the framework for implementing the others. The benefits that come from the object-oriented technology are many, including extensibility, reusability, reduced complexity, eases of use, etc (Gartner et al., 2001; Dollner and Hinrichs, 2000).

As explained earlier, the basis of this approach is performed by the observation of the objects in the world that are perceived to have interactions. The interaction among the objects can be seen as a command or message given to an object either verbally or by some action such as physical force (Egenhofer and Frank, 1992). Based upon common operations that can be applied to the object, or command to which they respond, they are grouped into object classes. This constitutes efficiency in the use of objects in geospatial data modelling because multiple representations can be produced (Figure 3.2).

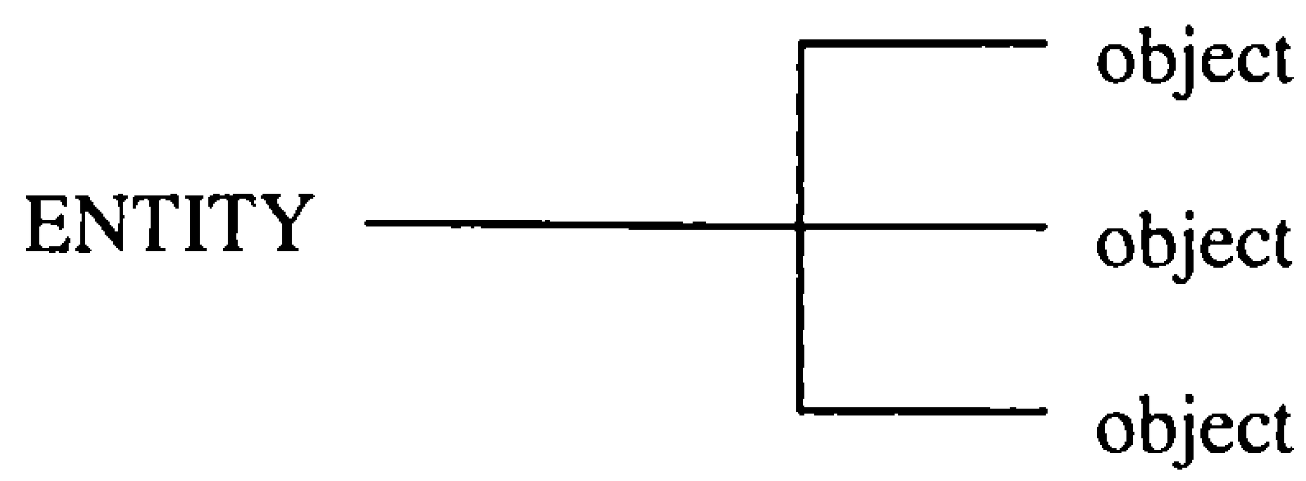


Figure 3.2: Several objects embody a single entity (i.e. there are many relationships between entities)

Object orientation has become an important design method in software engineering. It focuses, as a first line of structuring on modelling objects as a human perceives them in reality. Unlike previous approaches of modelling, it combines the modelling of the structure and the behaviour of the objects. Procedural abstraction primarily models the operation, while methods use for designing the database schema concentrate on the structure of the entities (Capretz, 2003). Object technology has been applied to many realms so that there are several alternative interpretations of the underlying concepts leading to object orientation. Different environments have interpreted different concepts and fundamentals of object orientation. The following section explains this.

3.4 Fundamentals of Object Orientation

An object has the basic concepts of object identity, attribute, and operation, whilst object orientation considers further aspects of objects, their interactions, relationships, and organisation. These provide the fundamentals for object orientation modelling and design. In the following sections, the fundamentals or key features that make object data models especially good for modelling geographic systems are outlined.

3.4.1 Abstraction

Abstraction is when vital parts of an object or group of objects are extracted while ignoring the unimportant details for the time being. Abstraction denotes a model that includes the most important, essential or distinguishing aspects of something while suppressing or ignoring less significant, immaterial or diversionary details (Booch et al., 1999). The process of forming a conceptual model of a part of the real world during database design is an abstraction process, since only those entities and their properties that are relevant to the application domain are collected. The resulting model is also an abstraction of the real world. Abstraction is one of the fundamental ways human cope with complexity.

3.4.2 Classes

Abstraction means taking the important part of an object while leaving the unessential part for the time being. Classification is the most crucial and well- recognised form of abstraction. An object is an instance or occurrence of a class. Objects of the same type are grouped into classes. Classes provide a mechanism for sharing across similar objects. A class is a description or specification of a collection of objects with the same attributes, common behaviour (operation), similar relationships with other objects, and common semantics (Blaha et al., 1988). Mattos et al. (1993) state three concepts that are associated with the use of classes. The first concept illustrates that a class gives the structural definition for the instances of that class, that is, the names and types of their attributes and methods. Objects being instances of the class, inherit the attributes and methods specified for the class. Second, a class serves as a template for creating new objects. New objects are created as instances of the class; therefore they have access to all the attributes and methods of the class to which they belong. This is called inheritance, and is discussed later. Thirdly, each class is associated with a group of objects, which comprises its instances, and the class has to control these instances.

3.4.3 Encapsulation

Encapsulation states that each object packages together its description of its attribute (state) and operation (behaviour) (Longley et al., 2002). This means that each object contains the attributes and the operations that will manipulate the object, and that the object's attribute are accessed only through its operation invocation. Object-oriented models thus impose encapsulation and information hiding (Kemper and Moerkotte, 1994). The organisation of the attributes is known only to the objects themselves, and is not accessible from outside the object except through a specific and carefully designed interface. Encapsulation thus separates the internal implementation of an object from its public interface. All other details concerning the objects are hidden behind the interface. Users of the object access the object's attributes only through the methods specified in the object's interface. Users will not know the internal implementation of the object, i.e. how its methods are executed. An important benefit of encapsulation is that since the internal implementation is totally hidden within the class, the implementation of the class can be changed without affecting the other objects in the system. In other words, the implementation can be extended and modified without affecting the users of the class.

3.4.4 Inheritance

Inheritance allows classes, and so the object, to share common properties. Inheritance acts in two ways: generalisation and specialisation, that is either top-down or bottom-up ways (Capretz, 2004). A generalised class will exhibit properties common to all its specialised classes. These specialised classes are known as subclasses. Similarly, classes specialise, adapt and add to the properties of generalised classes. As generalised classes sit on a higher level they are called superclasses. When inheritance is applied to classes, inheritance specifies that if an object is of Class B where Class B inherits from Class A then that object can be used just like an object of Class A. Class B is said to conform to Class A and all objects that are Class B's are also Class A's (Fussel, 1997).

Basically, inheritance is the concept of object orientation which provides the means to define new classes in term of existing classes. The new subclass inherits the characteristic and behaviour of its parent class or classes. The parent class in turn is a superclass of the child class. The superclass is more general than the subclass. The inheritance hierarchy does not have to stop at two levels. The subclass can be a parent for still another class, and so on. By this way we can construct an inheritance hierarchy or tree as shown in Figure 3.3.

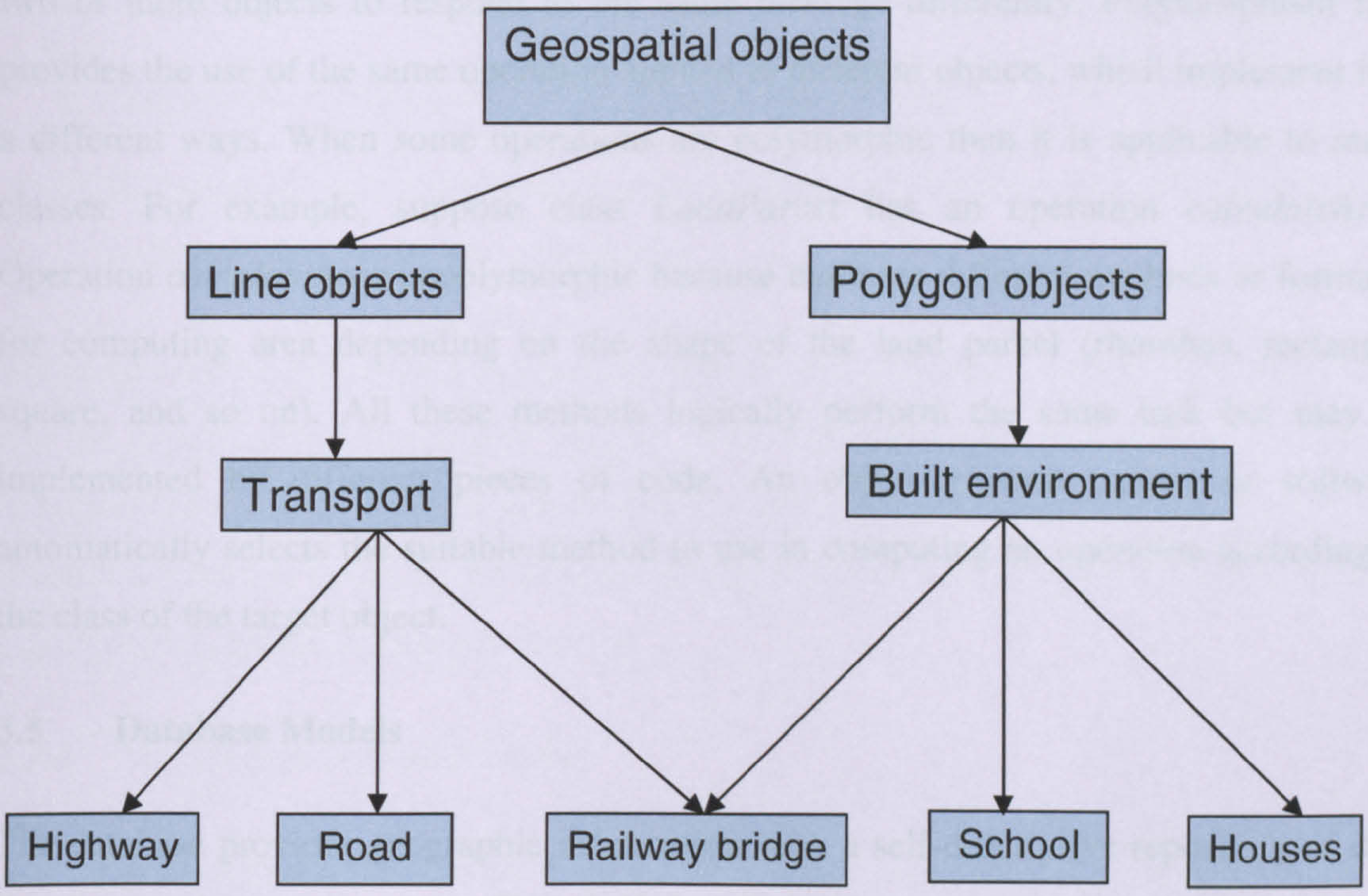


Figure 3.3: Single and multiple inheritances of object classes

Inheritance can be single or multiple. With single inheritance, a subclass may inherit attributes and methods from a single superclass, usually adding its own attributes and methods. Multiple inheritance means a subclass inherit from multiple superclasses. As in Figure 3.3, consider a railway bridge being a subclass of both Transport and Built Environment superclass. Inheritance allows code be reused and offers more flexibility for developing systems for specific applications (Pooley and Stevens, 1999). When it applies to classes, inheritance specifies that a class uses the implementation of another class with possible overriding modification (Beeri et al., 1999).

3.4.5 Polymorphism

The word polymorphism derives from Greek, to do with having many shapes (Booch et al., 1999). This refers to the constitution in which an entity could have any one of several types. Polymorphism means the ability of a single object to refer to different object having different forms and stages (Steimann, 2000). In an object-oriented system, this mainly refers to the fact that an operation with the similar name can be applied to different object, which may implement that operation differently. Polymorphism allows two or more objects to respond to the same message differently. Polymorphism thus provides the use of the same operation applied to different objects, which implement it in a different ways. When some operations are polymorphic then it is applicable to many classes. For example, suppose class *LandParcel* has an operation *calculateArea*. Operation *calculateArea* is polymorphic because there are different methods or formulas for computing area depending on the shape of the land parcel (rhombus, rectangle, square, and so on). All these methods logically perform the same task but may be implemented by different pieces of code. An object-oriented system or software automatically selects the suitable method to use in computing an operation according to the class of the target object.

3.5 Database Models

The database provides geographic information with a self-descriptive repository of data that is stored in one or more files. Self descriptive is what sets a database apart from ordinary files. The premise of the database is that the data structure is expected to be relatively static while the actual data may rapidly evolve. Database systems are normally classified according to the data model they use. There are old models called hierarchical and network data models that are now obsolete and not use for new applications. They are being replaced by three models known as relational, object-oriented and object-relational. The next few sections briefly explain the detail of these models.

3.5.1 The Relational Model

The relational model is based on a collection of tabular relations, often just called tables, that have simple structure (this is what makes it so powerful) (Worboys and Duckham, 2004). A relation has associated with it a set of attributes. The data in a relation is structured as a set of rows. A row, or tuple, consists of a list of values, one for each attribute. Each cell contains a single attribute occurrence, or value. The tuples of the relation are not assumed to have particular order. A relation schema is a set of attribute names and mapping from each attribute name to a domain (Guting, 1994). The relational schema gives the structure of the relation, and does not include the data. A relation is a finite set of tuples, which possess properties as below:

- The ordering of the tuples in the relation is not crucial.
- Tuples in a relation are all distinct from one another.
- Columns are ordered so that data items correspond to the attribute in the relation schema with which they are labelled.

Relational model is designed as a way of representing data into relation. This model also has the aspects of operation that may be performed on the relations (such as database manipulation) and the integrity constraints that the relation must satisfy. Intrinsically, the relational model is noted for its primary key and foreign key, which are used to provide links between tuples in different relation.

3.5.2 The Object-Oriented Model

The object-oriented model for a database system uses the combination of object orientation and the functionality that is available in a database (Gartner et al., 2001). An object-oriented database is regarded as a persistent store of objects created by an object-oriented programming language and is embodied by having an object-oriented logical data model (Egenhofer and Frank, 1992). The basic unit of data in this database system is an indivisible object. These objects have identity, attribute and behaviour (operation) defined through methods, all encapsulated within the object itself. The object-oriented model is intended to handle more complex application requirements, which the relational model cannot manage.

3.5.3 The Object-Relational Model

The object-relational model is an extension of the relational model. Through the compromise between relational and object-oriented model, object oriented features are incorporated into relational database or (built on top of it) in a manner that uses the power of object orientation while maintaining the full functionality of the relational model (Fussel, 1997). Presently, many database management systems and GIS systems are beginning to include object-oriented features into the existing relational database functionality. For example, Oracle 8 introduced this concept to overcome the limitations of the relational model. ESRI's recent model called the geodatabase fully combines object-oriented concepts with the relational model.

An object-relational database allows for user-defined object types, which makes the database more understandable. Object types can be used to map an object data model directly to an object-relational database schema, rather than restructuring the data model into the flattened row-column format of relational tables in a purely relational database. An object-relational database allows for the use of user-defined object types in application programs that access the database, which makes using the database more intuitive. Application programs can retrieve and manipulate the data as objects and call procedures that use the methods of the object type to perform operations on the object. Since the methods can be stored in the database, data-intensive procedures can be more efficient. Objects can be reused, which makes application development faster and more efficient since the use of objects relieves developers of the need to write a mapping layer between application program objects and database objects. The use of objects, based on the underlying software engineering principle of data abstraction and encapsulation, also makes it easier to understand application program code and to maintain application programs.

Basically this model is utilised for the implementation of database system in this research. The object-relational model has been adopted herewith because it makes use of the power and semantics of object orientation and the full functionalities of the underlying relational database system.

3.6 Geodatabase

Throughout the years, databases for GIS were storage of the spatial and attribute components of geographic data in separate files linked together by unique identifiers. The data model utilised often isolates the graphic data from the attribute data. In the old ArcInfo coverage, for instance, using the georelational model, the spatial data is stored in indexed binary files, and the attributes are stored in separate files (INFO tables). Recent advances in hardware and software engineering have now made it possible to combine GIS and database systems to create geodatabases. Geodatabases or 'geographic databases' keep both spatial and attribute data together in a single database management system. The spatial data is stored just like an attribute in the database. The concepts of relational and object-oriented model are integrated to enable the enhancement and possibility of both spatial and attribute data storage in a single database management system. Geodatabase has been acknowledged to provide 'gear and tool' to varied and extensive resources and developments across the globe (Arctur and Zeiler, 2004). An example of geodatabase model that holds basic objects is as shown in Figure 3.4.

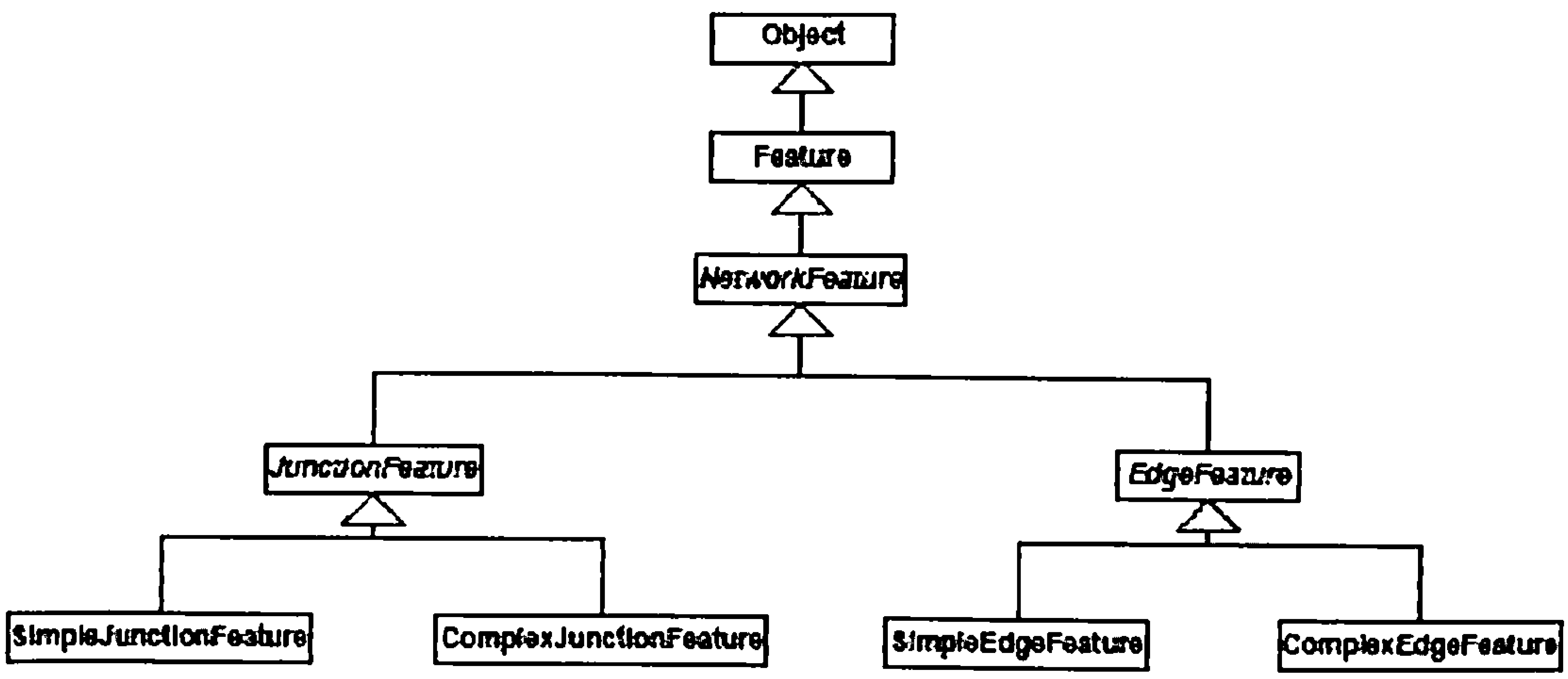


Figure 3.4: Basic objects in the geodatabase model (Zeiler, 1999)

3.7 The Unified Modelling Language

The Unified modelling language or UML is used to design the conceptual and logical model for this research. An introduction to UML is briefly explained in this section along with the conventions used in this thesis. UML is a language for specifying, visualising, constructing and documenting the artefacts of software systems, as well as for business modelling and other non-software systems (OMG, 2004). UML is a strong modelling language that has solid semantics and notation definitions. UML is widely accepted in software applications, especially in most GIS software such as ESRI's ArcGIS which directly supports it. UML contains a number of graphical elements which when incorporated will form models that provide manifold views of a system. It is important to know that UML models describe what the system is supposed to do and does not show how to implement the system. In this section, only the relevant elements of the UML class diagram used in this thesis are briefly described.

3.7.1 Objects

In the UML notation, an object is shown as a rectangle with two partitions. The top partition shows the name of the object and its class, all underlined. The second partition shows the attributes for the object and their values as a list (see Figure 3.5).

3.7.2 Classes

Classes are symbolised in the UML by a solid outline rectangle with three partitions separated by horizontal lines. The top partition defines the class name and other general properties of the class (e.g. stereotypes); the middle partition defines a list of attributes; the bottom partition describes a list of operations.

3.7.3 Interface

A class interface is an 'instructor' for the externally visible operations of a class. The interface of a class specifies the set of operations that the class presents to other classes. An interface is represented in the UML as a rectangle with two partitions, the top showing the interface name with the keyword <<interface.., and the bottom part showing the list of operations. Refinement or realisation indicates by dashed lines shows the relationship between a class and its interface (Figure 3.5).

3.7.4 Association

Association is represented in the UML as a line connecting two classes with the association name just above the line. This kind of association is called a binary association in UML. The association role is shown at both ends of the line next to the class. Association cardinality or multiplicity is shown just above the association line near the appropriate class. An association may have attributes and operations just like a class. In such case, it is called an association class. This is represented in the UML like an ordinary class with a dotted line connecting it to the association (see Figure 3.5).

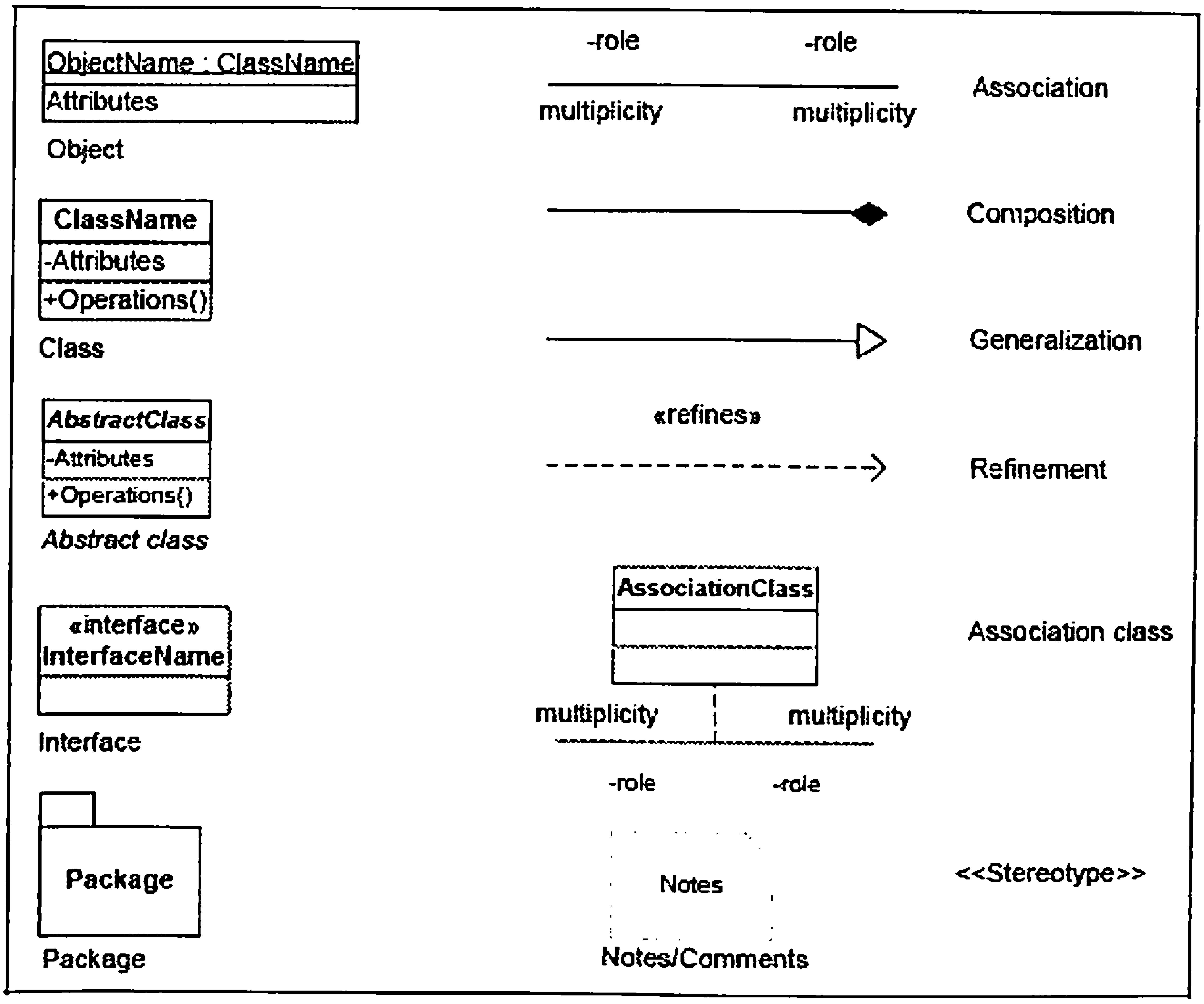


Figure 3.5: UML static structure elements (Tilley and Huang, 2003)

3.7.5 Generalisation

Generalisation is a relationship between a superclass and its subclasses. In the UML, generalisation is represented by a line that connects the subclass to the superclass, with an open triangle on the end of the line that point to the parent class.

3.7.6 Composition

Composition shows a relationship between a composite object and its basic objects. Composition may be shown by solid filled diamond as an association role adornment. The multiplicity of the composite end may not exceed one, i.e., it is unshared.

3.7.7 Abstract Class

Abstract class is not meant to create new object. Its properties are inherited by its subclass. It cannot be instantiated but is used to specify requirement for other classes to be provided by another code, e.g. a survey monument could be an abstract class for a numbered survey monument and temporary survey monument classes. In a class diagram, an abstract class is shown by its title in italics.

3.8 Chapter Summary

This chapter has explored three different technologies: object technology, database modelling for management systems and the Unified Modelling Language. The UML is covered as a base to derived conceptual and logical models. All these are significant to the successful implementation of a geographic information system. Both object technology and object orientation is an important key in enhancing the design and implementation of GIS. The database is the foundation of the GIS system. Three main database modelling concepts were described as input to database design intended in this research. Geodatabase was introduced and it was described how it incorporates these concepts to develop a system for geospatial data management.

The next chapter deals with JUPEM's survey practices and datasets, data models, data management approaches, concepts of geographic database technology for modelling raw captured data to the stage of producing GIS-ready information and delivery of GIS-ready information.

Chapter 4

Geospatial modelling for management and delivery of survey datasets

4.1 Introduction

In the preceding chapter description focussed on the object, database modelling and UML technology which are the underlying foundation of the geospatial data technology used in this research. This chapter explains data models, and the method of representing the real world. In addition, conceptual, logical and physical modelling are described. The data management approaches are discussed leading towards description of databases and DBMS. The geographic database as a database to manage geospatial data based on the object within an object-relational database is described in relation to the research.

Before proceeding on the data modelling, description of the existing work flow of surveys implemented from raw to the production of GIS-ready information in JUPEM is made in the next section to consider possible influence of its nature to the modelling technologies.

4.2 Existing Work Flow of Surveys in JUPEM

Historically, JUPEM began in 1885 with the initial establishment of state surveying departments and some trigonometrical survey departments in the states. The British who colonised Malaya strengthened its function with Colonel H.M. Jackson appointed as the first Surveyor General in 1908 (JUPEM, 2004). State surveying was concentrated in the creation of land title while trigonometrical survey led to the establishment of further topographic mapping activities. Eventually, JUPEM's main activities became cadastral surveying and topographic mapping (see Figure 2.6). These surveys are explained in the following sections.

4.2.1 Cadastral surveying

Cadastral survey originated from the Torrens System of land administration that was introduced in Malaysia in 1891 by the British. Cadastral surveying, which yields land title for a piece of land (parcel), involves the fieldwork of collecting survey information such as bearing, distance, coordinates, acreage, lot number, and detail such as disputed buildings within a land parcel, connections lines, boundary marks, sections and district administrative boundaries. Cadastral survey systems are implemented in every state according to the local coordinate system, Cassini Soldner Reference System. Surveying is carried out as the requisition for survey (RS) is received by JUPEM from the Land Office. According to the National Land Code 1966, normally four type of land are surveyed and resurveyed by JUPEM for cadastral purposes and recorded as well as updated in the national cadastral map base or sheet. The first is vacant or state land alienated by the state land office, secondly is the land after being approved for subdivision, partitioning and amalgamation, thirdly is old land title to be resurveyed for verification and upgrading, and lastly is land involved in an acquisition by government for public use.

An RS, which contains preliminary sketch of survey site and authorisation documents from state land administrator, is therefore produced to require surveys to be carried out by JUPEM. It is likely that RS and Land Office document paper, as a valuable source of raw information about the survey, be managed in GIS. RS and Land Office paper document can be scanned and form raster data and can be kept in JUPEM database. It is hoped that RS and the Land Office document paper can be sent electronically through the Internet to JUPEM where they can be stored instantly once accepted and served as part of source data for the survey.

Field surveyor equipped with instruments, workers and previously referenced survey information, implements the survey. Currently, there are three ways of survey being carried out. They are traditional survey which involves bearing recording and EDM measurement, total station surveying which uses data logger to capture the observation, and GPS survey. Traditional survey is carried out in some states when field surveyors are not concerned about using computer to compute and plot the survey plan. By total station surveying, survcy data can be downloaded into the computer in the office. This type of survey has currently been upgraded towards the concept of field-to-finish (F2F). Many JUPEM states have recognised F2F concept. Cadastral survey using GPS has been proposed to be used by field surveyors, but yet to be realised fully. GPS survey also provides the concept of F2F.

Traditionally, field book is used to record bearing, distance and details in the field observation. In the field book, there are observation pages (including astronomical sun observation for azimuth control) and diagram page. Connection lines and traverses are sketched, not to scale, in the diagram page. Monument to monument parcel boundaries are computed and sketched in the diagram page. Computation sheet is produced to show the final product of field work and the accuracy of the survey. Field book and computation sheet can be priceless survey source information that can be managed in GIS as raster scanned data. Using the detail in the computation sheet, certified plans are drawn and deposited in the strong room after the issuance of title of the surveyed land. Certified plan can be stored in the database by converting the hardcopy to digital scanned image. Field book, computation sheet and certified plan form valuable source data that can be archived using GIS. It is essential to access these source data in GIS especially when the end product, GIS-ready information need to establish their metadata or parcel polygon has changed due to amalgamation and subdivision.

Nowadays most of the state JUPEM departments are implementing the field-to-finish concept of surveying which produces information right to the stage of processed data. Total station surveying with data logger and GPS survey are used to carry out the cadastral survey in most states. The produced survey files are ASCII files and GPS data file. The processed data are in CAD files (in DXF and DGN) and Shapefile, downloaded and enhanced from the ASCII and GPS files before they can be converted to GIS-ready information. Data entry (keyboard) process has been the practice for old cadastral certified plan that were drawn manually. The data entry leads to the production of Shapefiles. In cadastral survey package, the end products, GIS-ready information should be the data and information that are in points (monuments or boundary stones), lines (land parcel boundaries) and polygons (land parcel, cadastral map sheet and administrative boundaries). GIS-ready-information may include the title deeds (scanned) of the land parcel as attribute that can be stored as raster image. The cadastral GIS-ready information should be disseminated to the public and other government departments who need cadastral certified plan and map sheet as backdrop to their geospatial data. Civil engineers, engineering surveyors, architects, planner, local authority and utility engineer, as well as lawyers are some of the professionals that may need the data for their use and development. Figure 4.1 forms the work flow of cadastral survey.

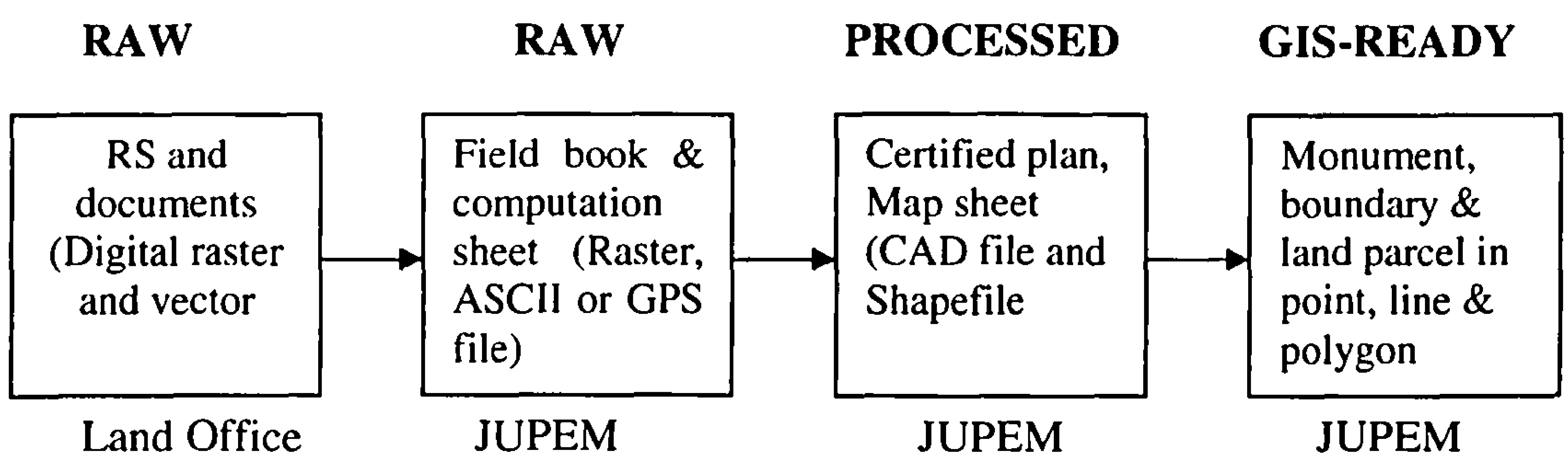


Figure 4.1: Work flow of cadastral survey in JUPEM

4.2.2 Topographic Mapping Survey

In JUPEM, topographic mapping activities for the purpose of producing hardcopy and digital map, and national base map include plane tabling, air survey for digital mapping (photogrammetry), topographic survey using total station, GPS survey, levelling, gravity reading and MASS (Malaysian Active GPS System). Presumably, these activities produce processed data and GIS-ready information for dissemination to the public. Plane tabling is not actively implemented unless the department is lack of modern instruments in some JUPEM districts. The rest of the activities will be discussed in this section.

4.2.2.1 GPS survey

GPS survey is an advanced positioning and data recording technology available to the land surveyor today. In the Mapping Division of JUPEM, GPS survey has seen the use of field-to-finish concept where all GPS coordinated points data are downloaded and stored in the computer at the end of the day. The data is processed to create CAD files and Shapefile. GPS survey in JUPEM also includes control station establishment for reference to others surveys carried out within the area. The GPS establishments are marked with cemented permanent pillar and include information such as name, location plan, map and its coordinate.

In this survey, raw data are the observation point data of the features in the field such as building, road, slope and others for the national map base and digital map. Coordinates of points are recorded in X, Y and Z (height). These GPS observed data are in ASCII with feature codes that can be converted to graphic data using specific software such as Microstation. Microstation creates processed data in DXF format. These data can be converted into GIS-ready information as point, line and polygon, as well as attributes data of the features mapped (see Section 5.7). GPS survey packages thus have raw and processed data, and its processing steps that enable the production of GIS-ready information. Can this data and processing steps be viewed and managed in one single portal of GIS system?

4.2.2.2 Air Survey for Digital Mapping

Air survey is carried out before the process of digital mapping by photogrammetry. Air survey or aerial photography is the activity to obtain digital photographs of the earth surface by taking photograph while flying at specific height. These photographs are captured in sequence with .the overlapping coverage within 60% between series of photographs. The photographs are then developed and produced as digital and hardcopy. Coverage at the edges of the photographs is not true and distorted. They are then ortho-rectified using the control points established in the ground survey by GPS. Mosaicing is then carried out to obtain coverage of the targeted survey area. Triangulation is carried to enable the right orientation of the image using real world coordinate of the ground control point. The ortho-rectified images are considered as the raw data in this survey which are used for the production of CAD files using the Microstation software. The whole process is mainly called photogrammetry. This raw data can consist of metadata and attributes which includes time of flying, surveyor, GPS ground control used, camera information and name of the photographs produced.

The next step covers the digitising of the mosaic coverage. Digitising is carried out using Microstation and CAD files are produced to show the polylines, closed polylines, points and annotations of the features. At this stage, the raw data are processed (see Section

5.7). The CAD file can be processed, edited and converted to GIS-ready information. In JUPEM, it seems that the raw data is produced in a section away from the section that processed the raw (different building and database or data storage) and GIS-ready information is created in another section. They should be managed and processed within one system that can be viewed by multiple users.

4.2.2.3 Malaysian Active GPS System (MASS)

Malaysian Active GPS Station (MASS) is established to track GPS satellite 24 hour a day continuously. It was established under the 7th Malaysian Development Plan with a budget of RM 5.3 million. Data is stored in file format every hour and transferred to the Geodetic Data Processing Centre in Kuala Lumpur every day. GPS data are updated automatically and are broadcasted in the department Web at 10 o'clock every morning. The data are then updated for the department used and the public are served with the actual timely coordinate information. They are in the file format called RINEX (Receiver Independent Data Exchange) Version 2. MASS stations include 15 locatins situated strategically across the country as depicted in the map below. All stations are equipped with GSP receiver Trimble 4000SSi and antennae system Choke Ring or TR Compact L1/L2.

Figure 4.2: MASS stations across the West and East Malaysia

MASS stations are used for the application of various systems which include surveying and mapping, positioning with a single GPS receiver, integrated GIS system, Coordinated Cadastral System, topographic field completion and international collaboration projects. MASS stations are established with the objectives:

- To enable JUPEM to carry out surveying and mapping functions more effectively.
- To maintain the National Reference Frame System at centimetre accuracy.
- To monitor and provide the user community with a coordinate system that is connected to the national and global reference frame.
- To provide observation data and correction to the public.
- To monitor the vertical and horizontal crustal motion.
- Scientific research on geodynamic.

MASS station data are online data and is used to automatically update the GPS data for the whole country. MASS station can be managed as GIS-ready information incorporated with national base map as points with all the attributes as the metadata, photographs of the station (Figure 4.3) and locational sketch as raster image embedded in the point data. MASS station data can be a source of valuable data for the processing of all map data for the whole country. GIS can be used to access the source RINEX data at a specified time and data in the case for any movement of the earth crust as the result of Tsunamis or an earthquake.

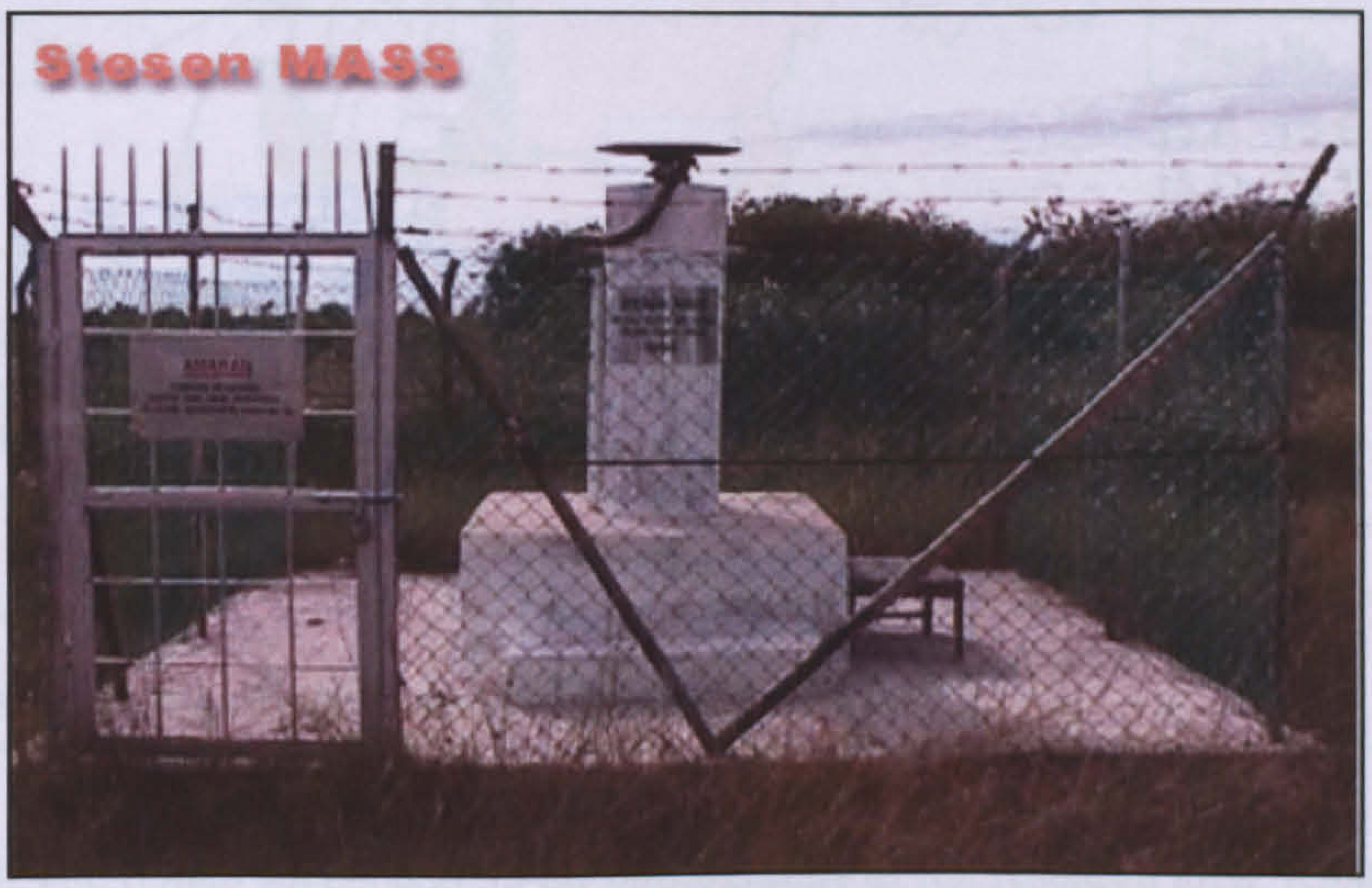


Figure 4.3: A typical MASS station

4.2.2.4 Detail Survey by Total Station

This survey is carried out for the purpose similar to the survey described in Section 4.2.2.1, except that total station is used to do the topographic survey instead of GPS. Topographical surveys measure the land profile and/or the contours of the land and any natural or man made structures or obstacles. Field-to-finish concept is embedded in this survey and data logger is used to get the raw into Microstation software for the processing to produce CAD files. CAD files can be converted into GIS-ready information using processing steps in GIS.

For this research, CAD files of map sheets 15A, 15B, 15C, 15D, 15E and 15F were used to process and convert them into GIS-ready information. The amount of 55 air photographs (research coverage for air surveys as described in Section 4.2.2.2) and the 6 map sheets used for the sample data are depicted in Figure 4.4. These data were sample data for the prototype implementation as described in Section 5.5, 5.6.and 5.7.

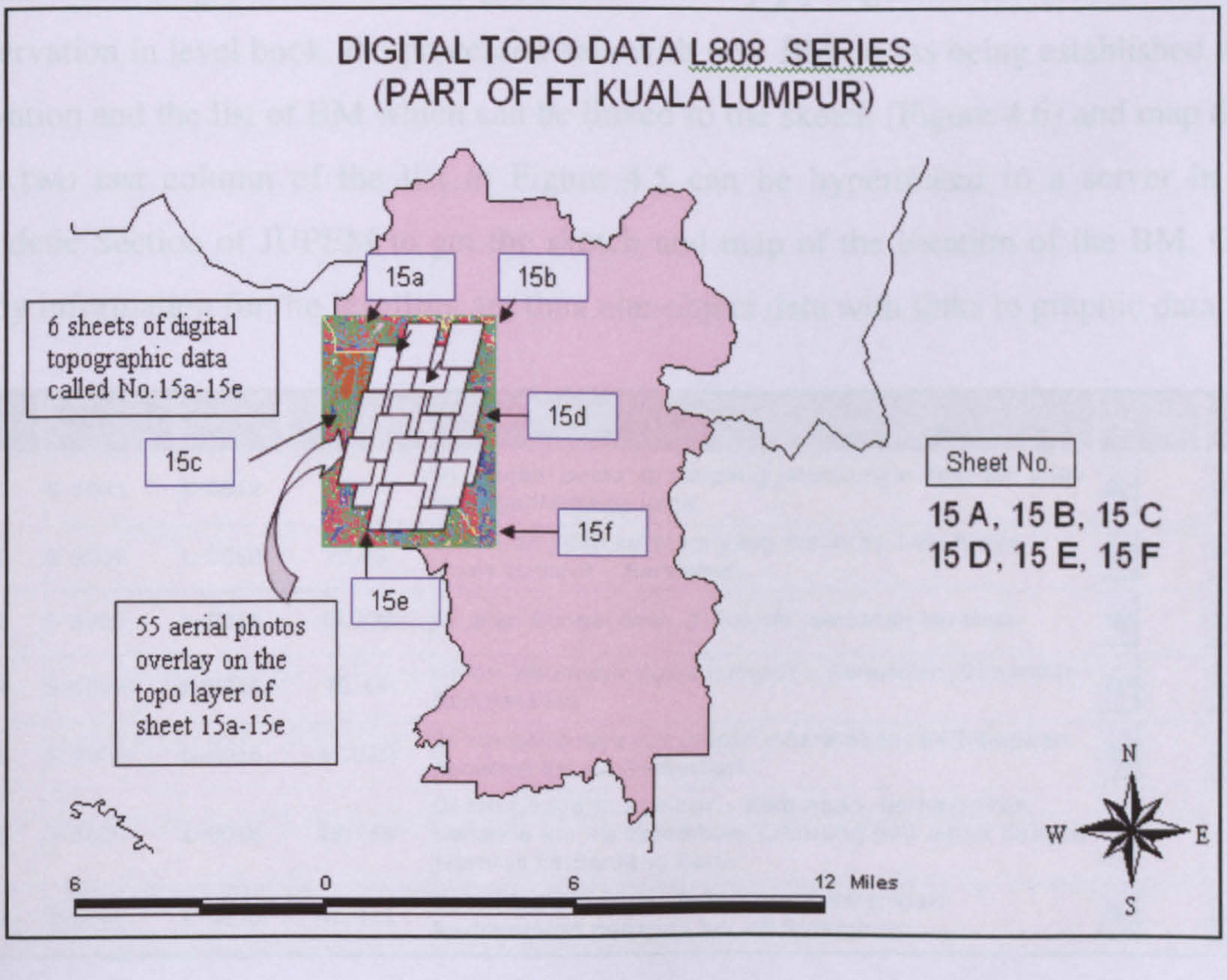


Figure 4.4: The 6 map sheets and 55 air photographs of the research area, Kuala Lumpur

4.2.2.5 Levelling

JUPEM's levelling data is meant to calculate height or elevation value of a point on the ground surface. Its vertical z-coordinate is significant for national infrastructure development for tasks that include survey and mapping, engineering, scientific and hydrography. Height of a levelling mark is derived using two levelling methods, that is, Precise Level Surveying and Second Class Level Surveying. The former confines to a process of measuring high accuracy point height that is used for geodetic task, whilst the latter is a second order accuracy level surveying that is carried out within the precise level network. Precise level surveying is concentrated around main road network and connected to 10 tide gauge stations nation-wide covering 5,400 precise levelling marks.

Levelling network of the Malaysia has existed many decades ago and levelling information are stored as list of the Bench Mark (BM) with station numbers, levelling paths, elevations, description of the location, sketch and map of the point location of the Bench Mark (Figure 4.5). Levelling can be a survey package which has raw levelling observation in level book, the processed data with new BM points being established with elevation and the list of BM which can be linked to the sketch (Figure 4.6) and map data. The two last column of the list in Figure 4.5 can be hyperlinked to a server in the Geodetic Section of JUPEM to get the sketch and map of the location of the BM. GIS-ready information for the levelling are thus non-object data with links to graphic data.

REC#	No Stesyen	Laluan	Tinggi Aras	Huraian	Lakaran	Map
1	B 0001	L-0052	35.54	Di tengah 'pulau' di simpang jalanSungai Besi dan jalan lapanganterbang lama.		
2	B 0002	L-0052	30.45	Di bawah jalan keretapi yang melintasi Lebuhraya Kuala Lumpur - Seremban.		
3	B 0003	L-0016	43.334	Di jalan Sungai Besi. Di bawah jambatan keretapi.		
4	B 0004	L-0016	70.48	Di kiri lebuhraya Kuala Lumpur - Seremban. Di bandar Seri Petaling.		
5	B 0006	L-0016	47.029	Di kiri lebuhraya K. Lumpur - Seremban. Berhampiran penanda km 50 Seremban.		
6	B 0007	L-0016	43.753	Di kiri L/raya K. Lumpur - Seremban. Berhampiran penanda km 48 Seremban. L/kurang 600 meter selepas jejantas ke Serdang Baru.		
7	B 0008	L-0016	41.911	Di kiri lebuhraya Kuala Lumpur - Seremban. Berhampiran penanda km 46 Seremban.		

Figure 4.5: List of levelling BM network

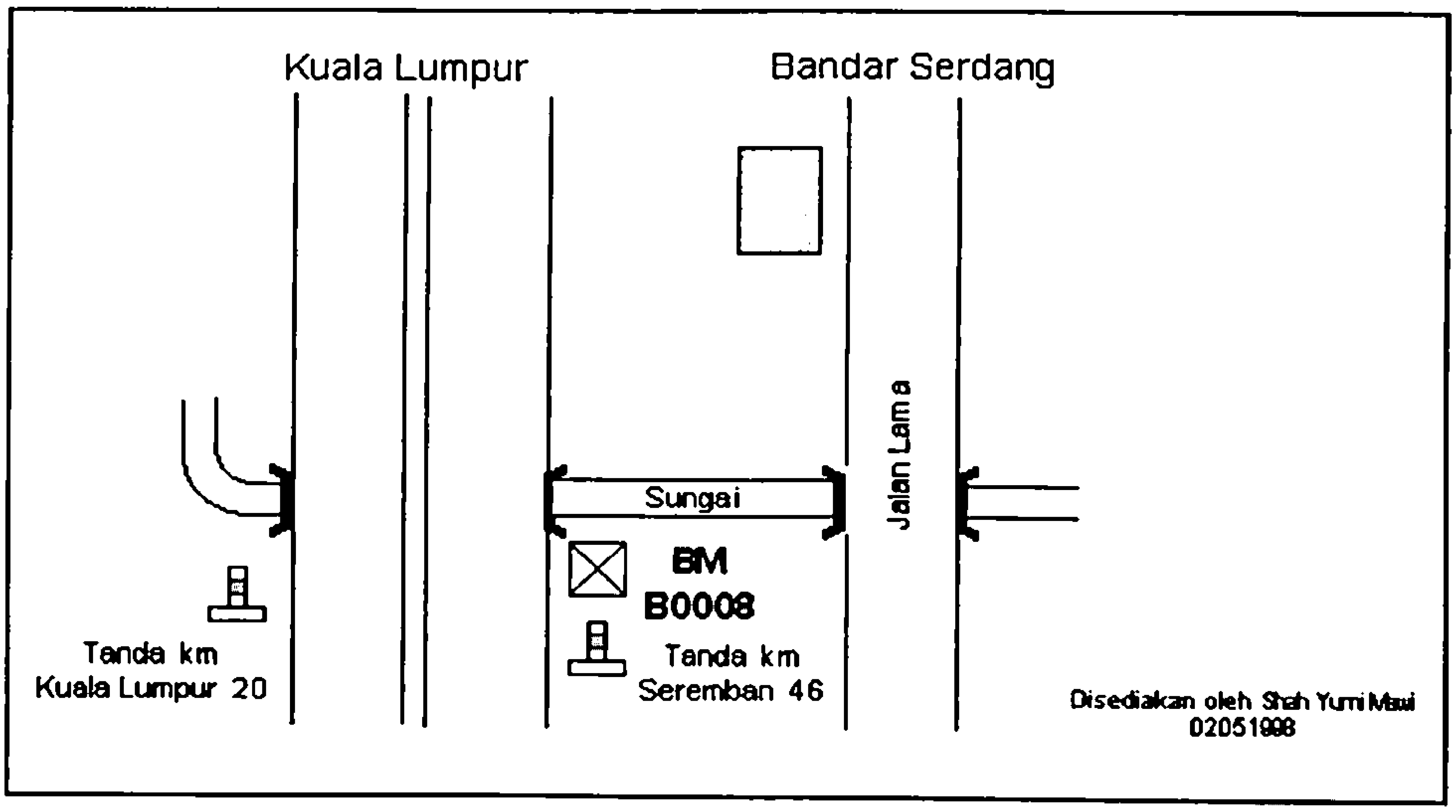


Figure 4.6: A location sketch pops out as 'lakaran' (or sketch) of BM 0008 is clicked

4.2.2.6 Gravity Measurement

Gravity is a resultant force due to the universal gravitation of the earth and the centrifugal force generated by the rotation of the earth. The strength and direction of gravity varies according to position and time. Gravity, in other words, is a physical index that contains much information about the earth. By measuring the distribution of gravity and its change with time it is possible to know the shape and size of the earth, to estimate underground construction, to study the seismic and volcanic activities and to investigate the viscosity and elasticity of the earth.

In JUPEM, gravity field is required for the following:-

i. To determine the local geiod enabling the reduction of geodetic measurements from the surface of the earth to the reference ellipsoid.

ii. To determine the orthometric correction to be applied to precise levelling measurements enabling orthometric heights to be computed.

iii. Other uses include prospecting for oil and other minerals, prediction of artificial satellite orbits, measurement of earth tides and other crustal movements and to detect possible changes in the gravitational constant.

Gravitational information is classified into two types, that is, first order and second order, measured using gravimeter. Main station for the measurement of first order gravitational measurement was numbered 02631B, located at a Physic Laboratory in University of Malaya in Kuala Lumpur. It is one of gravitational measurement station that was documented under the network system of the International Gravity Standardisation Net 1971 (IGSN71). This station was recognised as the control station for the establishment of the Peninsular Malaysia Gravity Standardisation Net 89 (PMGSN89). The gravitational information is used to produce orthometric correction in precise levelling and as well as to account geoids undulation for accurate mapping. The height different between an ellipsoid and geiod is called geoids undulation (Figure 4.7). For accurate orthometric heighting of a point this correction must be accounted for.

Gravitational data and its station can be modelled in GIS as point data with X, Y and Z coordinate, attributes and picture of the established station.

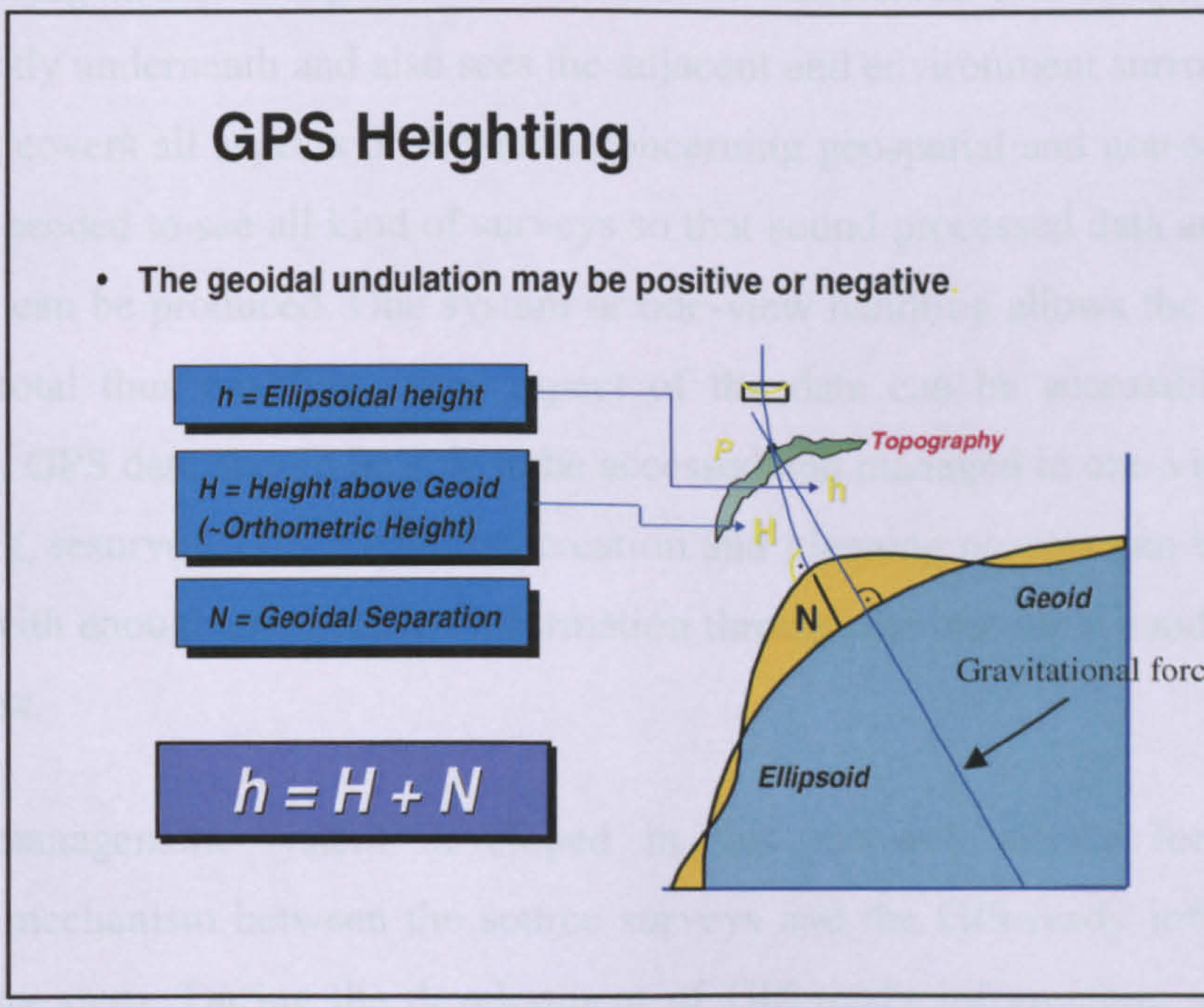


Figure 4.7: Gravitational information as applied to earth elevation (JUPEM, 2004)

4.3 Applications of GIS for Surveys in JUPEM

After discussing the surveys and data in the JUPEM, it can be assumed that these packages of survey can be modelled as layers of surveys that can be overlaid based on

GIS objects, which allow intuitive data management. Improved way of handling, querying and accessing by way of a 'drill down' search to query source data, processed data and GIS-ready information can be achieved when they are all stored and managed in the same real world coordinates and in one system of GIS (Figure 4.8). There is a need to search source historical data (such as the RS) when updating process of a cadastral base map is done. Using the cadastral raw survey, processed data and the GIS-ready land parcel, problems with subdivision or amalgamation can be solved, e.g. information such as status, RS, description of monuments and boundaries of a old survey of a land parcel may be needed to survey and plant new monuments in the subdivision of that parcel. Historical surveys are valuable to avoid resurvey of the same area and the same features.

By using GIS application to manage layers of survey, issues of birds-eye view and one-view handling can be applied. Birds-eye view means that all detail of the earth and the surrounding can be visualised and understood in three dimensional systems. The adjacent and surrounding details alongside the data can be understood and analysed. Bird sees feature directly underneath and also sees the adjacent and environment surrounding them. Survey data covers all aspects of the earth concerning geospatial and non-spatial. Birds-eye view is needed to see all kind of surveys so that sound processed data and GIS-ready information can be produced. One system or one-view handling allows the management of data in total thus enabling every aspect of the data can be accessible. Cadastral, topographic, GPS data should be able to be accessed and managed in one-view system so that updating, resurvey, GIS-ready data creation and planning process can be carried out efficiently with enough and detailed information through the one-view handling and drill down process.

The GIS management system developed in this research allows for a two-way traceability mechanism between the source surveys and the GIS-ready information and its processing steps. During the development of GIS-ready information, source survey data can be traced back to the origin of the survey or data capture (files, class of survey, standard or type of survey, date of survey, area measurement of polygons and who surveyed them) as well as the processing of the raw field data. Tracking of previous processing is sometimes necessary when GIS-ready information created are not accurately mapped or drawn, or are degraded during transformation or conversion.

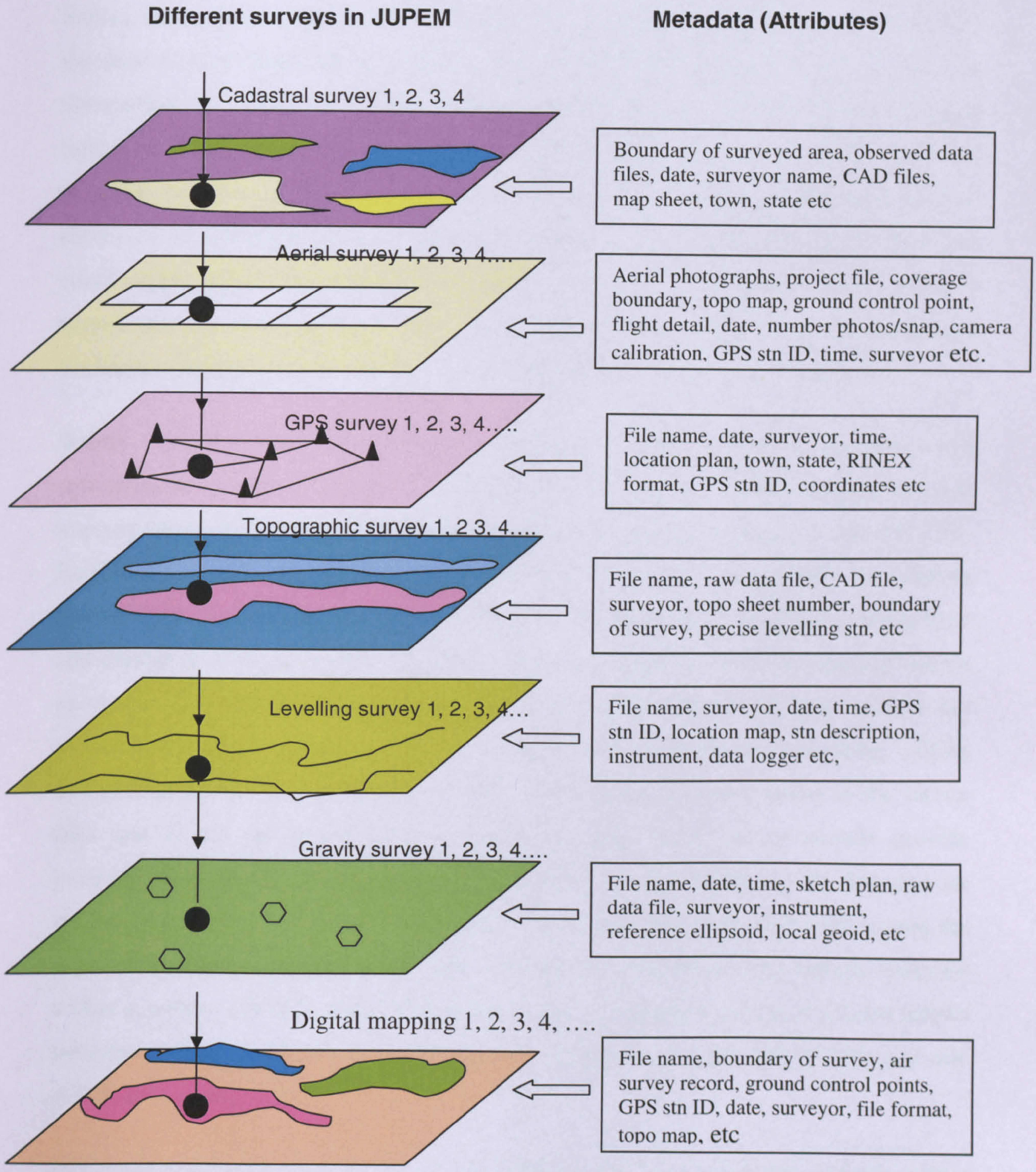


Figure 4.8: A spatially-distributed records and datasets of survey can be managed using geographic information system. A 'drill down' search can tell the history of survey being implemented on certain area

Survey measurements and computations assist in maintaining the integrity and quality of the data. Crucial historical description of surveys may be needed and tracked down so that complex processing in the development of GIS-ready information can be minimised especially when the previous survey has linked to other survey activities in the neighbouring area. Similarly, a surveyor in the survey department implementing new survey could also track in the system the existence of new GIS-ready data nearby that could supply information to minimise time of survey and save unnecessary survey. A survey may not be needed when GIS-ready development using remote sensing images has been completed for the same area by an other section in the same department.

Source survey measurements (bearing, distance, survey accuracy, azimuth) and computations can be used and maintained in a GIS database so that enhancement and improvements in the spatial quality for survey data can be propagated into the GIS. Requirements will inevitably change in the fast-paced geographic information environment that systems for geospatial data and GIS-ready information management and dissemination are targeted towards. Two-way traceability between captured source geospatial data and the processing of GIS-ready information could help achieve the requirements of the system by allowing a single view, comprehensive tracking, editing and processing of the geospatial data. History records are normally stored in the survey files and in the production files of processed data. These history records provide information regarding all measurement and processing steps applied to the data and are critical to providing full processing traceability. At any time during GIS processing the operator is able to examine all relevant records. All observations and records collected within a survey, and GIS-ready information can be arranged into a data layer that allows traceability back to the file name, surveyor in-charged, surveyed region, date and raw ASCII data collecting file.

GIS-ready information in JUPEM can be widely used to implement other surveys within the organisation. Mapping for national base map needs cadastral GIS-ready information in order to map properly the building and road within the administrative cadastral boundaries. Cadastral survey can be very easy if building structure of residential parcel can be accessible within one view in the working computer. In future, cadastral survey is not anymore practical to carry out by bearing and distance survey or GPS survey. With topographic detail of the residential structure, cadastral survey of land parcel can be carried out with little accuracy without planting the monument anymore. This survey may be carried out by just tracing the topographical detail of the building structure. Land parcel survey can also be carried out using satellite or aerial photographs (digitisation) as alternative which actually save money, time and manpower. There are therefore two ways of interaction of surveys in JUPEM especially when the GIS-ready information being managed together with the processed and raw data.

GIS-ready information in JUPEM has been important to the general public especially for issue of land title deed and for the national development of the country. Lawyers and land administrator are interested to access land parcel information for their daily work concerning land disputes and council tax or quit rent. National base map in GIS –ready information are very significant to the general public including the decision makers and politician. GIS-ready information should be available online to the executive level of all ministries, state level authoritative officer and federal department. JUPEM is truly important to drive into the geospatial world as the main provider of GIS data.

As a result from the discussion above, the management of these surveys and data needs modelling. This will be focussed in the next section.

4.4 Data Models

A model is a simplified representation of something, either as a physical object or some aspect of a problem or description of the object. Data modelling is a type of data abstraction that is used to provide conceptual representation of the real world. Data modelling normally defines specific groups of entity, and their attributes and the relationships between these entities (AGI, 1999). A sample of a certified plan for a land parcel produced from a cadastral survey, or vector data digitised from aerial photography within a national survey and mapping department are examples of data models. They are also objects for an object class of processed data.

Traditionally, GIS displays information in the form of maps and symbols. Maps provide knowledge about places; where, what and how they can be reached. Information within maps can be achieved through GIS functionality interactively, not appreciable on a printed map. A list of all attributes of a feature is an example of a query. Performing buffer analysis for features within 100 metres of a river is another kind of GIS interface. The way this functionality can be achieved depends upon how geospatial data is modelled from the real world. Basically geospatial data can be modelled in three ways (Zeiler, 1999), namely, vector, raster and triangulated irregular network (TIN).

4.4.1 Vector Model

Geospatial data represented by vector uses the basic units of spatial information of points, lines and polygons. Arcs are part of a line unit. Each of these units is composed simply as a series of one or more co-ordinate points, for example, a line is a collection of related points, and a polygon is a collection of related lines. Figure 4.9 shows simple point, lines and polygon vectors with coordinates. They are also represented in the form of coordinates that make them spatial.

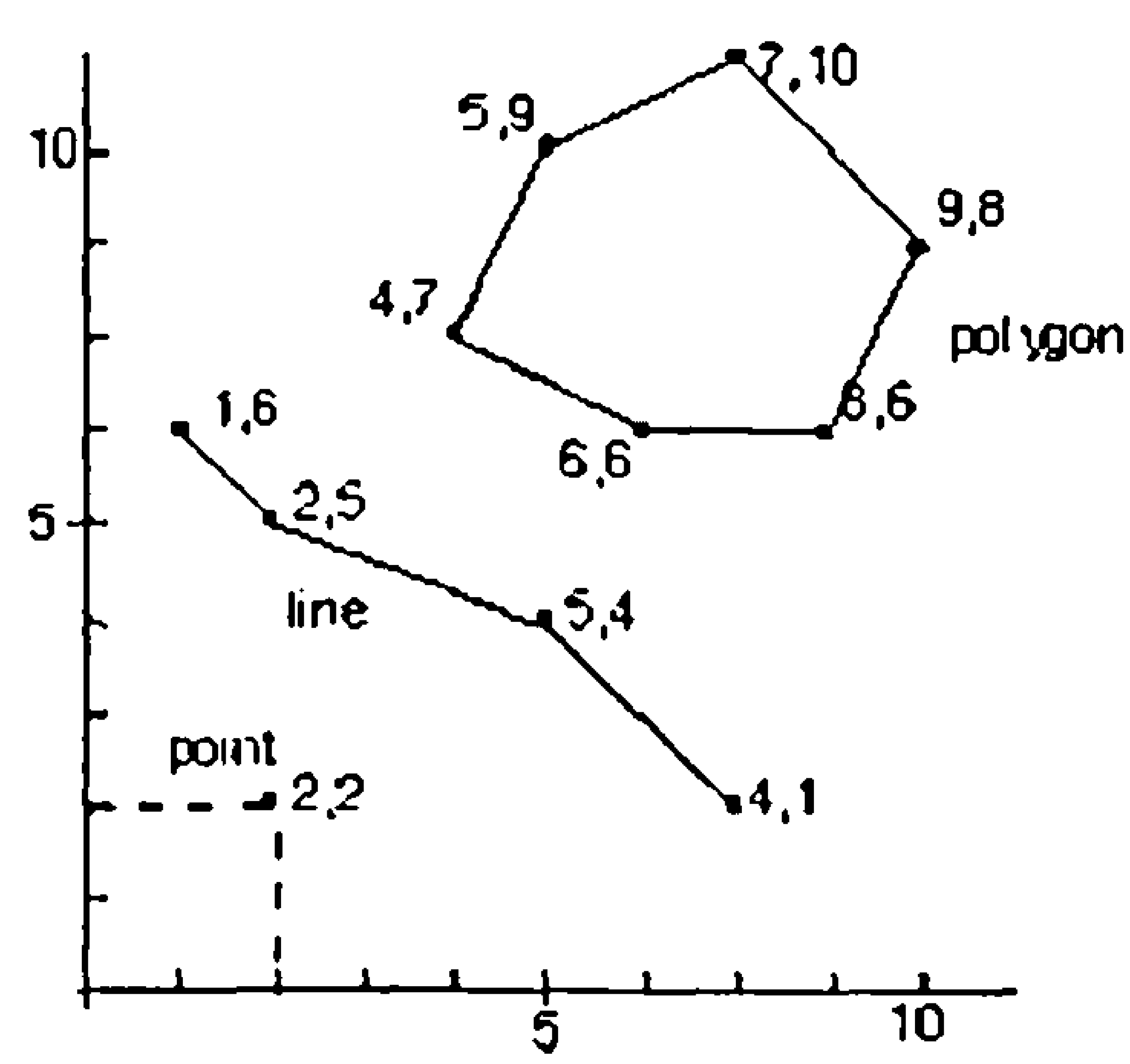


Figure 4.9: Representational geometry of vector format

A survey carried out in the field yields vector forms of models of the captured or observed objects on the survey site. They can be boundary limits (polygon), lines of a survey network and points for the survey station or land parcel monument. A basic survey network consists of survey stations, framework control survey, traverses, radial shots to located detail (e.g. a building or road), referenced monument or GPS station. This data and information are vector form models of the surveyed reality. Non-spatial information recorded for a survey are surveyor, date, ASCII files of the observed information as captured on a data logger or total station. There are many survey data collection technologies that produce models of survey. As an example, the latest STAR*NET PRO, a survey software suite produces non-spatial survey observation and GPS vectors. GPS survey produces coordinated points which can be plotted to form vector point, lines and polygons. The latest technology allows these models to be plotted in the field within a concept called 'Mobile GIS'. A survey using total station also allows the concept of 'mobile GIS' to be carried out.

Vector data is therefore a collection of discrete features represented as points, lines and polygons. It is best graphically applied to discrete objects in which shapes and boundaries can be defined. In this way features have a precise shape and position, attributes and metadata, and useful behaviour (Fotheringham and Wegener, 2000).

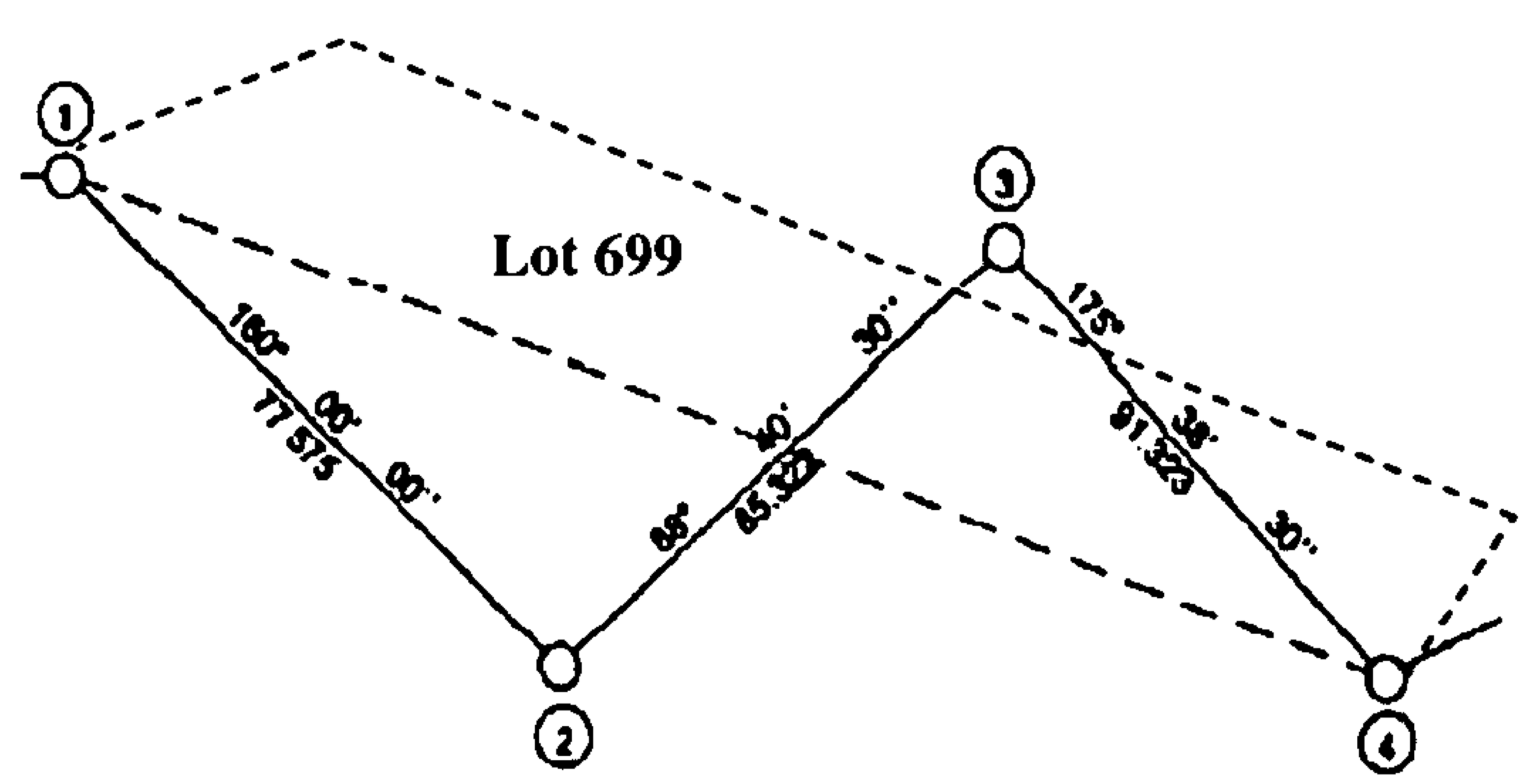


Figure 4.10: Survey stations and lines of observation are vector forms of model (lines)

As in Figure 4.10, survey lines produce line vectors with attributes of bearing and distances with the end of the lines being survey stations (peg). Line 1-4 is a boundary line to be established for a land parcel, the survey from 1 to 2 to 3 and 4 would result in the computation of the line 1 to 4. This line is established as a vector boundary line for a land parcel. A closed boundary lines produce a polygon vector which is a geospatial object called land parcel (Lot 699). Survey monuments are planted at 1 and 4 as the point data for the boundary marks.

A polygon boundary limit of a survey represents the polygon vector. A small detail survey by radiation observation from an established and referenced station to locate a building or road produces vector radial lines and the building or road is computed in the field or downloaded to a data logger and computed in a survey office (Figure 4.11). A concept of field-to-finish has been a fashionable and practical way of carrying out complete survey from field to office for final plotting.

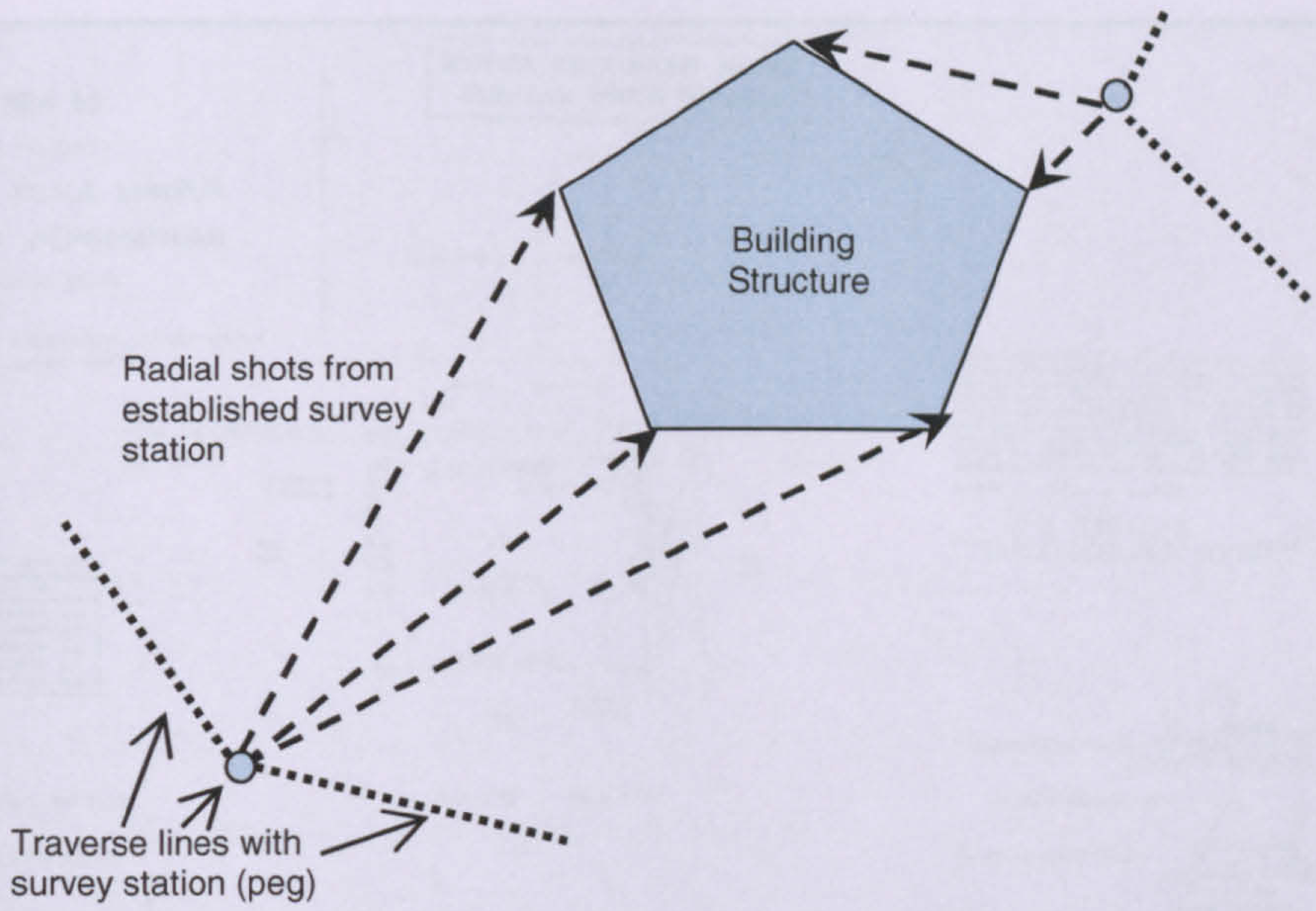


Figure 4.11: Radial shots that result in a polygon vector of building in detail survey

The data and information created within surveys as in Figure 4.10 and Figure 4.11 are considered objects that can be managed using geospatial handling techniques. They are able to be stored as spatial and non-spatial objects within the object of geospatial data capture. The processed data consists of drawing files in CAD that have been processed from the survey. CAD files have enhanced features that illustrate the model survey reality in a graphic presentation. CAD drawing files present a vector form of data. CAD files are also produced after processing from raw air survey via photogrammetry.

For a cadastral survey, a certified plan is a lawful end product used for the graphic documentation in a land title for a certain land parcel. Certified plans drawn manually in survey offices when field surveyors have given the details of survey are shown as in Figure 4.12. A digitally certified plan produced through a process of data entry (keyboarding) of information from old certified plans is vector data. Digitally certified plans are products of cadastral surveys which can be categorised as data within the class of processed data. To-date, in JUPEM a field-to-finish project for cadastral surveys that will complete the drawing of certified plans directly without much intermediate computation is on the way to realisation.

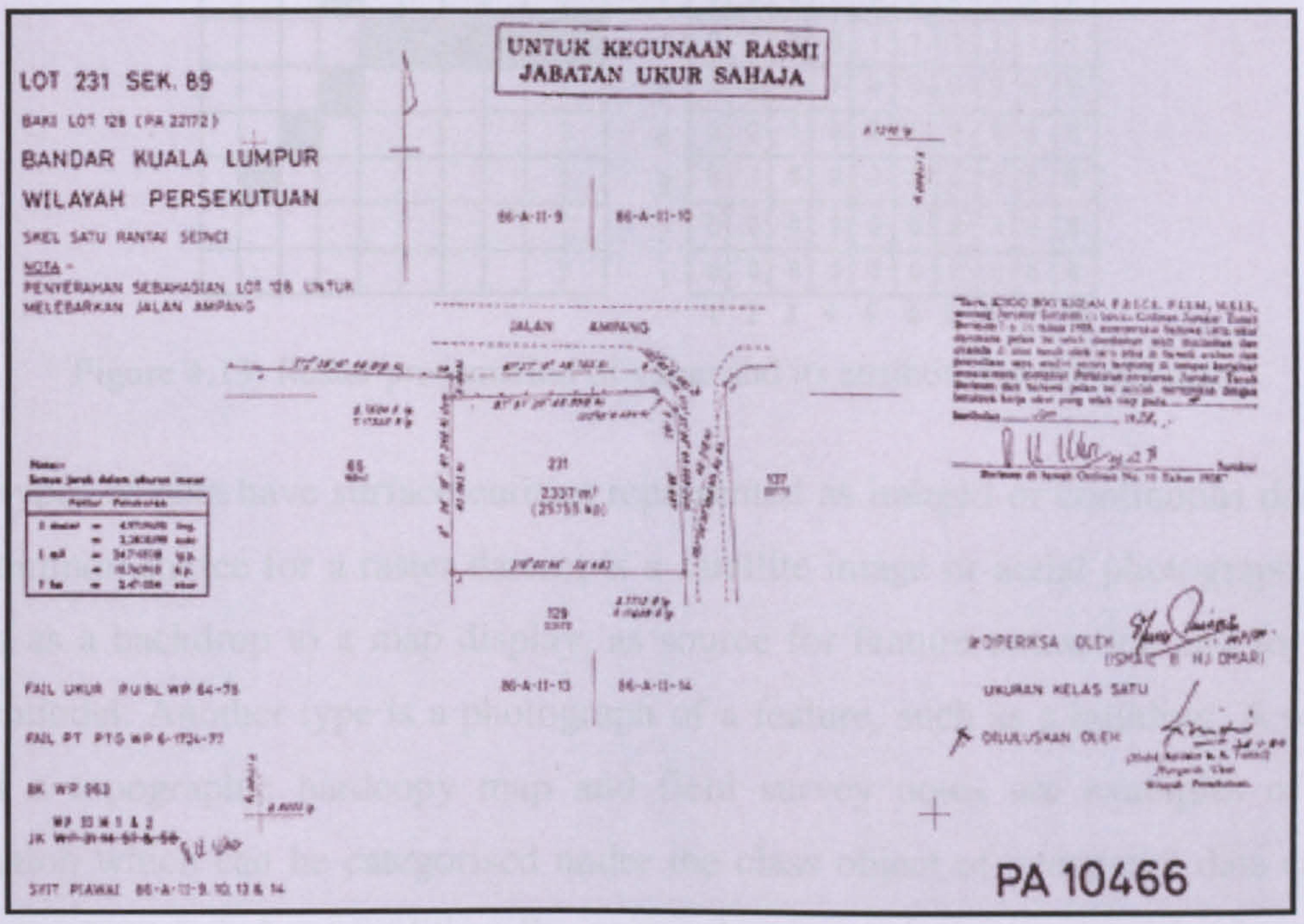


Figure 4.12: A certified plan produced from cadastral survey

4.4.2 Raster Model

The raster data model is based on the division of reality into a regular grid of similar shaped cells. Each cell is assigned a single value, which represents the attributes for the area of that cell (Figure 4.13). The area each cell represents varies from a few metres to kilometres and is known as the resolution of the grid. The higher the resolution of the grid, the more cells are required to represent a given area. The raster data model is simple but it supports a rich variety of data types.

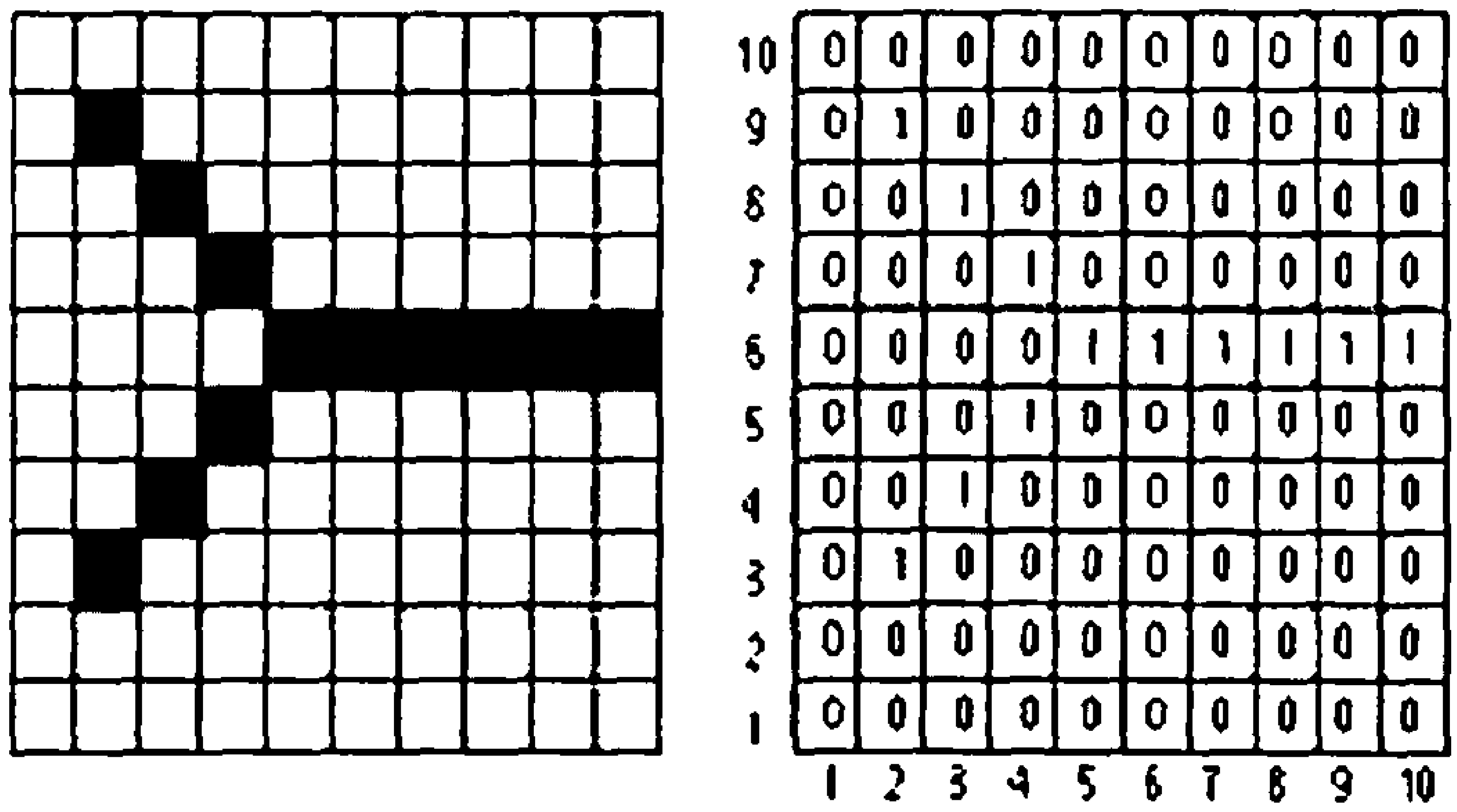


Figure 4.13: Raster presentation of value and its attribute for the cell area

Raster types of data have surface entities represented as imaged or continuous data. The most common source for a raster dataset is a satellite image or aerial photograph. It can be used as a backdrop to a map display, as source for feature extraction and for a grid surface model. Another type is a photograph of a feature, such as a building. A scanned map of a topographic hardcopy map and field survey notes are examples of raster information which can be categorised under the class object of geospatial data capture. As surveys are carried out traditionally, scanned computation notes can also be a raster type of data that can be archived as historically valuable for reference when a resurvey needs to be carried out.

In the national survey and mapping department there are several datasets that can be modelled as raster. In JUPEM, certified plans of cadastral survey can be scanned as media output for access by users within the department and by authorised licensed surveyors who need detail of surveyed parcel as reference to start a cadastral survey. This kind of data is considered part of processed data for a cadastral survey; therefore they are to be modelled as objects within processed data.

JUPEM carries out air survey for national mapping. The produced aerial photographs are raw survey in raster format (Figure 4.14). According to JUPEM's experience, a typical aerial photography has produced 55 raw photographs, which are used in this research. This data is numerous considering each photograph holds a space of 23 Mb. Therefore a proper management for this type of raster data is needed for efficient access and processing.

Figure 4.14: An aerial photograph is an example of a raster format of raw survey

4.4.3 TIN Model

A TIN is a vector data structure, which depicts geographic surfaces as contiguous non-overlapping triangles (Molenaar, 1998). The vertices of each triangle match the elevation of the terrain exactly. This means that a topographic surface is represented by several triangles, with each triangle face having an approximate slope, aspect, and surface area. The irregularity of the triangles comes from the scattered nature of the (x,y,z) points (the triangle vertices) used as a background elevation source (Figure 4.15).

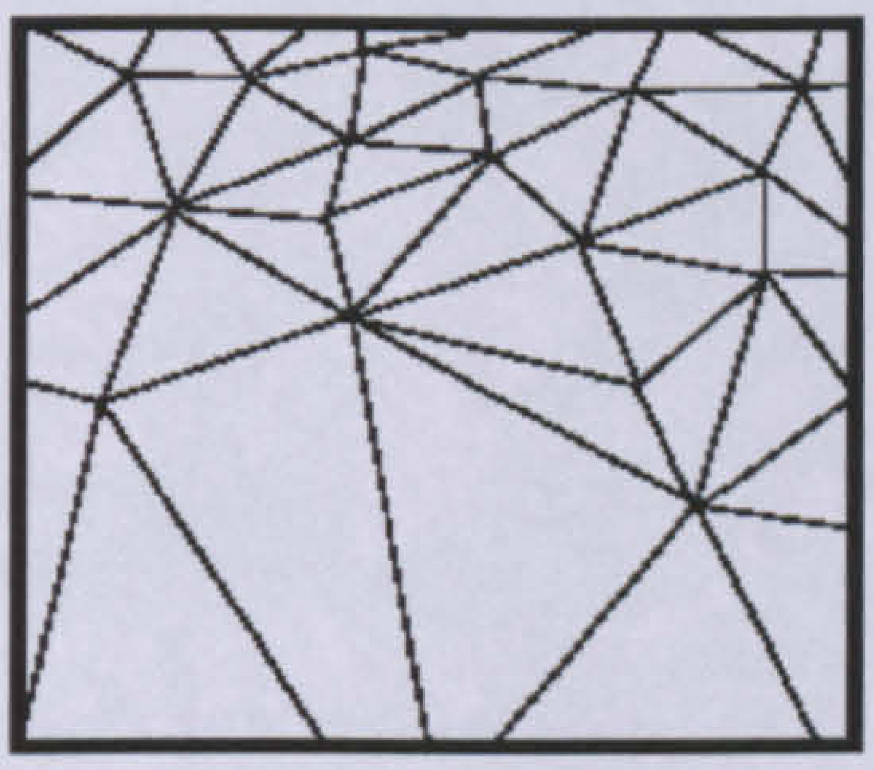


Figure 4.15: A TIN comprises of irregularly spaced connected triangles

TINs are particularly effective to represent surfaces that are highly variable and contain discontinuities and breaklines (Chrisman, 2002). This is possibly because they connect significant locations which are regularly spaced and these locations define a point where surface changes in elevation. Downhill slopes can be shown as all neighbouring points at the peak of a hill and adjoining points along a stream show an uphill slope. Research in this type of geospatial data for GIS has been significant and active presenting such issues in lossless compression (Kidner and Smith, 2003) and accurate representation compared to real elevation (Carlisle, 2005).

A surface data, Digital Elevation Model (DEM) refers to the continuous representation of change in relief of terrain or height variation of terrain. It has a third dimension space of elevation as Z-co-ordinate. The difference of TIN to that of DEM is that the TIN model produces irregular triangles of network which makes it a 2½ dimensional model. DEM are typically used to represent terrain relief and also referred to as digital terrain model (DTM).

In JUPEM, raw air survey is followed by digitisation, triangulation and photogrammety compilation processes. CAD files are then produced through photogrammetry method using Microstation software. DEM, DTM and TIN data can be produced using CAD files through appropriate processing steps in the department. DTM data and TIN are used as objects to be modelled under processed data. A sample of DTM data produced by JUPEM is shown in Figure 4.16 to demonstrate hill shading. The JUPEM has the responsibility of producing and distributing DTM for the government and public uses.

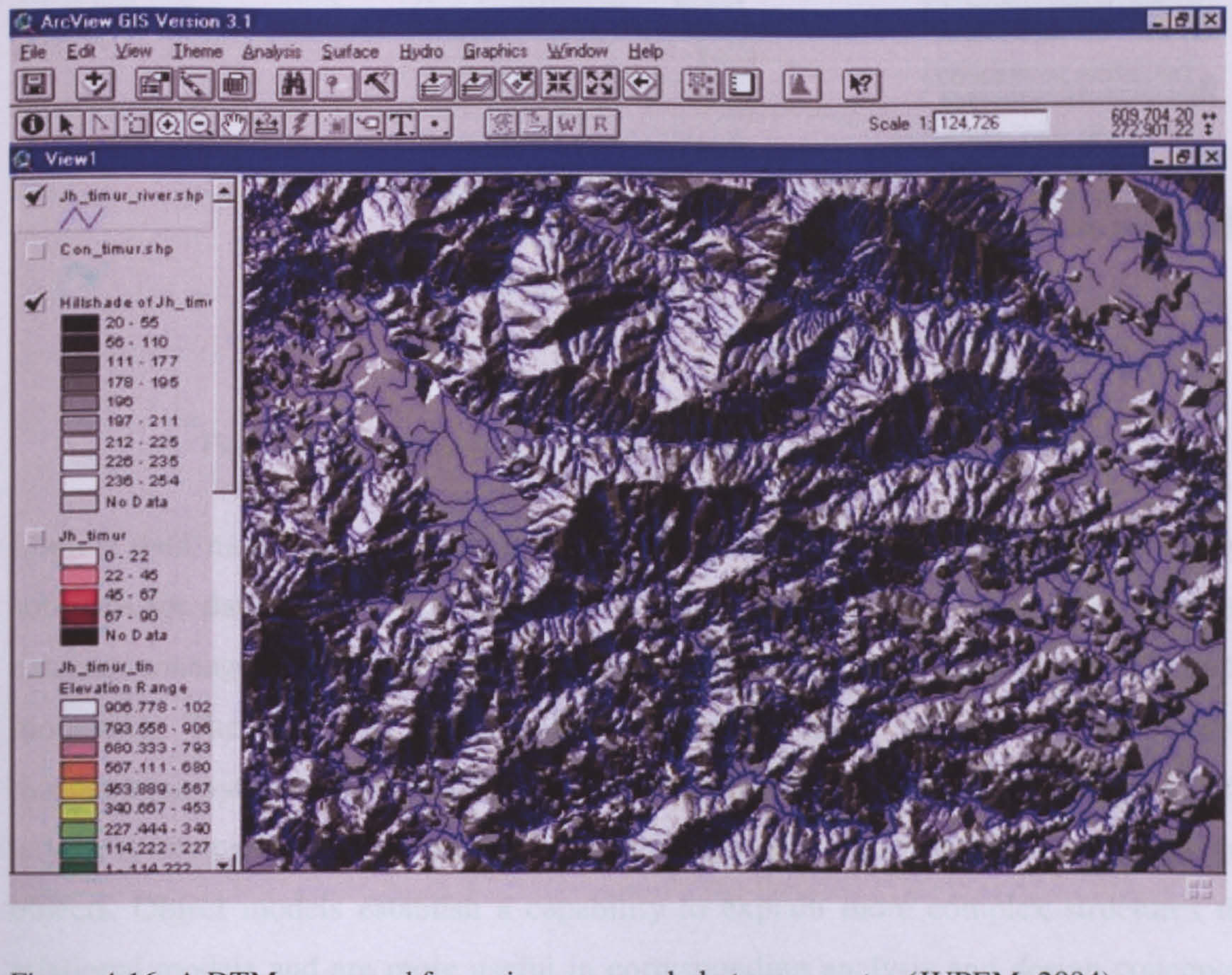


Figure 4.16: A DTM processed from air survey and photogrammetry (JUPEM, 2004)

4.5 Data Modelling

Data is modelled to conceptualise it and is structured so that it can be processed by computer. Computers can handle data if it is stored in bits, organised in bytes, records and pages. When data is mapped in these ways it is said to be a physical model. At this stage, the model is implemented. This step is represented by the inner most circle in Figure 4.17. The modelling of the data at this level is not fully understandable because the way it works is not really accessible by most users of spatial information systems (Fotheringham and Wegener, 2000). Presently and in the last few years, people have become aware of and developed at a higher level than the physical model, that is the logical data model which is also known as the database model (Molenaar, 1998).

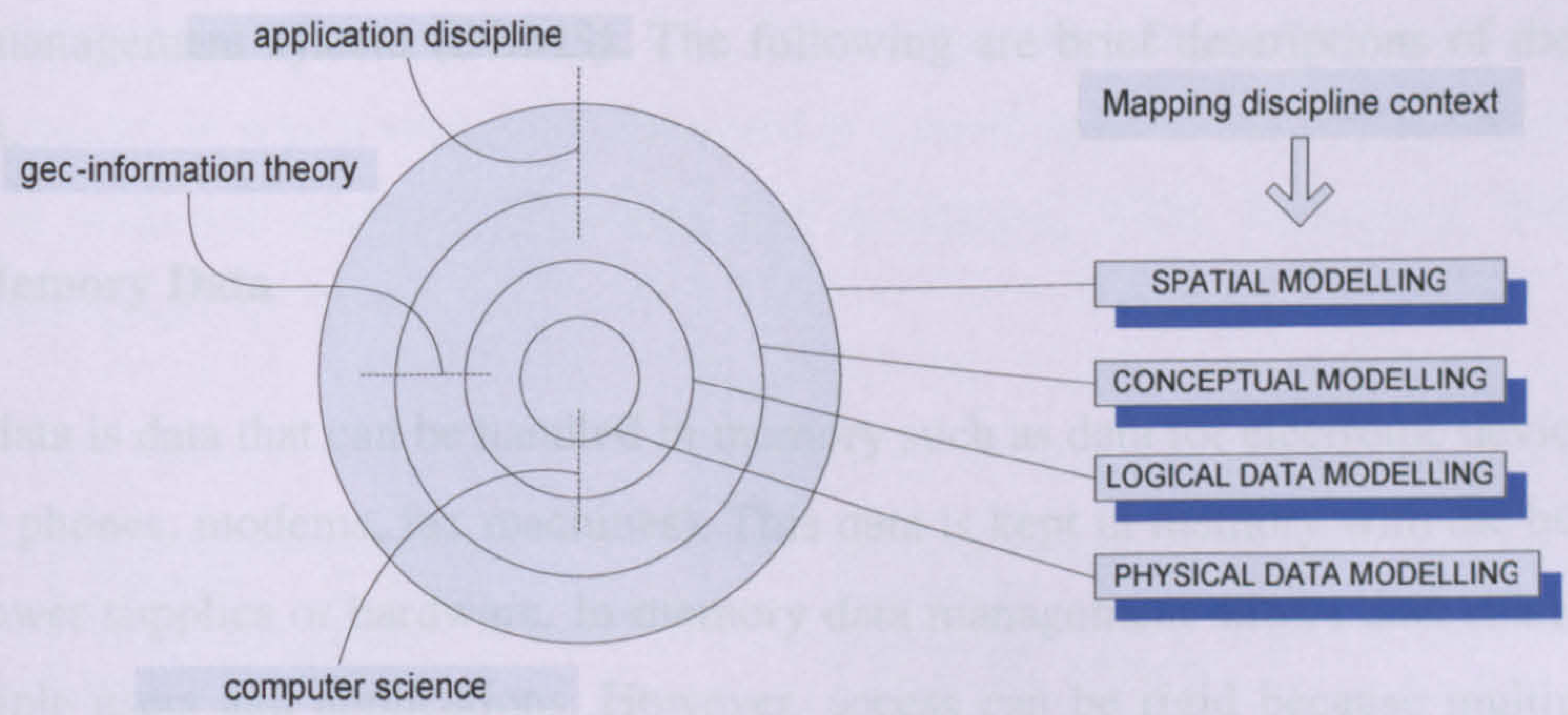


Figure 4.17: Several levels of data modelling (after Molenaar, 1998)

Logical data modelling, as represented by the second inner circle, involves the use of software for data input and storage, and for querying the object data entrenched in a database management system. The relational model is a popular one used in this step of modelling. The data can be organised and manipulated by a relational database management system. More recently models use objects to populate a database. In other words, to represent relevant entities of feature through a database, users handle them as objects. Object models establish a capability to explain more complex structures than relational models and are more useful in corresponding analysis and design concepts to users.

Before formulating a logical data model, relevant data should be first conceptualised and selected. Relationships between spatial data should be described and represented correctly. In the development of a GIS application, it is important for the data model be designed to answer appropriately spatial queries and other spatial questions. Decisions are, for example, how the features should be represented geometrically, which data features representation can be used to answer spatial questions and which spatial representation is best to describe the data, either in vector or raster structure. The second outer circle in Figure 4.9 represents the conceptual data modelling.

4.6 Data Management Approaches

There are several options for data management; in-memory data, files, hybrid systems or a database management system (DBMS). The following are brief descriptions of these three options.

4.6.1 In-Memory Data

In-memory data is data that can be handled in memory such as data for electronic devices (e.g. cellular phones, modems, fax machines). This data is kept in memory with the help of special power supplies or hardware. In-memory data management allows data sharing among multiple users and applications. However, access can be rigid because multiple applications may find difficult to agree on a format. In addition, the inflexibility of memory data inhibits the extension and evolution of large systems. Persistent memory in this type of data management has a limited capacity and can be costly. Geospatial data and information which involves bigger bytes is not suited to this way of data management.

4.6.2 Files

Data and information that are handled in files are normally sequential or random-access files. Data filing is a simple mechanism for attaining data persistence. Data that belongs in files includes raw data that is summarised in the database, such as from data acquisition. Data that is voluminous and of low information density such as archival files and detailed historic records is another example of file management. As data increases in quantity and variety the complexity grows rapidly. Files normally use applications that have a common format. They can be difficult to maintain, extend and co-ordinate for concurrent access. Large sequential files are inefficient for accessing small, random portions of data especially for geospatial data which handles mostly graphic formats.

4.6.3 Hybrid System

A hybrid system of managing geospatial data comprises of separate spatial data files and non-spatial data files. Figure 4.18 shows schematically the hybrid architecture. The left-hand box represents the geometric and topological graphic data and the right hand holds the non-spatial (not graphic) data. Hence for a building feature, the polygon vector representation with geometry is held in the left box while the building type, name, parcel ownership, value is held in the right box. Typically this system is based on a georelational model, in which spatial data is stored in a set of system files and non-spatial data stored in relational database (Worboys and Duckhams, 2004; Rodriguez, 2004; Adam and Gangopadhyay, 1997).

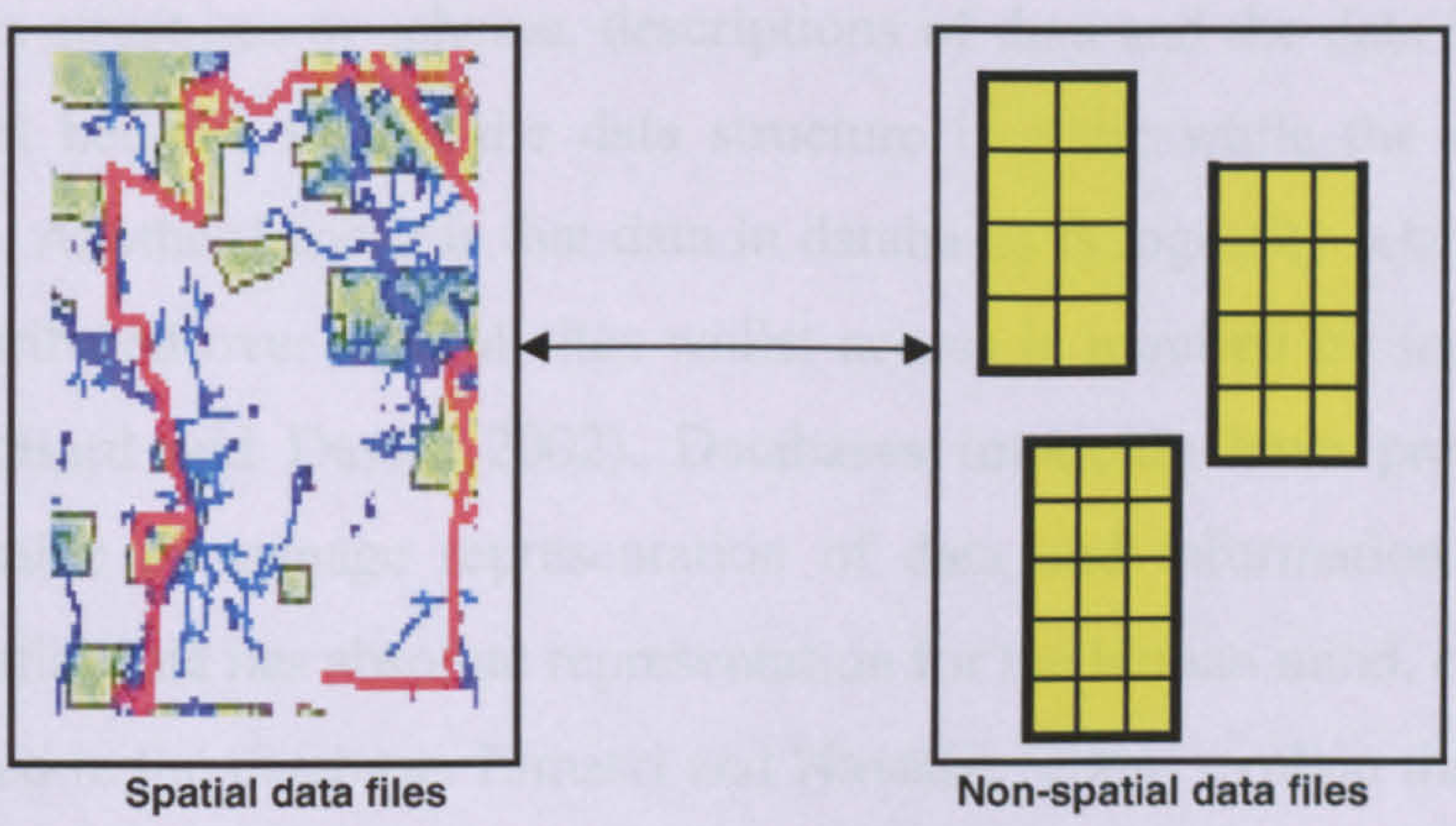


Figure 4.18: The hybrid database design

An early hybrid geospatial data management involving the georelational model is ArcInfo, produced by ESRI. In the first system *Arc* was introduced as the graphics and spatial data engine while *Info* was the non spatial database (Rigaux et al., 2002). This system is still maintained with ESRI's Shapefile format, frequently used for storing and transferring geospatial data. The Shapefile format consists of three separate files. The one to store geometry has a file extension *.shp,* for non-spatial data table *.dbf*, and the file to store an index to relate the geometry to a tuple in the data table comes with *.shx* extension.

Putting spatial data and non-spatial data in separate storage makes the data management system quite rigid to maintain its integrity, security and reliability (Devogele et al., 1998). Lack of integration of spatial and relational data causes the system to be unpopular. The hybrid system also has a propriety data structure and format that cannot be read or updated by external applications (Newell and Theriault, 1995). As a result, geospatial data management that would contain numerous complex data and information, spatial and non-spatial, modelled as objects integrated with relational model concept needs an integrated system such as in DBMS.

4.6.4 DBMS

It is important to organise a method in which spatial and non-spatial files can be stored and linked in the computer to model real world phenomena. Similarly, efficient storage and retrieval of the data is ensured. A database acts as a permanent, self-descriptive repository of data, which is managed in one or more files (Blaha and Premerlani, 1998). It contains data structures or schema, descriptions of data and the data itself. Databases give significant benefits in that the data structure is static while the actual data may rapidly evolve. Another benefit is that data in databases is logically related, but possibly physically distributed over several sites whilst access is required by many applications and users (Voisard and David, 2002). Databases implicitly have properties that are especially suitable to manage representation of data and information of the world's reality. Geospatial data has absolute representation for the human mind, catering for what can be modelled in the database. Elmasri and Navathe (2004) explain that the properties that are relevant are as follows:

- A database represents some aspect of the real world in which changes to the real world are reflected in the database.
- A database depicts the natural meaning of a logically coherent collection of data: an arbitrary range of data cannot technically be referred as a database.

- A database is designed and populated with data for a specific purpose with an intended group of users and applications in which these users are interested.

A database is created and maintained using a general-purpose piece of software called DBMS. A DBMS provides routines and protocols for managing large quantities of data and isolates applications from detail of the physical storage (Worboys and Duckham, 2004). Together database and DBMS software can be described as a database system. A typical database environment for GIS using DBMS is shown in Figure 4.19.

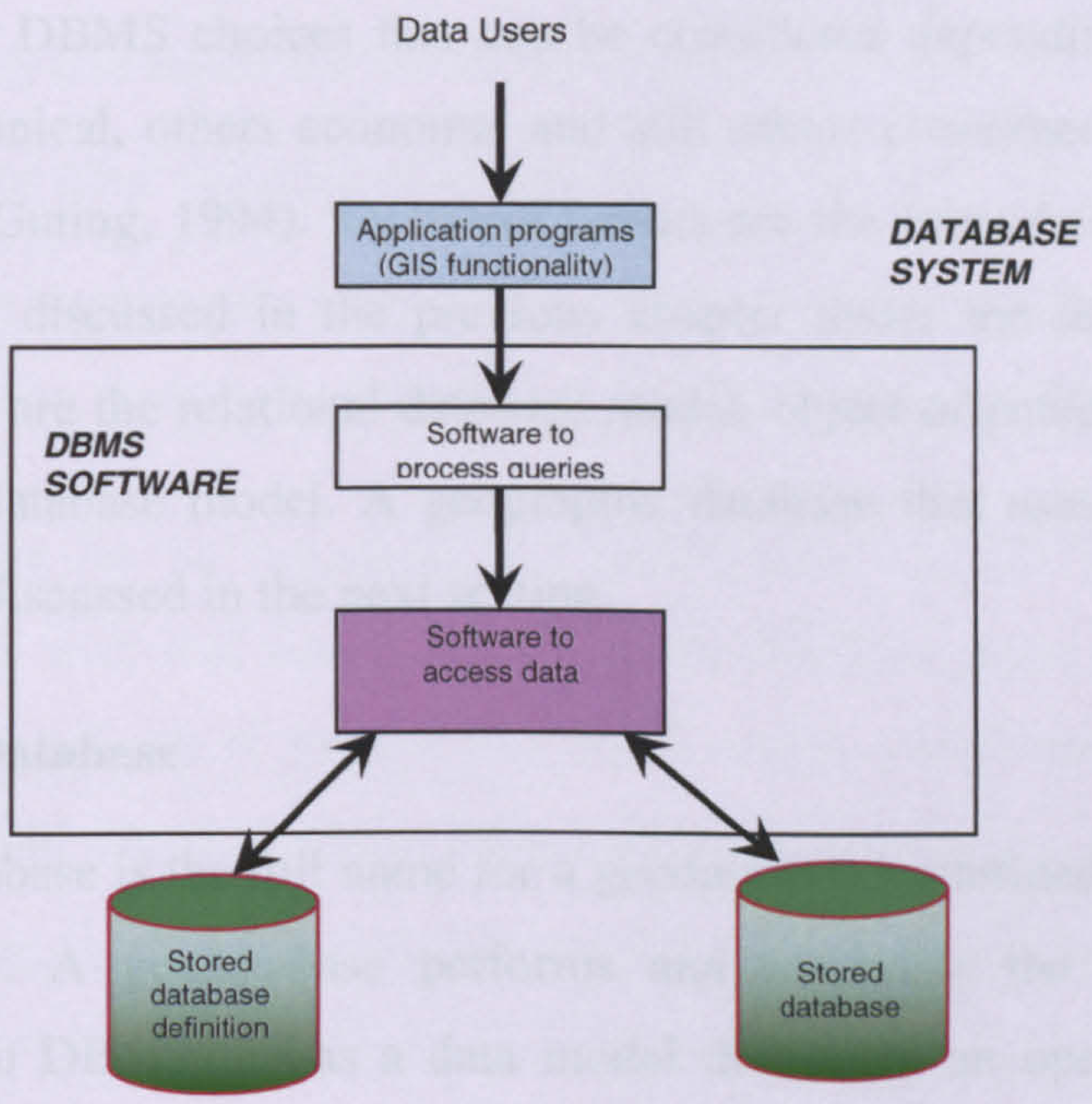


Figure 4.19: A simplified database system (after Worboys and Duckham, 2004)

Geospatial data that is managed in a database system, not in-memory, file or a hybrid system, needs certain criteria so that maximum benefits can be achieved. These criteria include

- Data that needs updating at fine level of detail by multiple users.
- Data that requires access by multiple application programs.
- High volume of data that must be efficiently handled.
- Data that is long lived and highly crucial and costly to an organisation.
- Data that must be carefully secured against unofficial and malicious access.

- Data that is mostly focussed to complicated analysis for decision support.

The process of transferring geospatial data capture to GIS-ready information involves variety, voluminous and complex datasets. The use of database systems is therefore suitable. Basically, a GIS is very much dependent on a structured database which enables a description of the real world geographic terms (Lemmen, 2005). Therefore using a database system along with the environment of GIS functions allows efficient geospatial data management.

There are various DBMS choices that can be considered depending on the number of factors: some technical, others economic and still others concerned with the politics of the organisation (Guting, 1994). Technical factors are the focus herewith. These factors and choices were discussed in the previous chapter under the database model. The choices of DBMS are the relational database model, object-oriented database model and object-relational database model. A geographic database that uses an object-relational model concept is discussed in the next section.

4.7 Geographic Database

A geographic database is the full name for a geodatabase as defined in the Dictionary of GIS Terminology. A geodatabase performs and maintains the physical storage of information within DBMS. It has a data model that holds an operational view of the database and supports objects with attributes, relationships and topology. A geodatabase uses object-based features to represent geographic features and attributes and they can be hosted inside a relational DBMS. The database design in a geodatabase conceptualises the object-relational model to structure geographic data. Hence, a geodatabase is a relational database that contains spatial and non-spatial objects. A geodatabase is described as a modern container for geospatial data because most data models can be supported for GIS applications (ESRI 2001).

A geodatabase is conceptually analogous to the well-known file-based coverage and shapefile data models but extends these models in some important ways. These include support for topologically integrated feature classes, advanced geometry (for example three-dimensional co-ordinates), linear measures, continuous non-tiled data storage, feature-linked annotation, complex linear networks or geometric networks, user-defined relationships among features classes, address locators, and efficient storage of raster data (MacDonald, 2001).

There are three main features of the geodatabase (Zeiler, 1999):

- It is a centralised management of wide variety and rich data types of geospatial data and information within DBMS.
- It stores and maintains intelligent features that have sophisticated relationships, editing rules and behaviour as well as transactional views of the geodatabase.
- It can access large volumes of geographic data stored in both files and databases.

In comparison to traditional hybrid GIS data models, geodatabases have several important advantages, such as the following (Arctur and Zeiler, 2004):

- GIS data can be visualised in a fashionable way. By modelling geographic features with real-world objects in the database, users can work with more intuitive objects of interest like water mains, transformers, roads, or parcels, versus the generic "points, lines, and polygons."
- GIS tools within a geodatabase. An editing environment provides precision feature "sketching" tools such as intelligent snapping, relative or absolute co-ordinate entry, extending and trimming features, parallel or perpendicular constraints, and more. Although these tools work on other data models, like shapefiles, they are optimised for use with geodatabases.
- Different representations of geography – vector, raster and TIN, can be stored in the same geodatabase. They can be built, stored, and managed centrally, if desired. As geodatabases scale from the project to the enterprise they can also be shared and distributed via the Internet thus enabling single portal data management.
- Using geodatabase models, multiple users can edit geographic data simultaneously. The model permits work to flow, where many editors can work on the same geographic area and then reconcile conflicts that emerge.

4.7.1 Objects and Relational View of the Geodatabase

The geodatabase is an object-oriented database. In this model view of the reality, entities correspond to objects with properties, behaviour and relationship. A mixture of various geographic object types are mapped and supported in the geodatabase system. Examples of these objects types are simple objects, geographic objects (objects with location), network features and annotated features. Relationships and rules can be built and defined between objects thus maintaining the reliability and reference among objects. Geodatabase supports two views of this data model, the object view and the relational view. The object view of the database is used by a GIS application to define and work with the database as an integrated system of geographic objects (Verbyla and Chang, 1997). The relational view provides a simpler, non-object view of the data model. The relational model corresponds to the simple feature model used in GIS. The view can be presented, using modules within the geodatabase in which the datasets are stored, in the relational database via Microsoft Access database software. It can also be opened with Microsoft Access if needed to see how it is organising the data.

4.7.2 Data Elements of the Geodatabase

As mention earlier, there are three general types of geographic data models, namely vector, raster and TIN. Geodatabase implements the vector model as represented in feature datasets, the raster model in raster dataset and catalogue, and the TIN model in the TIN dataset (Zeiler, 1999). These datasets make up elements of the geodatabase apart from feature classes, object classes (attribute table) and relationship classes. The following subsections describe the six data elements of the geodatabase.

4.7.2.1 Feature Dataset

A feature dataset contains spatially related feature classes together with the relationship classes but does not contain tables (Figure 4.20). Feature classes are bound together to hold a system of geographic objects. A feature dataset is similar to a directory or folder in a computer Window operating system. It is used to help arrange the data into a logical structure with a common geo-reference system. In a geodatabase, relationships are stored based on tabular information but feature datasets store them based on geographic information. An example of a feature dataset is infrastructure that contains roads, airports and railtracks.

For this research, the feature dataset is used to model objects of geospatial data capture, processed data and GIS-ready information within a DBMS. Feature classes within the feature dataset enable the modelling of all spatial and non-spatial data in the raw geospatial data, processed data and GIS-ready information. They may be all the data and information within survey packages: examples of the spatial type include limit of boundary, framework control survey or radial shots to located detail, whilst the non-spatial type are ASCII text file of GPS observation or survey software, surveyor, date, flight line detail for aerial photography or traverse computation. The conceptual model which uses feature dataset concept for modelling geospatial data management from raw captured data right through to the GIS-information is illustrated in Figure 4.21. There are many more feature classes that can be produced considering other types of surveys such as LIDAR survey, terrestrial laser scanning, hydrographic survey etc.

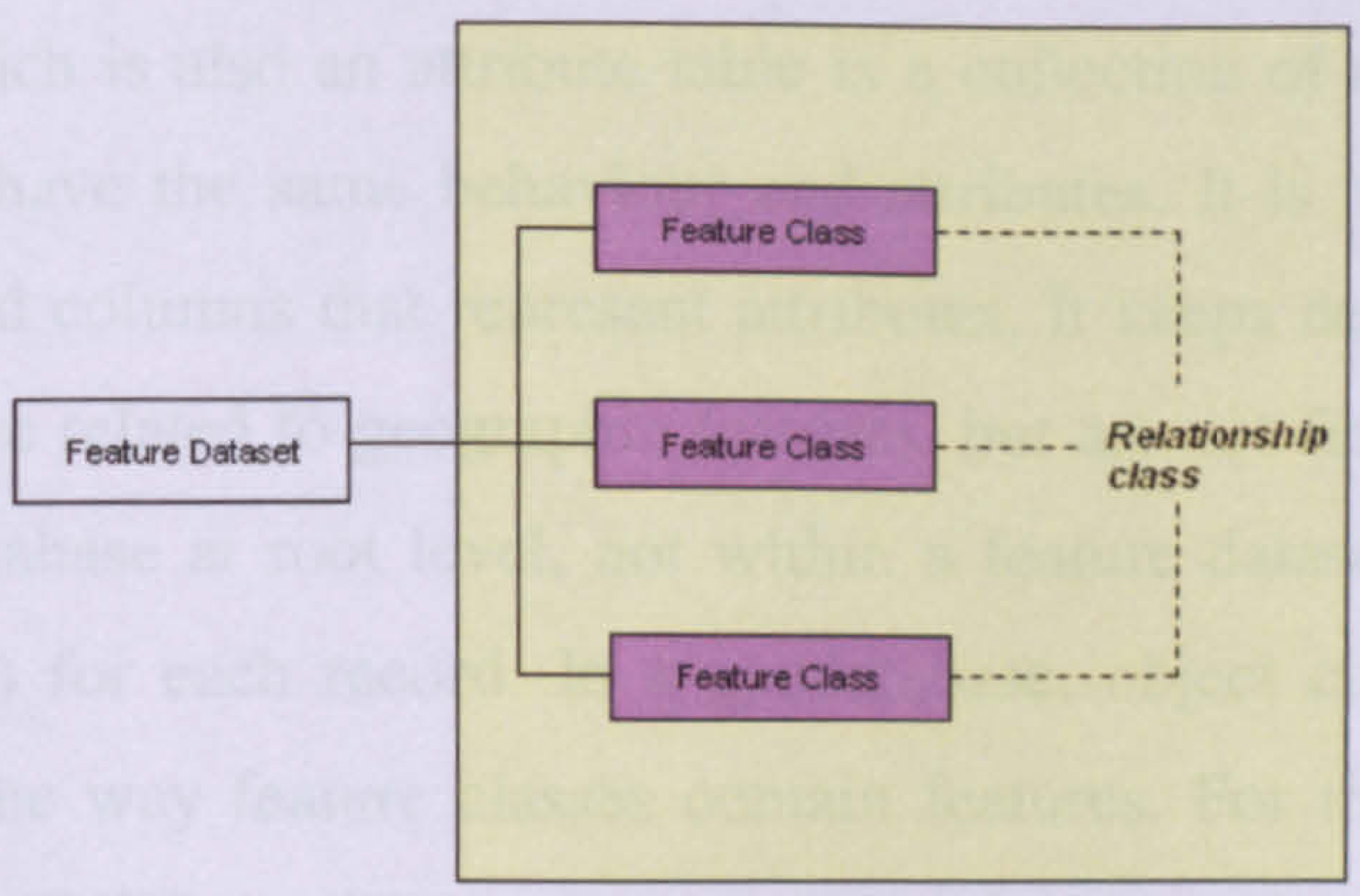


Figure 4.20: Feature dataset and feature classes

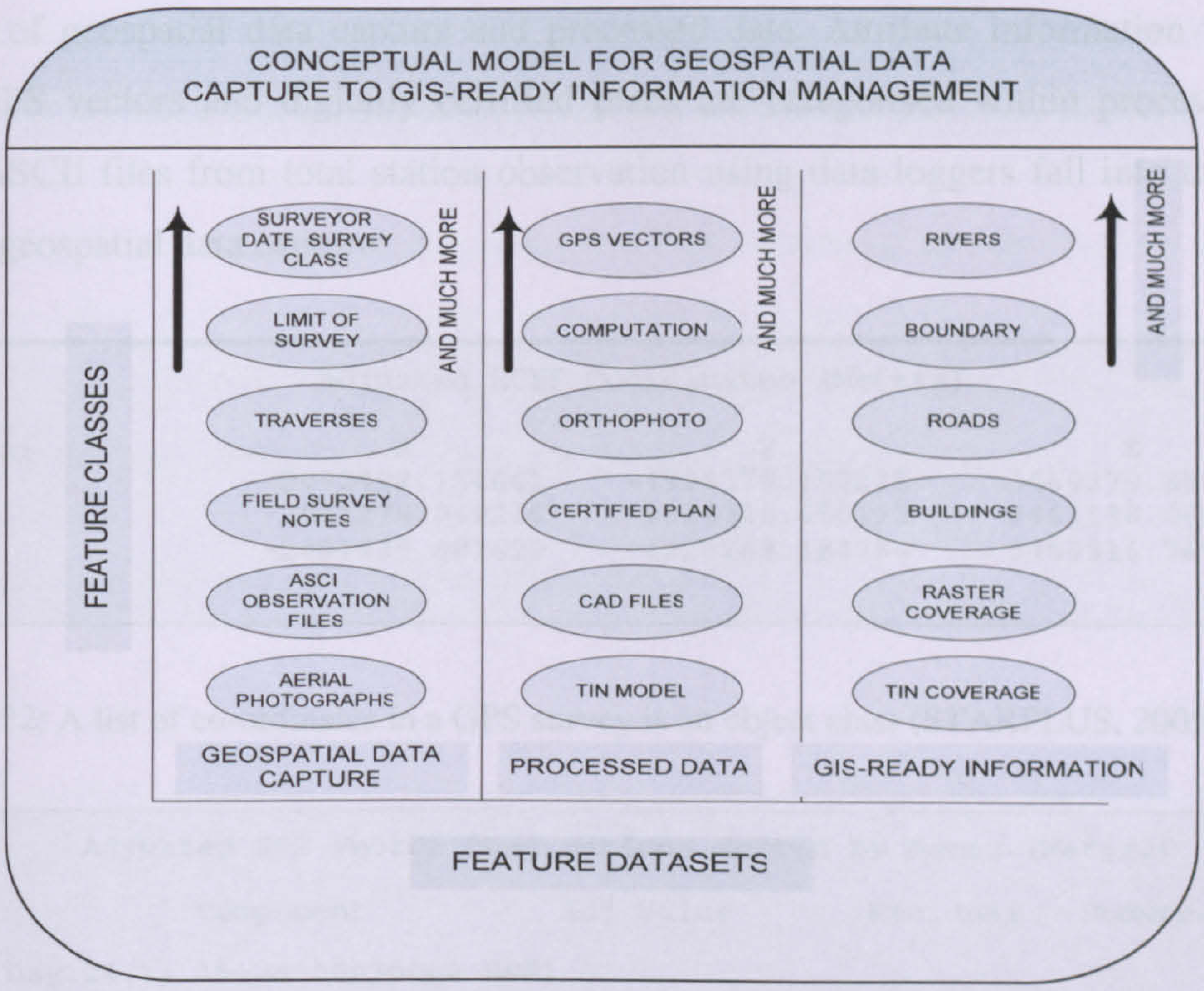


Figure 4.21: A conceptual model for geospatial data management

4.7.2.2 Object Class

An object class, which is also an attribute table is a collection of non-spatial objects in tabular format that have the same behaviour and attributes. It is a matrix of rows that represent objects and columns that represent attributes. It keeps descriptive information about objects that are related to geographic features, but are not features on a map. It is stored in the geodatabase at root level, not within a feature dataset. There is a unique identifier (ObjectID) for each record. In a geodatabase, object classes contain related objects, similar to the way feature classes contain features. For the research, the table storage format is in INFO or dBase. An example of object class for this research is survey monument for a cadastral land lot. Herewith it is possible to establish a database join between cadastral land lot (feature class) and an object class for survey monuments.

A GPS co-ordinate listing is an object class achieved from field geodetic observation or GPS survey for topographic mapping (Figure 4.22 and Figure 4.23). A total station survey produces results in ASCII files before being processed in the survey office to build vector form information. Object class or non spatial data can exist in feature datasets of geospatial data capture and processed data. Attribute information in tables about GPS vectors and digitally certified plans are categorised within processed data whilst ASCII files from total station observation using data loggers fall into the object class of geospatial data capture.

Adjusted ECEF Coordinates (Meters)

Station	X	Y	Z
0001	-2092498.154661	-4924379.150235	3460379.958174
0002	-2091276.069216	-4924320.650192	3461198.060514
0003	-2091439.881629	-4925183.184356	3459816.765030
etc.....			

Figure 4.22: A list of co-ordinates in a GPS survey is an object class (STARPLUS, 2005)

Adjusted GPS Vector Observations Sorted by Names (Meters)

From To	Component	Adj Value	Residual	StdErr	StdRes
(V117 Day124(1) 15:34 00010002.SSF)					
0001	Delta-N	975.7438	-0.0003	0.0037	0.1
0002	Delta-E	1101.8681	0.0013	0.0033	0.4
	Delta-U	0.7383	-0.0004	0.0055	0.1
	Delta-X	1222.0804	0.0013	0.0038	0.3
	Delta-Y	58.5045	-0.0003	0.0049	0.1
	Delta-Z	818.0993	-0.0005	0.0041	0.1
	Length	1471.7981			
(V109 Day124(3) 19:24 00010003.SSF)					
0001	Delta-N	-649.9083	0.0072	0.0078	0.9
0003	Delta-E	1288.4278	0.0023	0.0059	0.4
	Delta-U	-34.0012	0.0029	0.0128	0.2
	Delta-X	1058.2721	0.0027	0.0061	0.4
etc.....					

Figure 4.23: An object class of a listing of GPS vector observation (STARPLUS, 2005)

4.7.2.3 Feature Class

Basically features are objects with a fundamental shape of points, lines and polygons, which symbolise the world entities in a layer on a map. A feature class is defined as a collection of these features in tabular format with the same behaviour and attributes. In other words, it is a table with a shape field containing point, line or polygon geometries for geographic features. They are stored in the geodatabase root or in a feature dataset. Feature classes share a spatial reference within a feature dataset. In geodatabase design, new feature classes can be created or imported from coverages, CAD feature classes and shapefiles. Feature class types are mainly point, line, polygon and annotation. Each feature class can contain a set of features and each feature class can also be further divided into subtypes. In this research feature classes are data and information within the geospatial data capture, processed data and GIS-ready information which includes limit of boundary, field notes, framework control survey, sketch of survey site, text file from field observation and radial shots and much more. Each feature class can be associated to a "related object" which is actually a table (object class), as in Figure 4.24. Feature class emerges as the combination of object class and spatial references, that is,

Feature class = object class + spatial coordinates

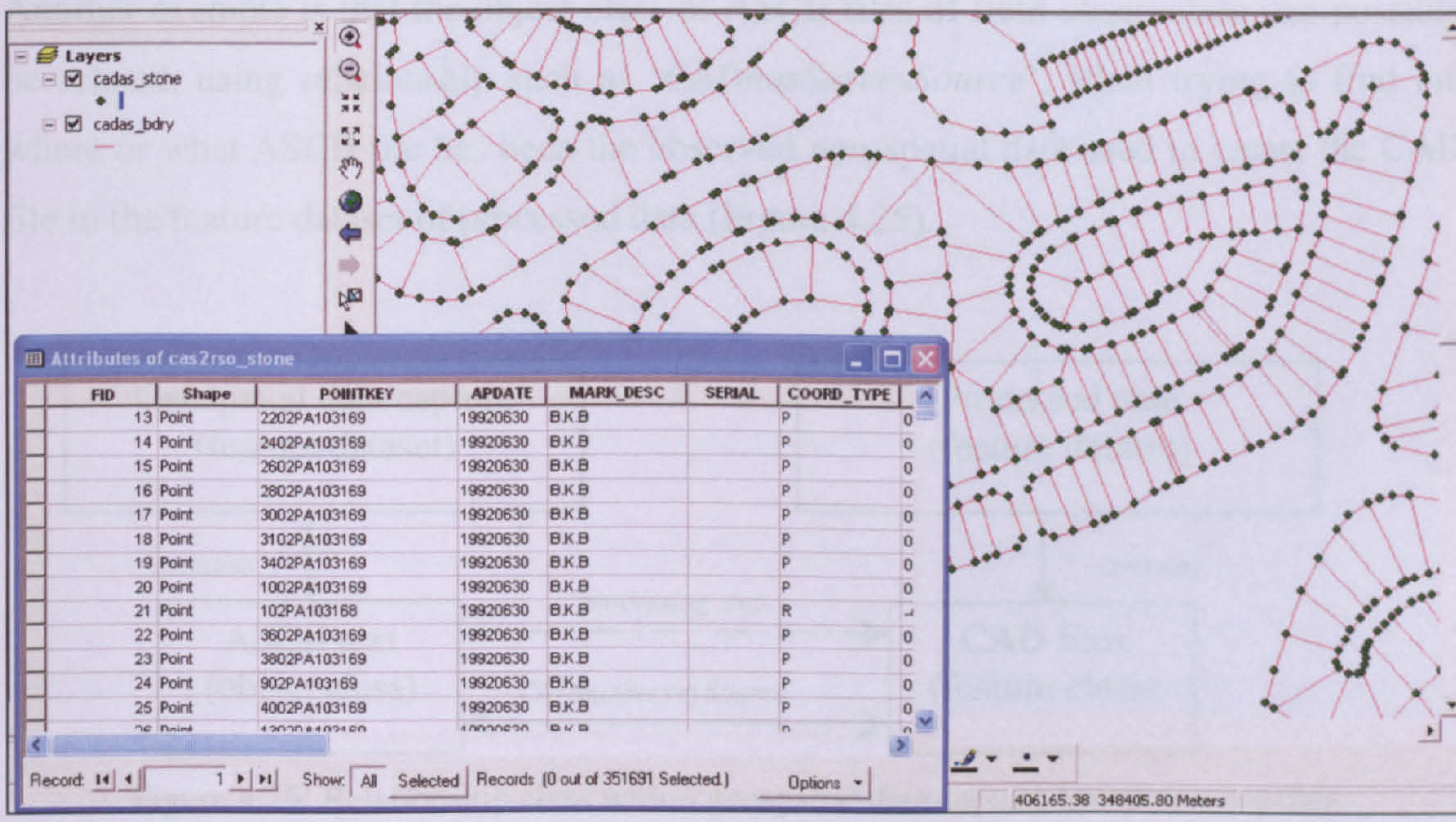


Figure 4.24: A sample of the research data (from JUPEM): survey stone (monument) and boundary in cadastral survey in the form of table (object class) and geographic features

4.7.2.4 Relationship and Relationship Class

Relationships are an association or link between two or more objects in a geodatabase. A relationship can exist between spatial objects (features in feature classes), non-spatial objects (rows in table) and between spatial and non-spatial objects. A relationship class is a table that stores relationships between features or objects in two features classes or tables. Relationships enable object dependencies between each other. Hence, with relationships, control over what happens to an object can be managed when its related object is removed or changed. Similar to the way feature classes contain features, and objects classes contain objects, relationship classes contain relationships between features and related objects. This relationship of features to related objects can be defined as one-to-one, one-to-many and many-to-many.

In an improved management of survey datasets and processes, it is possible to create relationships between data and information in geospatial data capture, processed data and GIS-ready information. An aerial photograph captured can be related in the database to CAD files produced in photogrammetry and to the GIS-ready information created in the end. The processing steps between them can act as the relationship class for all the information and data within the raw capture right through to the GIS-ready information. Another example is that the object class of ASCII files of field observation can possibly be related, using relationship such as '*CADhasSurveySource*', when trying to find out where or what ASCII file has been the observed non-spatial data used to create the CAD file in the feature dataset of processed data (Figure 4.25).

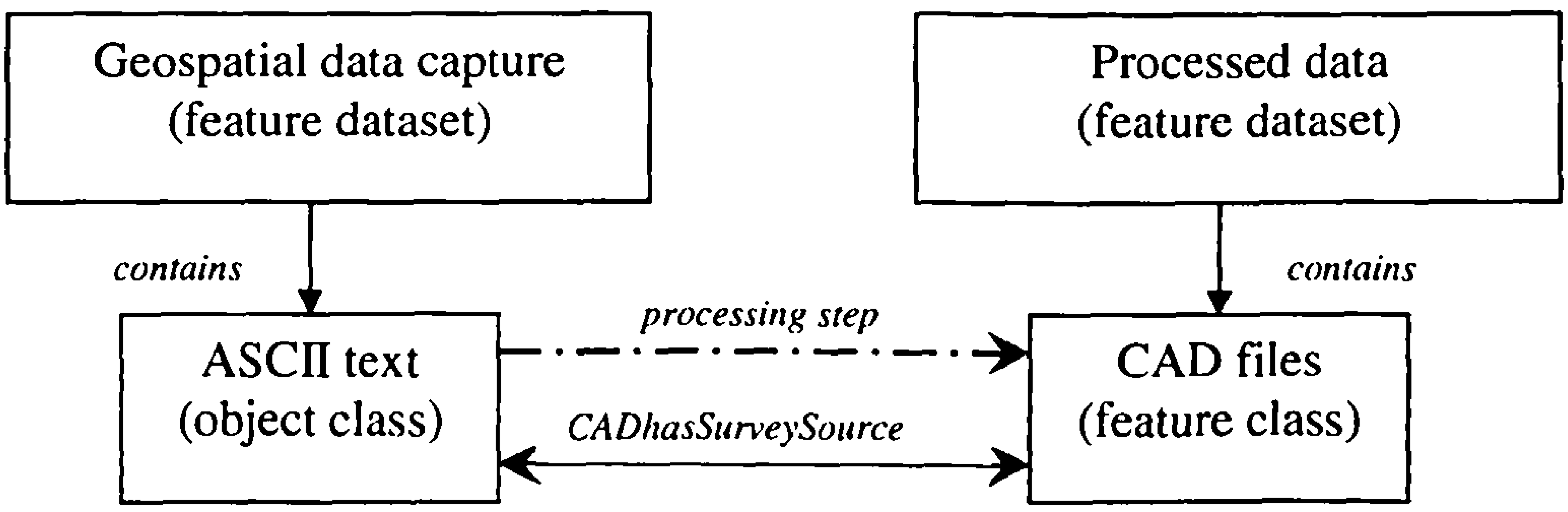


Figure 4.25: Relationship class within geospatial data capture and processed data

4.7.2.5 Raster Dataset and Raster Catalog

A raster dataset consists of a simple dataset of raster format data or several datasets with multiple bands for distinct categorical values. In other words, it can have individual images or raster tables. This component of the geodatabase enables rapid overlay of stacked raster datasets. Raster formats such as TIFF (Tagged Image File Format), JPEG (Joint Photographic Experts Group) and GIFS (Graphic Interchange Format) are composed of one or more bands. As mentioned earlier feature class, object class and relationship class are managed in feature datasets but raster data is stored as raster datasets or in a raster catalog. Geodatabase stores raster in raster datasets in a similar way it stores a geometry object. A raster column is added to a business table and each value of the raster column contains a reference to a raster stored in a separate raster table. Therefore each row of a business table references an entire raster. A raster catalog is a collection of raster with thematic properties, which may be geo-referenced with associated footprint polygons that can be spatially searched (Figure 4.26).

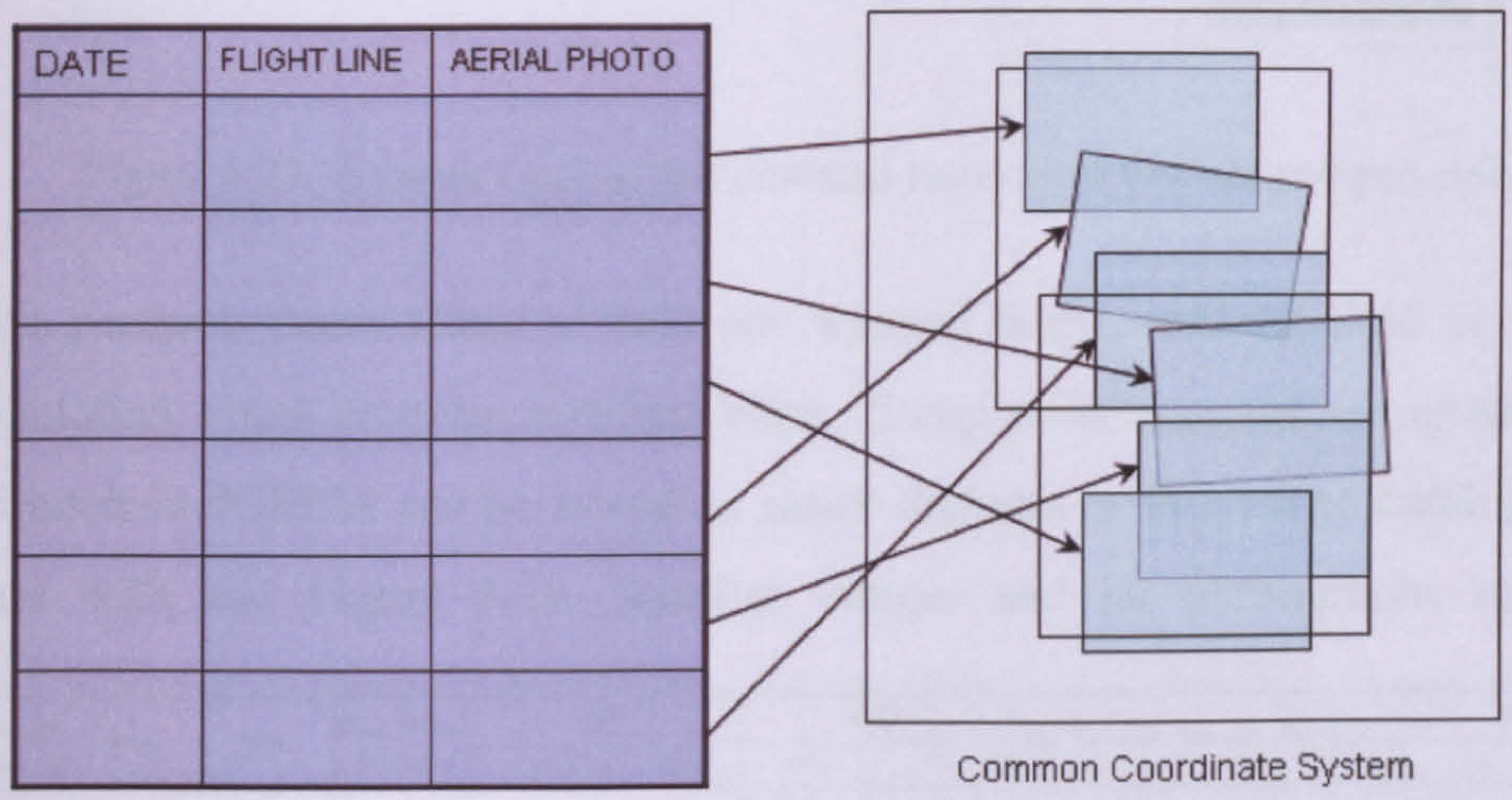


Figure 4.26: A raster catalog of aerial photo flight series

There are also raster datasets that do not need geographic reference. An image can be used as an attribute to a feature. A photograph of a home property can be an attribute in a GIS feature where a buyer can click a symbol to display an image, facts about the house and the price. These photographs may be stored in a geodatabase as a raster catalogue (Figure 4.27). Other examples of images as feature attributes are scanned documents such as deeds, field data forms associated with locations and blueprints or schematic diagrams of a master GPS station structure.

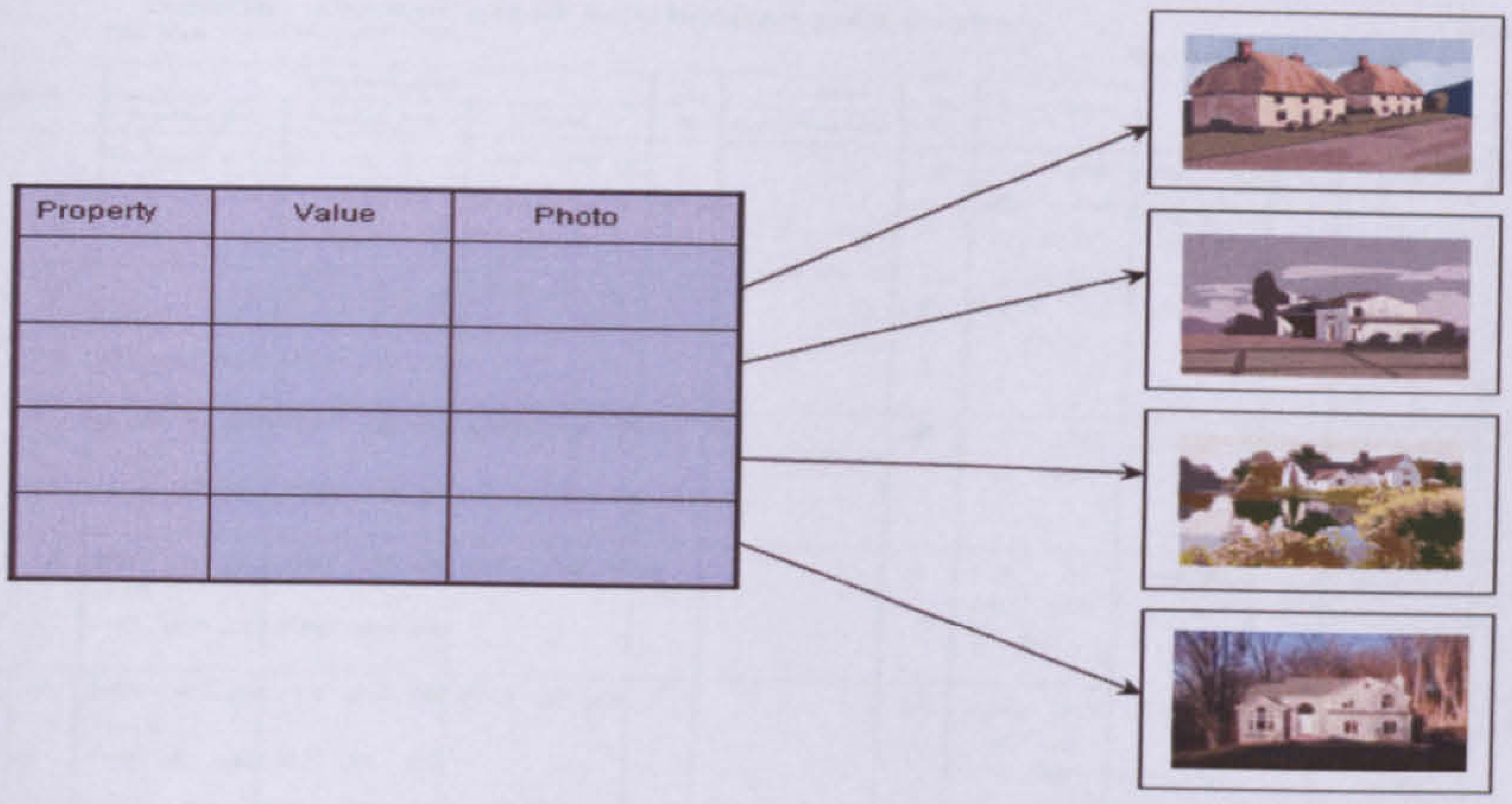
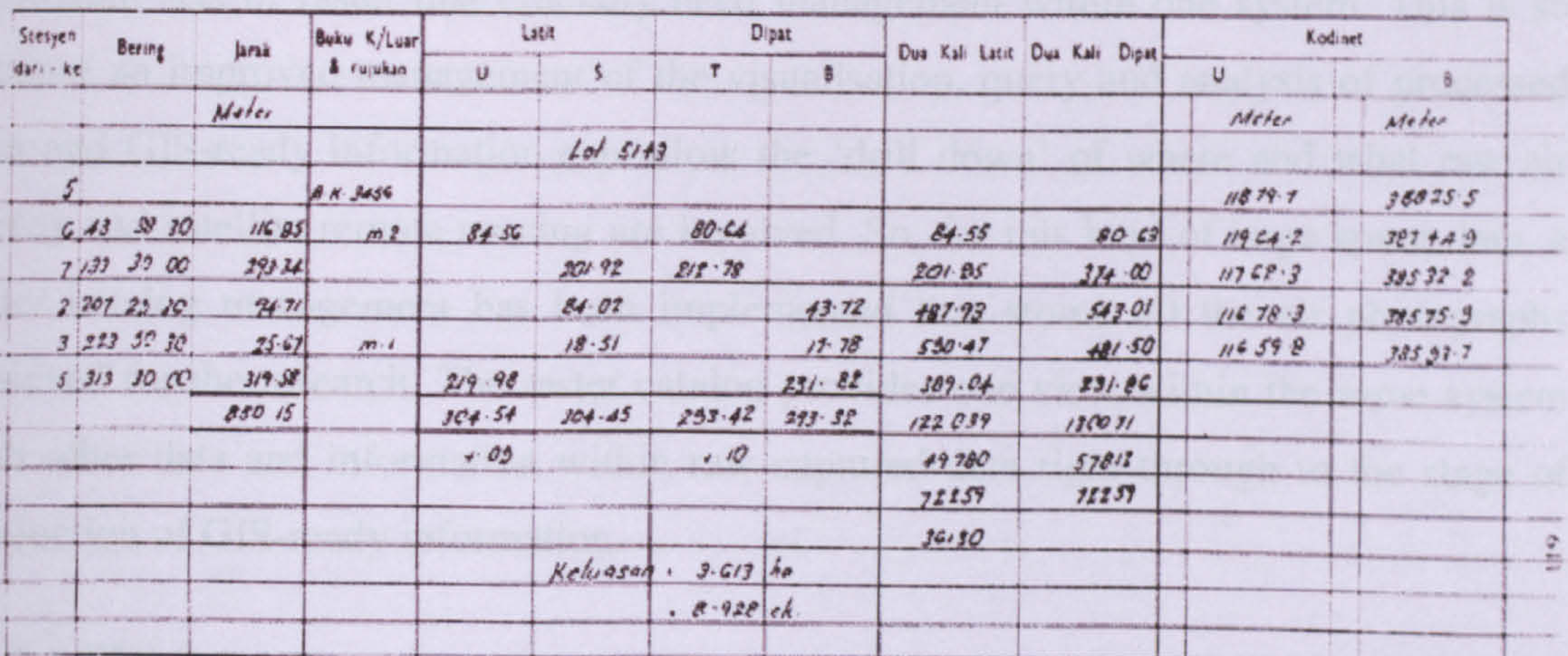


Figure 4.27: A raster catalog that contains raster data that are not geo-referenced

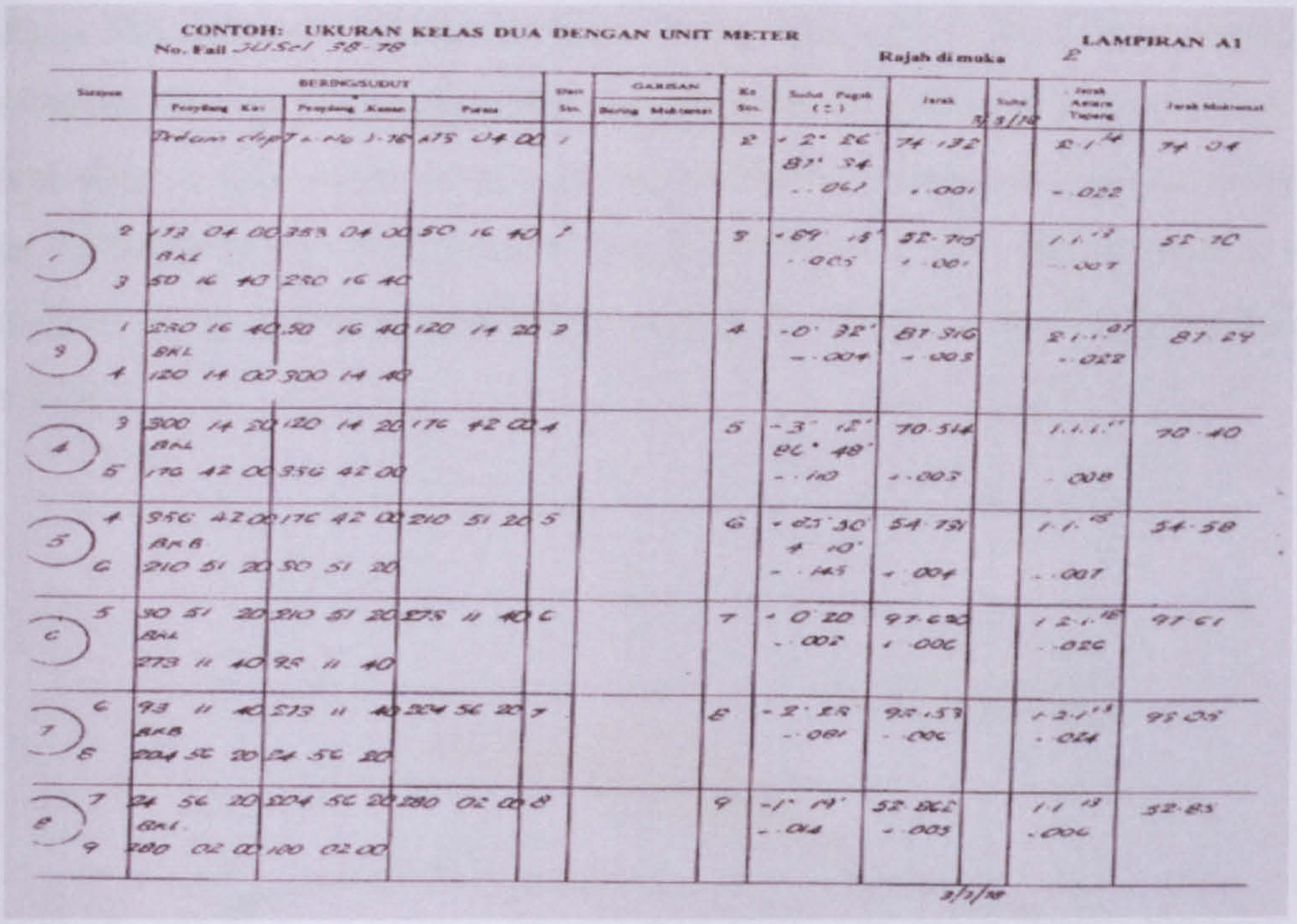
In this research raster forms of data are scanned maps, scanned field notes or scanned computation notes or even certified plans. Samples of scanned computation note and field book in JUPEM can be stored as raster datasets or in a raster catalog as shown in Figure 4.28 and Figure 4.29. Satellite images and air photographs in JUPEM are



Stesyen dari - ke	Bering	Jarak	Buku K/Luar & rujukan	Latit U	Latit S	Dipat T	Dipat B	Dua Kali Latit	Dua Kali Dipat	Kodinet U	Kodinet B
		Meter								Meter	Meter
					Lot 5143						
5			B.K. 3456							118 79.7	388 25.5
6	43 58 30	116.85	m 2	84.56		80.64		84.55	80.63	119 64.2	387 +4.9
7	133 39 00	293.34			201.92	212.78		201.85	374.00	117 62.3	385 37.2
2	207 22 20	94.71			84.02		43.72	487.93	543.01	116 78.3	385 75.9
3	223 50 30	25.67	m.1		18.51		17.78	530.47	481.50	116 59.8	385 97.7
5	313 30 00	319.58		219.98			231.82	309.04	231.86		
		850.15		304.54	304.45	293.42	293.32	122 039	130071		
				+.09		+.10		49780	57812		
								72259	72259		
								36130			
					Keluasan :	3.613 ha					
						. 8.928 ek.					
									stn. 5	U 59053.6 l/g	B. 193000.5 l/g
									stn 7	U 58470.0 l/g	B. 191542 [illegible]

Dihitung oleh Goh Seray ..... Tarikh 20.5.78 No. Kertas Ukur PUSel 1.77 ..... No. Sylt .....
Disemak oleh Rahim b Ali ..... Tarikh 26.5.78 Diukur oleh Mohaleyem ..... No. Pelan .....
Diluluskan oleh B.J. Ratnam ..... Tarikh 19.6.78 B. K Luar No. & mukasurat 3456 ..... No. Lot 5143 .....
Buku Kerjakasar & mukasurat ..... Mukim Jeram .....
456 / 123

Figure 4.28: A scanned computation note can be modelled as raster (JUPEM, 2004)



CONTOH: UKURAN KELAS DUA DENGAN UNIT METER

LAMPIRAN A1

Rajah di muka

Figure 4.29: A scanned field book can be modelled as raster (JUPEM, 2004)

significant sets of raster that crucially need management within one system. This is so because an improved management of the visualisation, query and analysis of processed data and GIS-ready information can allow the 'drill down' of where and what raw air survey and satellite remote sensing are involved. So, for this kind of huge space data, a raster catalog management has been implemented that stored all the air photographs achieved for the research. The raster catalog provides one view within the same system with other data and information within raw captured data right through to the stage of production of GIS-ready information.

4.7.2.6 TIN Dataset

In a geodatabase, a TIN dataset consists of a set of triangles that exactly span an area with a z value for each node that represents some type of surface. Inherent GIS tools with three-dimensional extension within a geodatabase allow the display of perspective views of surfaces with draped images, contours, grid lines or other features. An example of the tools are ArcScene from the ArcGIS software suite that enables the creation of TIN and interactive visualisation and analysis.

Basically, a TIN dataset is built by using a set of points with X, Y, Z co-ordinates which were collected through photogrammetric instruments, GPS data and other means. From this point data, a GIS application such as ArcScene creates an optimal network of triangles, called Delaunay triangulation (Zeiler, 1999). In a TIN, each triangle is created to be as close to equilateral as possible. Each triangle forms a face with a gradient slope (Figure 4.30).

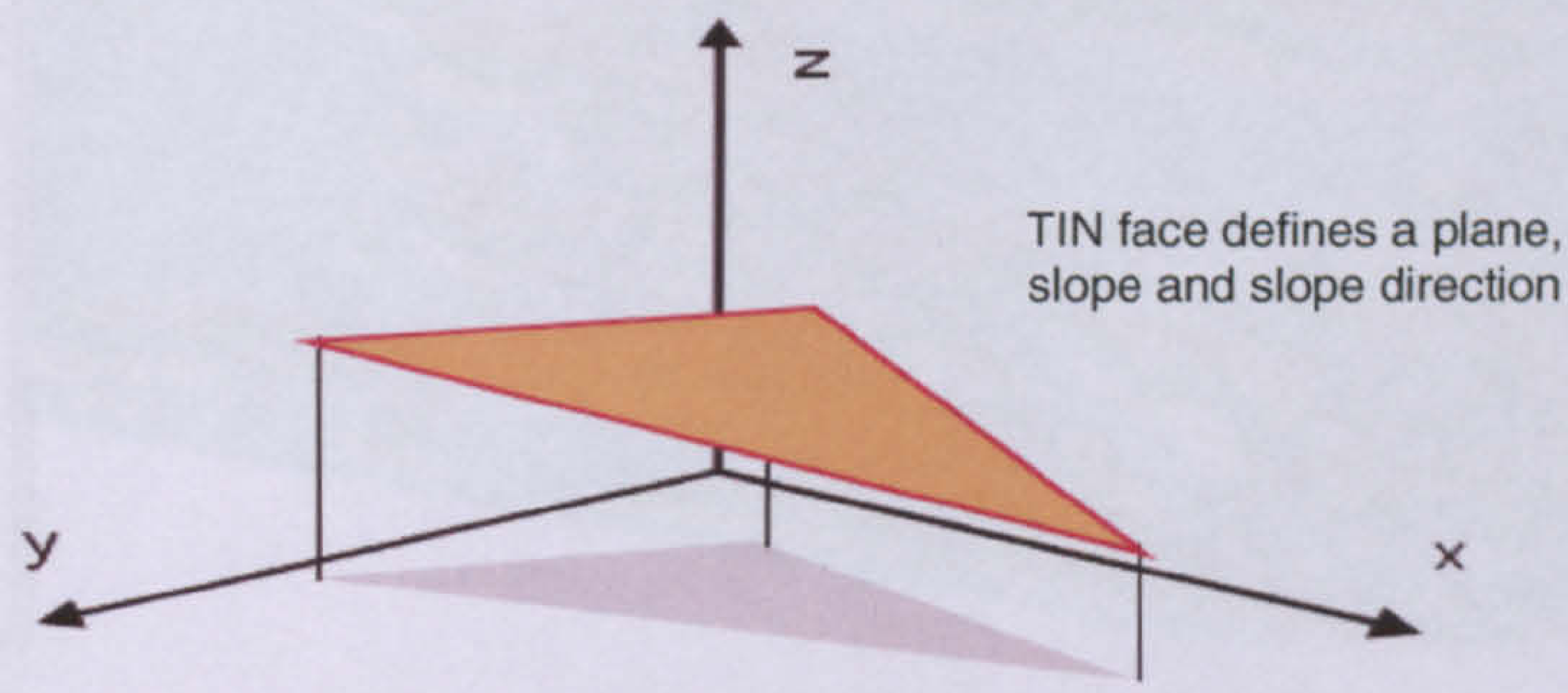


Figure 4.30: A TIN face is a triangle floating in 3-D space

JUPEM has surface data that may be managed in a system within a raw survey dataset, processed data and for the geospatial management. There are normally digital terrain models (DTM) produced using the photogrammetry method. Field topographic surveys and air surveys are the raw surveys that produced DEM data that is categorised as processed data in the GIS-ready information stage. Field topographic surveys produce elevation data for all details collected and for the generation of contours, which then can be used for the production of DTM data. Contour surface data is part of processed data in this case.

Air photographs produced in air surveys are rectified to become orthophotographs, which are then scanned using a modern digital scanner leading to the process of triangulation and digitisation of the captured features to create CAD data. A mapping software called SOCET SET, an extensive package for digital photogrammetric workstations, is used for this purpose. DTM data is produced from CAD files using Microstation software. DTM and TIN coverage with real world spatial reference is GIS-ready information. Figure 4.31 shows the TIN model generated for an area, Damansara in Kuala Lumpur. It is created from CAD data using inherent software that extends to generate surface data. The processing step from processed data (CAD data) is thus achieved to provide GIS-ready data (DEM data) within the same system of geospatial data management.

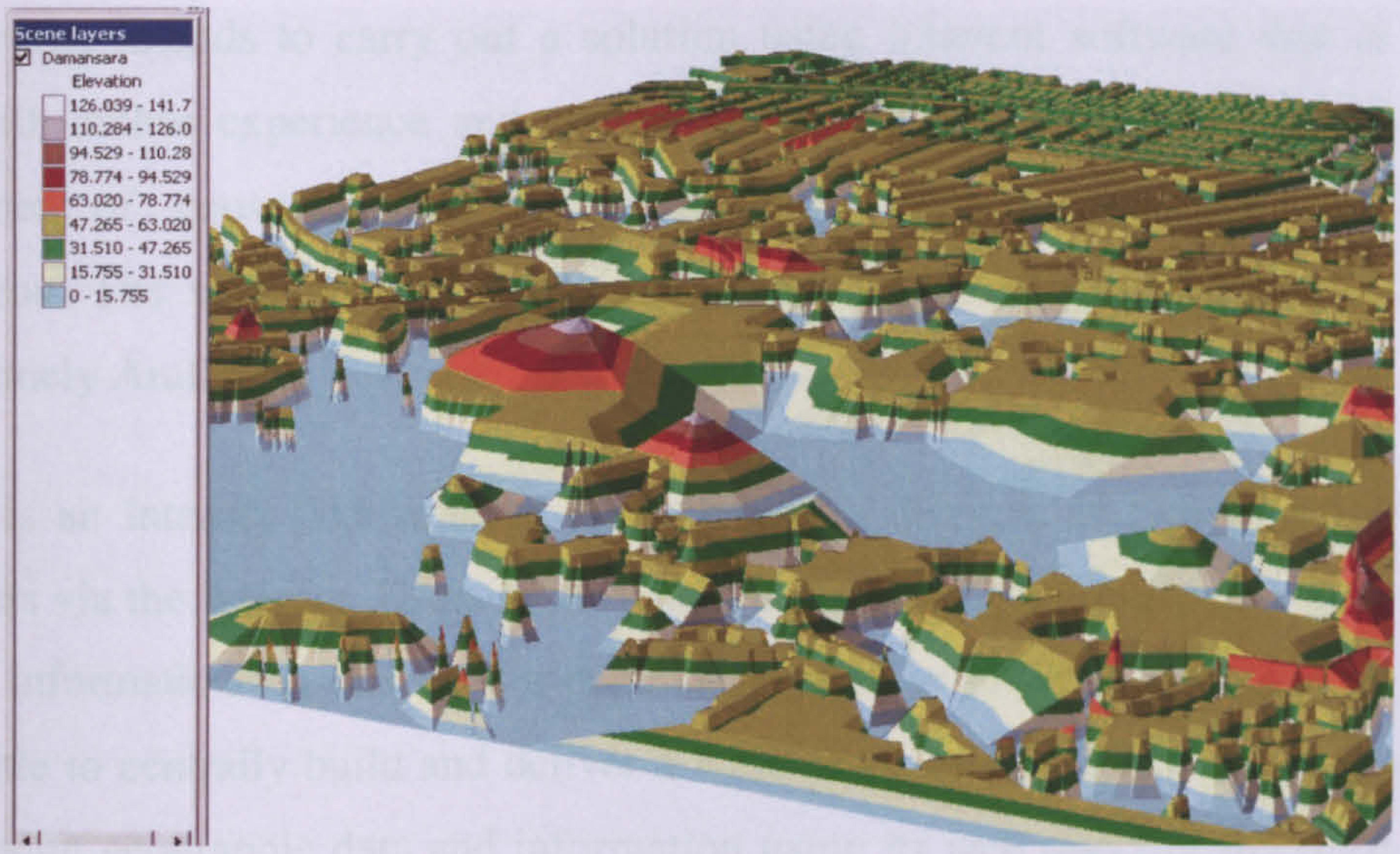


Figure 4.31: A TIN model from CAD file used for the research

4.8 Delivery of GIS-Ready Information

Extensive adoption of GIS technology has accumulated a large amount of geospatial data. The need for this information to be made available online has provided sharing and access of data among staff, organisations and the public. Inter-department solutions to data access have been provided using Intranet networks whereby internal staff are able to use data for business strategies or development. Internet solutions provide access to the wide range of the public domain. This section examines the technology of vendor specific software and Open Source Technology to deliver GIS-ready information.

4.8.1 Delivery Using Vendor Specific Application

Vendor specific software for delivery of geospatial data across the Internet has been developed since the Internet was introduced. Commercial GIS software vendors and developers compete to provide solutions to the access of geospatial data. There are many vendor solutions to provide interfaces and applications within GIS for data to be accessed and transferred into different locations in the world. There is a current trend of these applications being developed by vendors such as Mapinfo, Intergraph, ESRI, Autodesk, LaserScan, PCI Geomatics. There are also some small scale companies that are able to develop their own solutions to be used by small scale organisations such as public authorities and retailers.

This research intends to carry out a solution using inherent software that is licensed, suited with author experience and can be integrated with the GIS tools used for the management of geospatial data capture right through to the production of GIS-ready information. The software comes together with the GIS software ArcGIS suite from ESRI, namely ArcIMS.

ArcIMS is an Internet GIS solution that allows the delivery of GIS data and maps to other users via the Internet. Users in the same organisation and outside are able to access data and information via this vendor-dependent application. With this software, the GIS user is able to centrally build and deliver a wide range of GIS information. It is capable of displaying geographic data and information using its web operator interface. It is the only software that enables users to simultaneously access Web data, local shapefiles, Spatial Data Engine (SDE) layers and images for display, query, and analysis in an easy-to-use Web browser (Tang and Selwood, 2003).

ArcIMS supports the need and requirement of most organisations for analysis, business planning and decision making using graphic live presentation with significant GIS functionality (Peng and Tsou, 2003). The software ArcIMS Version 4.0.1, used in this research is designed to easily create maps and GIS data from any data sources connected with it. The manager component has been designed to administer the Web mapping site. Basically, ArcIMS allows Web clients, map servers, data servers and Web servers to communicate with one another.

There are many reasons why ArcIMS is popular for Internet GIS. ArcIMS provides scope for expansion of GIS data across the Web with its set of HTML and Java Viewers that can be freely distributed to users. ArcIMS lets users exchange, integrate, and analyse data by combining data and information accessed via the Internet with local data for display, query, and analysis. ArcIMS can be scaled and extended to fit the needs of map services as the volume of requests increases. Data holdings and accessibility can be controlled with choices of public access via the Web, on a secure network, or within the organisation on a local area network

ArcIMS runs in a distributed environment and comprises both the client and server components. The ArcIMS architecture has been developed specifically for Internet applications. It has a multilevel architecture consisting of presentation, core business level, and data levels (ESRI, 2003b). In addition, ArcIMS has a set of applications for managing Web mapping sites (Figure 4.32).

- The presentation level includes the ArcIMS client viewers for accessing, viewing, and analysing geographic data.
- The components in the core business level are used for handling requests and administering the ArcIMS site.
- The data level includes all data sources available for use with ArcIMS. Data sources can be either from a variety of file formats or from a relational database system via the application named ArcSDE.
- The manager application provides Web-based interfaces to support three main tasks in ArcIMS application in i.e. authoring services, designing Web-pages and administer sites.

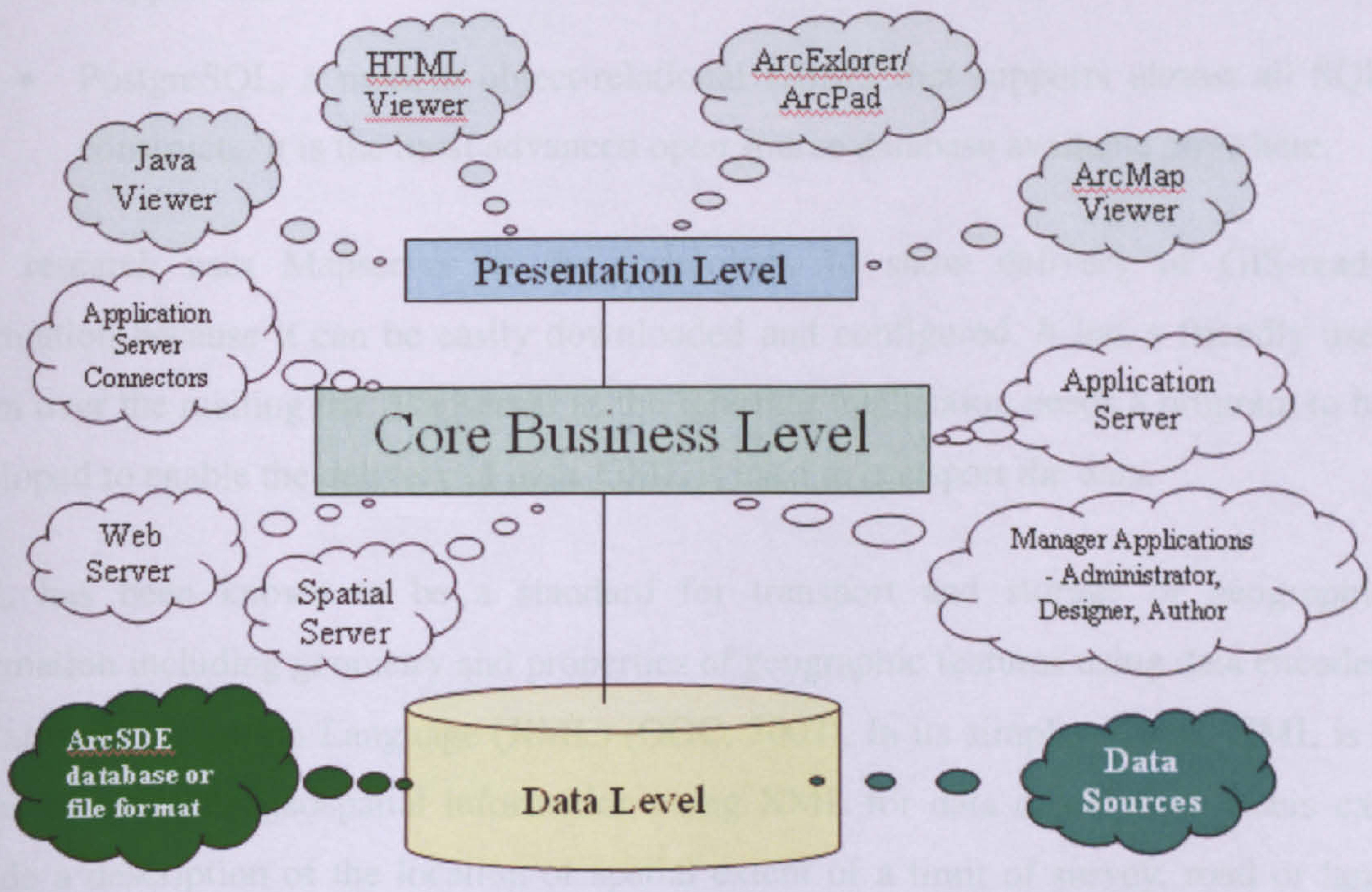


Figure 4.32: ArcIMS multi-level architecture (after ESRI, 2003b)

4.8.2 Delivery Using Open Source Technology

Open Source Technology is a technology for GIS whereby software and application for spatial data delivery and management uses open source software that is available over the Internet. This technology has achieved a broad diffusion covering all application in the Earth environment and land management field (Diviacco, 2005).

GIS based on open source software is nowadays common practice for public institutions and professionals because they are freeware and can be easily accessed and used. The open source software systems available to manage and deliver geographic information include (Neteler and Mitasova, 2002):

- GRASS (Geographic Resources Analysis Support System). It is a GIS that manages many types of geospatial data including 3D data and images.
- Mapserver, an open source development environment for building spatially enable Internet application.
- GRASSLinks, a WWW interface to GIS application that offers public access to mapped data.
- PostgreSQL, a modern object-relational DBMS that supports almost all SQL constructs. It is the most advanced open source database available anywhere.

This research uses Mapserver as the technology to show delivery of GIS-ready information because it can be easily downloaded and configured. It has a friendly user forum over the mailing list. Mapserver as the interface application needs a program to be developed to enable the delivery of data. GML is used to transport the data.

GML has been known to be a standard for transport and storage of geographic information including geometry and properties of geographic features using data encoded by Extensible Mark-up-Language (XML) (OGC, 2003). In its simplest terms, GML is a means of encoding geospatial information using XML for data description. Users can encode a description of the location or spatial extent of a limit of survey, road or land parcel which captures its geometry and other properties including surveyor, type of survey, road classes and survey monument and boundaries in Malaysia (Mukim or State).

There are clear distinctions between geographic data, which is encoded in GML and graphic interpretation of that data as it appears on a map or any other form of visualisation. Geographic data in GML is concerned with a representation of the world in spatial terms that is free of any visualisation of the data.

If the data representation is made on a map, which colours or line weights being used is very different when GML is concerned. As shown in Figure 4.33 and Figure 4.34, for comparison just as XML helps the Web to clearly separate content from presentation, GML does the same in the world of geography (Lake, 2005). In the evolution of geospatial information infrastructure, GML is as critical as HTML was to the development of the conventional Internet.

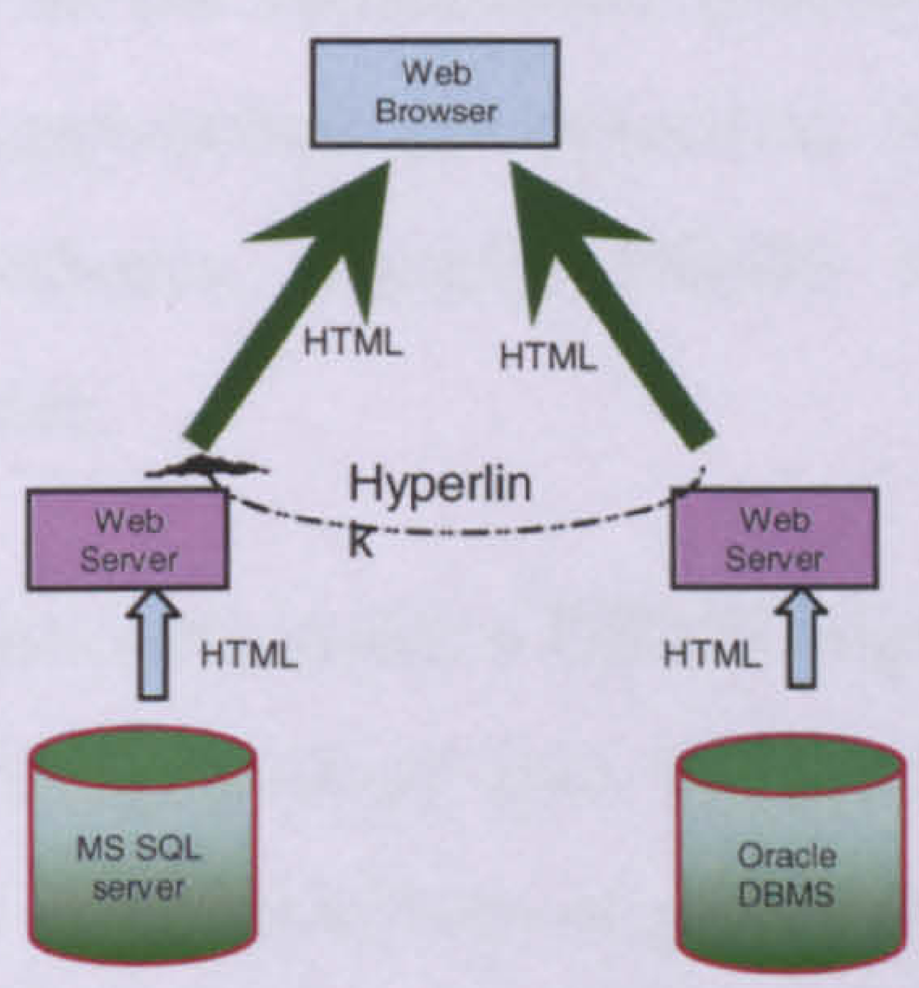


Figure 4.33: HTML version in Internet development
(after Peng and Zhang, 2004)

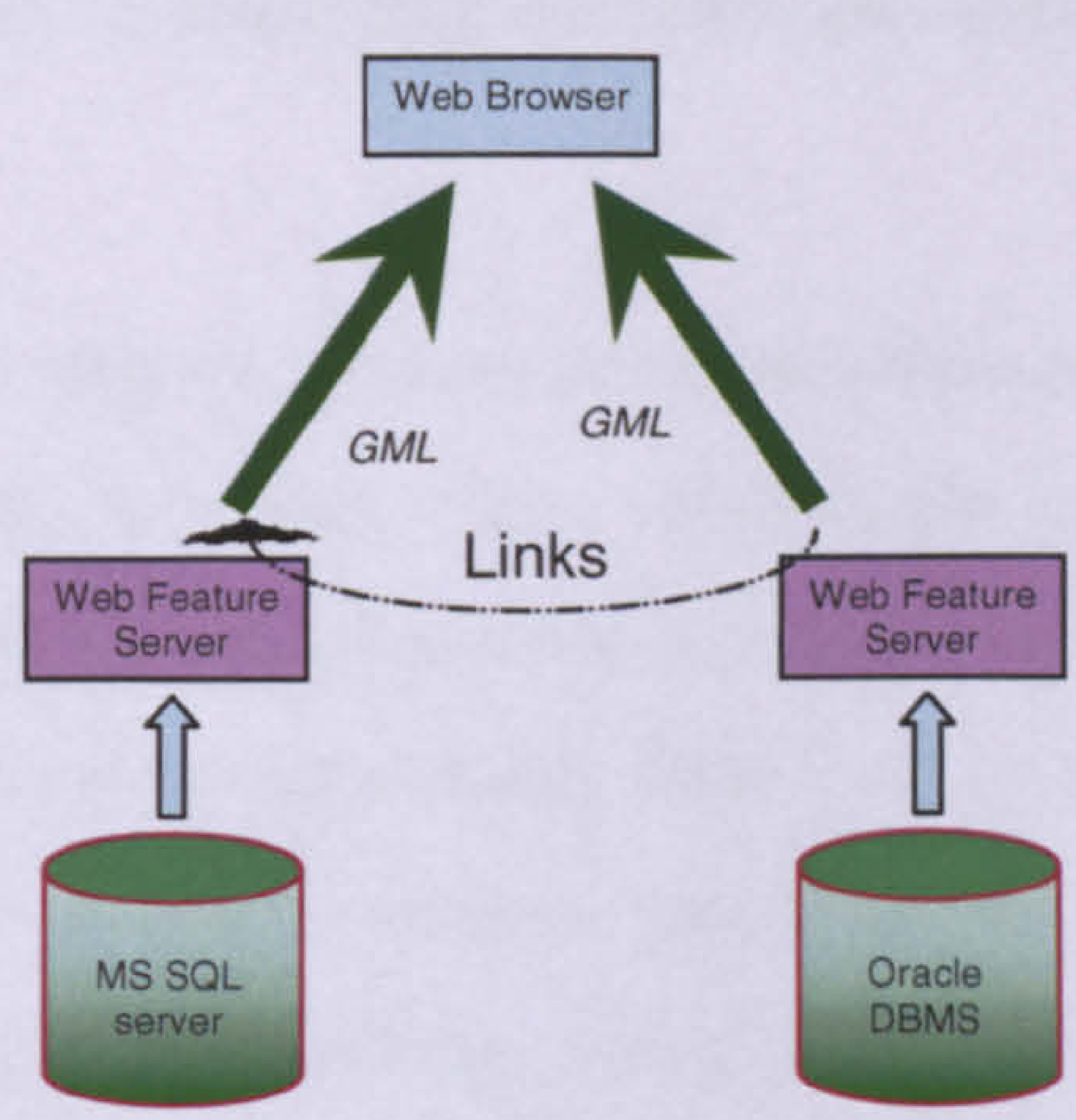


Figure 4.34: GML for geographic information links
(after Peng and Zhang, 2004)

4.9 Chapter Summary

The content of this chapter has encompassed some aspects of survey work flow and kind of datasets that exist in JUPEM and description of how geospatial modelling can be applied to support the management of surveys and datasets. The chapter then focused on the data modelling, data management approaches and aspects of geographic database as core concepts for the management and delivery of the surveys and datasets. Generally there are three types of data representation models as stated by Zeiler (1999): they are vector data, raster and triangulated data modelling. Data models were discussed with relation to data used in the research and produced by the JUPEM. Spatial modelling techniques using the basic flow of conceptual, logical and physical models were discussed, which would serve as the fundamental process for data modelling within a database. Data management approaches are described leading to the more thorough discussion of geographic databases. Clearly, DBMS should not be separated for management by GIS environment.

As geodatabases store geographic data inside a DBMS engine and are object-relational in nature, they offer efficient GIS technology into the realm of mainstream information technology. Details of benefits and capabilities of geodatabase models were examined to provide senses for the implementation of the geospatial data capture to the GIS-ready information management process. The geodatabase has been seen to provide the capability to address issues of managing complex geospatial data from raw until to the GIS end user.

The use of a geodatabase and its element to model datasets in JUPEM was described. Feature dataset, feature class, object class, relationship class, raster catalog and TIN model are relevant technologies that have the capability and concept to model objects in a national survey and mapping department from raw to the production of GIS-ready information. All datasets and information found in the JUPEM are suited for one management system using a geodatabase. The delivery of data has been examined and discussed focusing on vendor specific and Open Source Technology. The research uses these two technologies to develop GIS-ready information delivery and contrast them.

The next chapter implements a DBMS that leads to the prototype of the management of datasets and information from geospatial data capture right through to the GIS-ready stage.

Chapter 5

A prototype of a geospatial data management system

5.1 Introduction

The concepts of object, database model and UML technology were discussed in Chapter 3. In Chapter 4, some aspects of survey work flow and kind of datasets that exist in JUPEM and description of how geospatial modelling can be applied to support the management of surveys and datasets were described. Data models, management models, database models and methods of delivering data online were also discussed. How these models transpired to be suitable to model JUPEM datasets was revealed. The conceptual model derived is an impression of three objects represented as feature datasets, namely geospatial data capture, processed data and GIS-ready information. Feature datasets' down-the-line geospatial objects of feature classes, non-spatial object classes, raster and TIN model can be very useful for JUPEM datasets modelled together in a geodatabase. The conceptual model needs to be converted to a logical model and a physical model. In this chapter, the database (DBMS) is implemented and tested, and spatial and non-spatial data objects are loaded into, processed and tested as the final product in the GIS-ready information stage.

A proprietary GIS technology from ESRI is used for the implementation of a prototype. Using ArcGIS tools of module, an improved management of the raw survey, processed data and GIS-ready information is shown. The reasons why these specific tools are used will be discussed in the next section.

5.2 Why Use ESRI's ArcGIS?

There have been many GIS projects, spatial database designs and GIS processing implementations which utilise current vendor dependent technologies such as Autodesk (Ahern, 2004), Intergraph (Dugacek et. al, 1997) and Mapinfo (Turker and Kocaman, 2003). ESRI is another leader in GIS software and services. These vendors with their up-to-date software are capable of delivering quality services to the GIS community. In this research, ESRI's products including Erdas Imagine, ArcGIS, ArcSDE and ArcIMS are used as the major software to implement the research prototype. There are a few reasons for using them: firstly, they are very widely available and licensed in the University of Newcastle upon Tyne's Geomatics Laboratory; secondly, the online technical support through the user's forum, white papers, conference papers, knowledge base articles, extension software downloads, developer's forum, and live seminar and courses is seen to be more user-friendly, helpful and more interactive when compared to others. In addition, the author's organisation had been using ESRI's previous product, ArcView 3.0, and had human resources very much used to ESRI's software, its popularity among most spatial data dependent government and private office, and its establishment being strong in Malaysia and Singapore. Most preliminary data in the cadastral land parcel database, which was set to be the base map for national land administration, were in ESRI's formats. Moreover, ESRI's agent in Malaysia has been the pioneer in the development of the country's national spatial data infrastructure (NSDI) since its first pilot project in 1996. Currently, ESRI is on its way to implementing the upgrading of the NSDI delivery system. The application and system upgrading of the country's biggest spatial database development, the cadastral database system, is currently going ahead with the use of ESRI's products.

Developing countries are beginning to develop huge spatial data capture activities and storage as well as making them GIS-ready by contracting and outsourcing them to private ICT, GIS and software professionals. By using the related technology in this research, it is also intended that professionalism and training in ESRI's GISs, hands-on spatial data handling and experiences be achieved. It is hoped that experienced and trained government staff will develop their integrity and reliability in facing modern and professional vendors who are agents or GIS distributors from developed countries. Eventually, there should be an equal level of ability in project proposal development, delivery and operation of staff within the currently developing countries with those in the currently developed countries. The public sector is not going to depend fully on the private sector when system and application development have to be implemented by contractors who are from developed countries. Finances can be saved and more GIS experts can be created within the government. Government spends huge amount of money to get hold of ICT and spatial data development by outsourcing and contracting the whole project without the vendors transferring the technology to the government.

Funding is then lost again in the upgrading exercise by the same contractor whereas if there was experienced staff a better idea could be foreseen to have the system and application prolonged for a longer time before modifying them and dumping the old system.

The ArcGIS family of products is an integrated system that can provide excellent tools to store, manage and process data from raw captured stage up to the point where geospatial objects can be produced and disseminated as GIS-ready information to the clients. The flow of storing, processing and converting the data and information, spatial and non-spatial, can be very smooth with a standard unique platform that can be used and accessed within the same system by multiple users (Figure 5.1). Considering the voluminous raw survey data, processed data and the product GIS-ready information in JUPEM, this system is admirably suitable and has great influence on the datasets (elements of geodatabase).

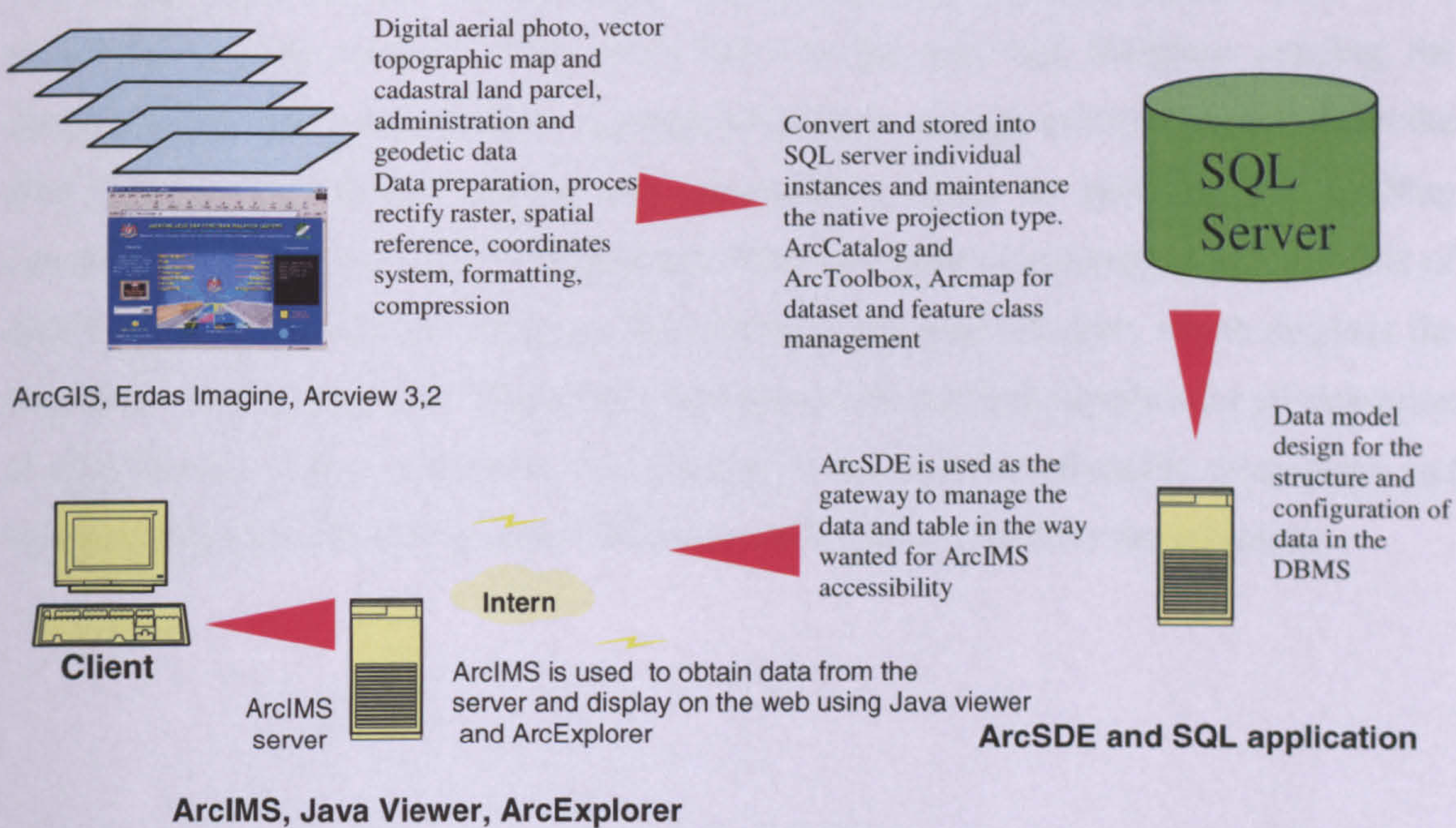


Figure 5.1: The integrated system for all data and information from raw capture up to the GIS-ready information stage

The following section briefly describes some modules of the ArcGIS family of products.

5.3 ArcGIS's Desktop Modules

This section describes two modules in the ArcGIS desktop engine that are used generally to map data and manage data in the database. These modules are ArcMap used as a mapping tool and ArcCatalog as a data operation tool. ArcToolbox is another component but will not be discussed in this section because it has not been used.

5.3.1 ArcMap

A program under ArcGIS, ArcMap is the focal component in the software. There are two other components namely ArcToolbox and ArcCatalog. ArcMap plays a main role in presenting, visualising, analysing and operating map and data with its interface. It enables viewing, adding and changing the way data is represented thus providing data analysis and interactive visualisation. ArcMap allows the creation of new maps and presentations of the data in layouts. ArcMap's maps and data are saved in ArcMap documents, which have an MXD extension file. Basically these documents do not store or manage copies of data. They store links to the data and therefore copying the document will also copy the links. ArcMap documents manage information as to how the data is to be represented, queried and presented in layout for printing. The ArcMap interface consists of two main components. First, the table of contents shows the lists of data that has been added to the map. The second is the map window, which displays the data in the format specified. These two interfaces work and link together for all operation of map layers. Figure 5.2 shows the display of the polyline shapefile (transport) and building polygons, the end product GIS-ready information, used for the research.

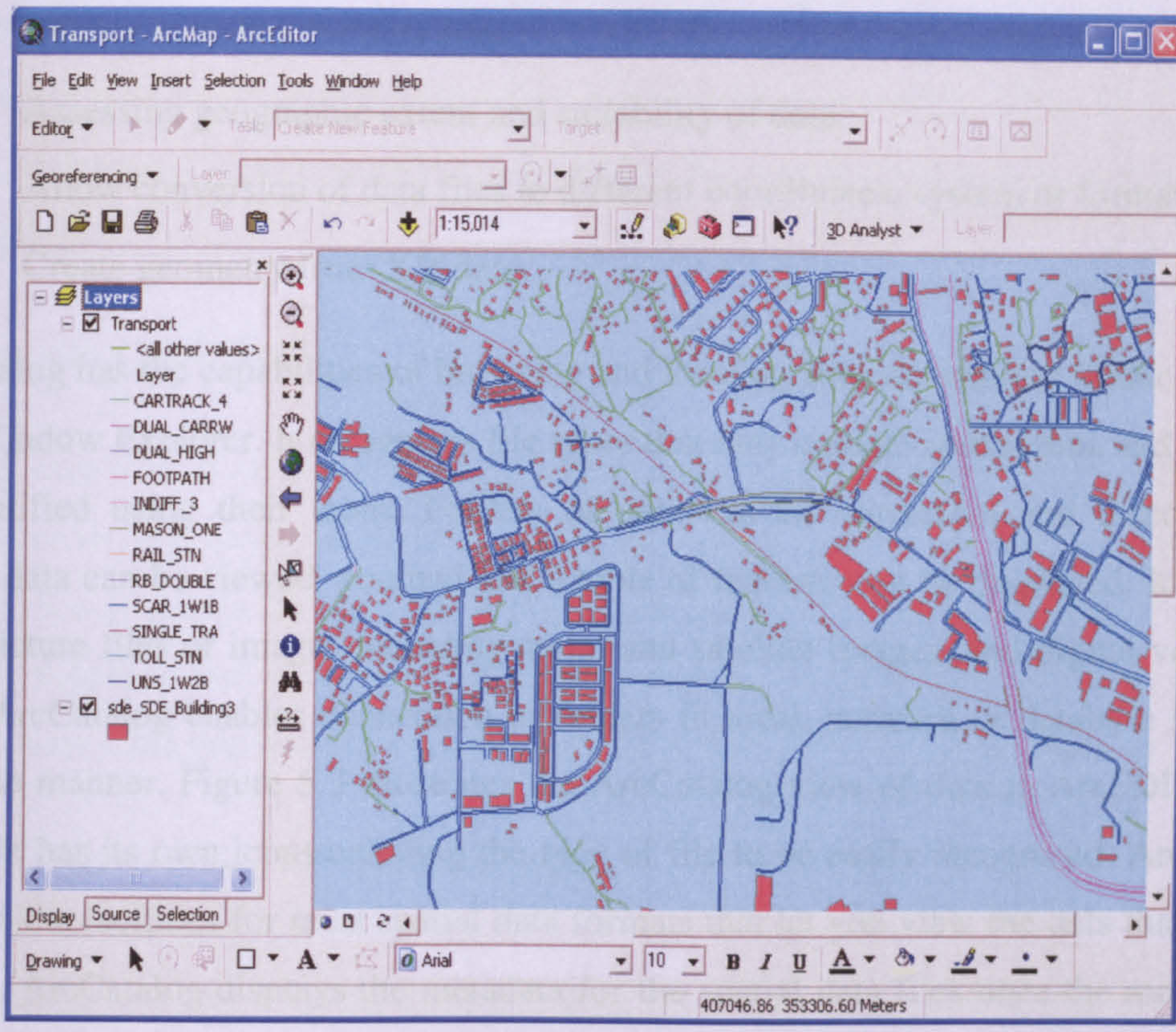


Figure 5.2: A sample data for the research is viewed in ArcMap

5.3.2 ArcCatalog

ArcCatalog is another interface component that works as an application tool for exploring, accessing and managing data from various sources. This application plays the role of user interface to establish, change and refine the structure of the geodatabase. It is designed to handle data-oriented procedures which may include the following capabilities:

- Browsing for spatial data files.
- Enable data files be copied and moved.
- Allow connection to spatial database server.
- Create and format geodatabase and new data files.
- Assigning coordinates systems to data files.

- Editing metadata for data files.
- Assessing geographic extent and suitability of data.
- Allow conversion of data files to different coordinates, system or format.
- Create geometry from XY data.

ArcCatalog has the capabilities of browsing and locating files resembling folder concepts as in Window Explorer. It recognises file types that may contain spatial data and they can be identified using their icons. Folders of data can be connected and disconnected. Spatial data can be viewed, zoomed and a table of features can be displayed. It can also show picture files or images including aerial and satellite images, and digital version of maps. ArcCatalog enables connection to folders in local, network or database server in the same manner. Figure 5.3 illustrates the ArcCatalog view of data in ArcSDE Server. Each file has its own icons enabling the type of file to be easily recognised. ArcCatalog has preview facilities for most spatial data formats that let you view the data that is right for you. ArcCatalog displays the metadata for the spatial data files once the metadata is created.

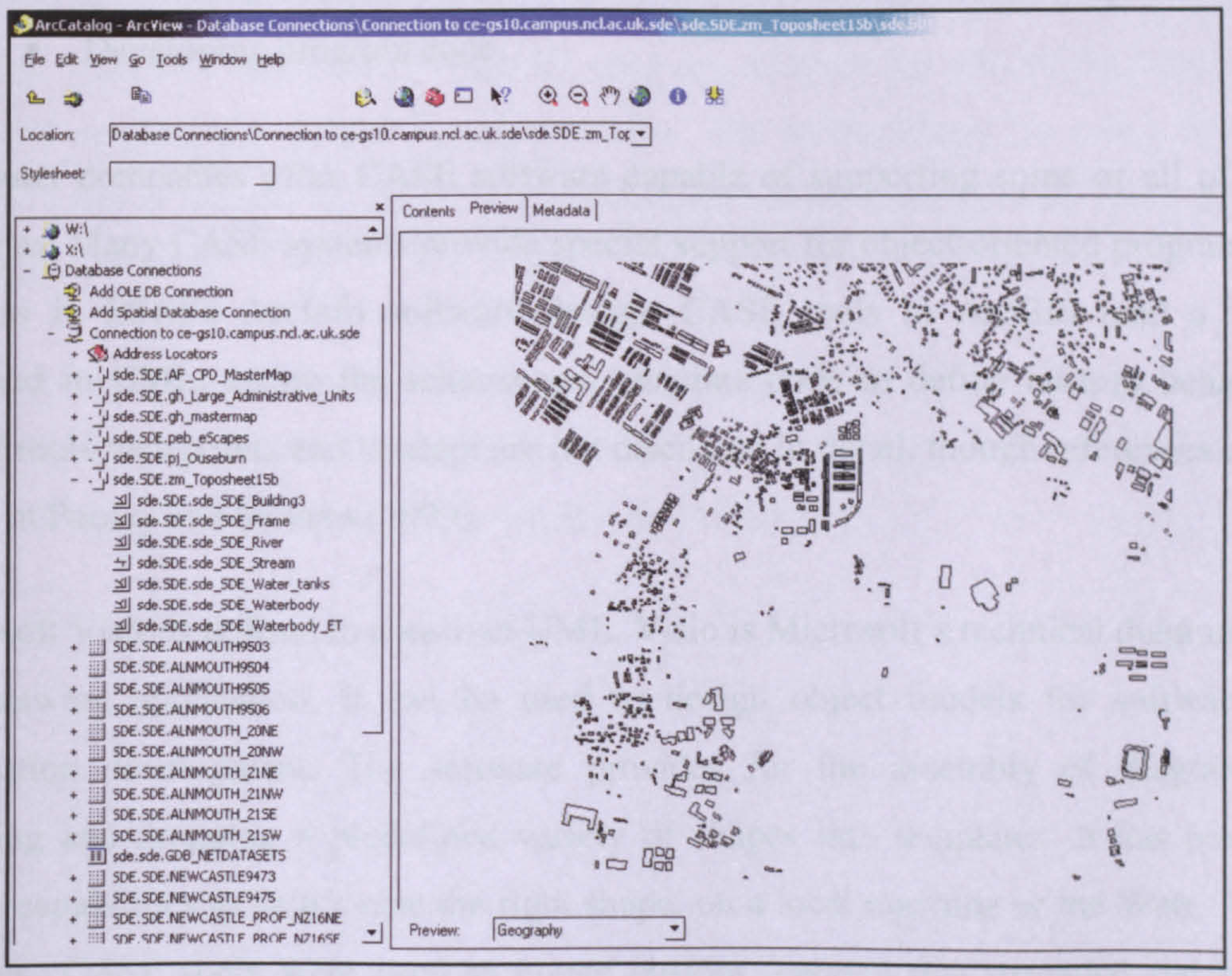


Figure 5.3: ArcCatalog manages data in ArcSDE Server used for the research

ArcGIS's desktop modules use object-oriented programming supported by CASE tools. CASE tools are also used to design a prototype database involving UML which defines schema of database. CASE tools are described in the next section as a basis to design the geodatabase.

5.4 What are CASE Tools?

A CASE tool is a set of integrated programs and application development tools that help in software development (Kang et al., 2004). CASE programs are used to understand, simplify, and automate the development methods employed throughout the software life cycle, to eliminate data redundancy or conflict, improve productivity and reliability, and to capture reusable functionality in terms of design and code. These can include tools for:

- Summarising initial requirements.
- Developing flow diagrams.
- Scheduling development tasks.
- Preparing documentation.
- Controlling versions.
- Developing program code.

Computer companies offer CASE software capable of supporting some or all of these activities. Many CASE systems provide special support for object-oriented programming such as in ESRI's ArcInfo software design. CASE tools in ArcGIS read a model designed in UML, create the schema and generate code to define custom behaviour. CASE tools viewpoints and concept are not discussed in detail, though references can be found in Pooley and Stevens (1999).

Microsoft Visio is utilised to construct UML. Visio is Microsoft's technical diagramming and drawing application. It can be used to design object models for software and application development. The software provides for the assembly of diagrams by dragging and dropping a predefined variety of shapes into templates. It has powerful search capabilities to help locate the right shape, on a local machine or the Web. In this research, CASE tools were used to create custom features that extended the object-oriented data model of ArcInfo.

The following section describes the use of UML and CASE tools to implement a DBMS for the project.

5.5 Implementation of a DBMS

The implementation phase for spatial database involves the use of ESRI's ArcInfo technology. The data model design of ArcInfo uses a CASE tool subsystem that supports UML. The creation and design of a new database intended in this research is basically a geodatabase as discussed in the previous chapter. A class diagram in UML provides clear representation of database elements such as feature datasets, feature classes, object classes, TIN datasets, raster datasets and the relationship among them. A CASE tool subsystem assists the design of applications and generation of application code. It also allows the creation of blueprints of the structure of the database using UML.

The approach for utilising CASE tools to design and create databases involves the use of UML to define all schemas for the database, generating that schema then populating the schema with data. ArcInfo object model makes use of Visio, which supports the UML, to automate the process of creating the logical model, generating code and creating the database schema. A program in ArcGIS provides wizards for generating initialisation code and creating the spatial database schema. The methodology for using Visio and ArcInfo tools to design and create the geodatabase can be accomplished using the steps below (Figure 5.4):

1. Creating the logical object model design in UML.
2. Exporting the UML model to Microsoft Repository.
3. Creating the geodatabase schema from XMI file.

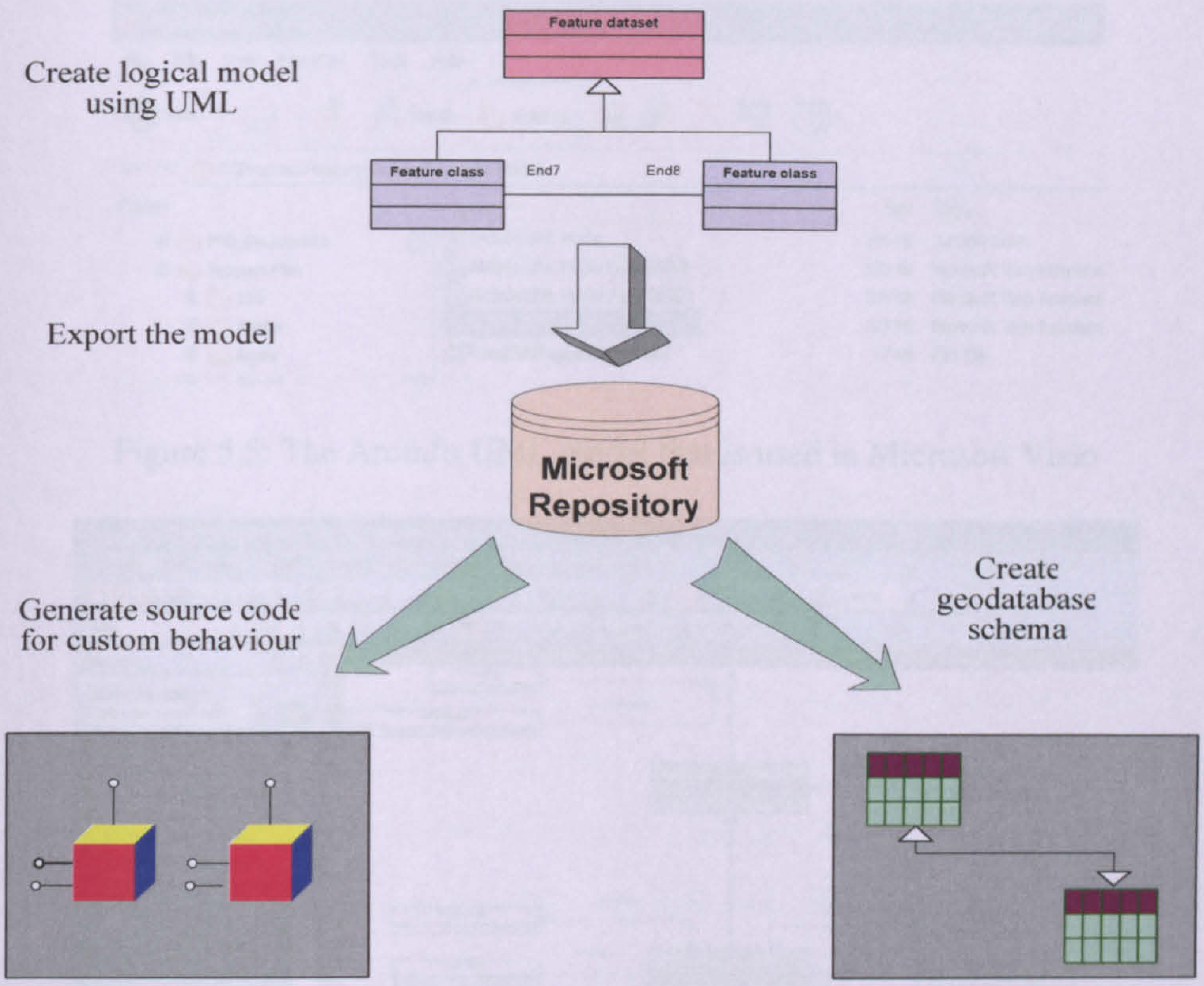


Figure 5.4: Strategy for using UML and CASE tools to design a database model

The following subsections describe the steps in designing the database. A logical model is constructed using UML. Exporting the UML model and creating geodatabase schema are carried out in the subsequent sections.

5.5.1 Creating the Logical Model

The logical model is created using the basic database model in UML, namely the ArcInfo UML Model that is provided in ArcGIS software. The ArcInfo UML Model contains the database model and other relevant parts of the ArcObjects Library needed for the production of custom models. The diagrams in the model are Visio drawing templates. Since Visio 2003 is used to design the model, the ArcInfo UML Model template for Version 2003 is loaded and displayed in Visio. The template is located under ArcGIS software in the Case Tools/UML Models directory (Figure 5.5). Once open, the ArcInfo

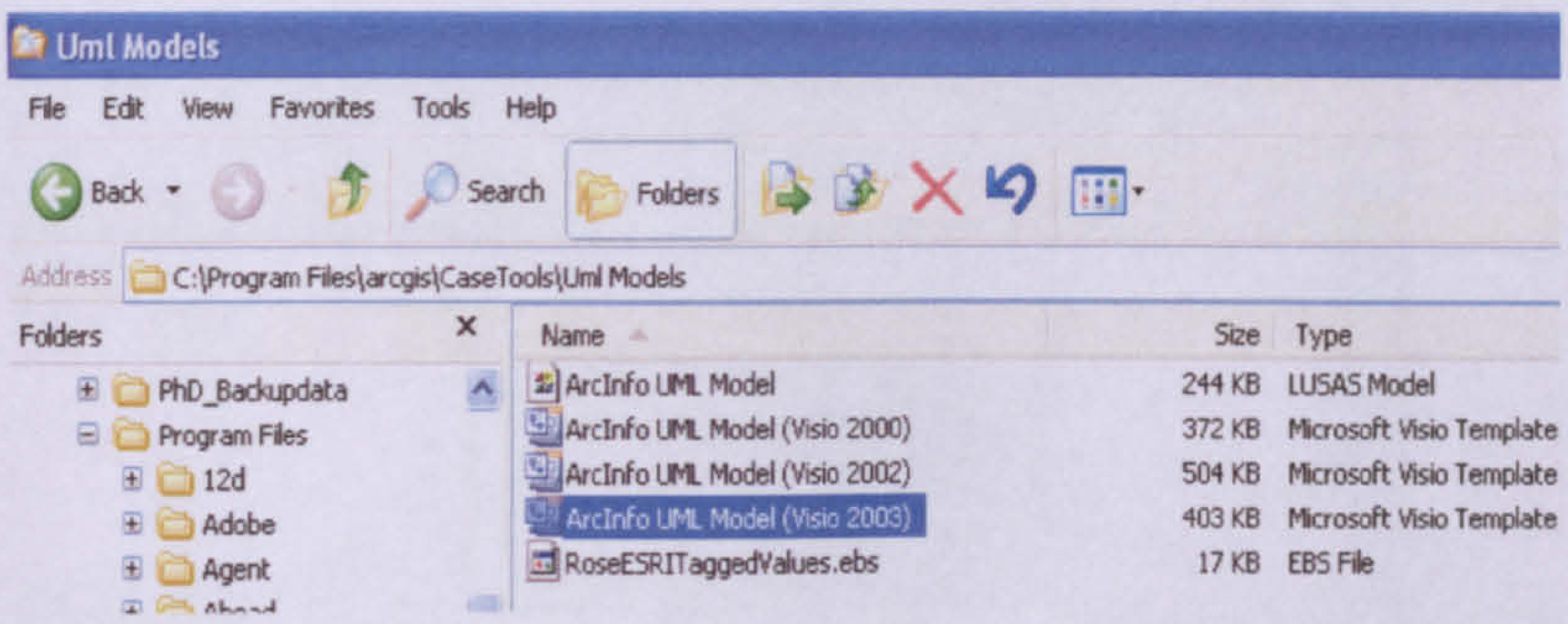


Figure 5.5: The ArcInfo UML model that is used in Microsoft Visio

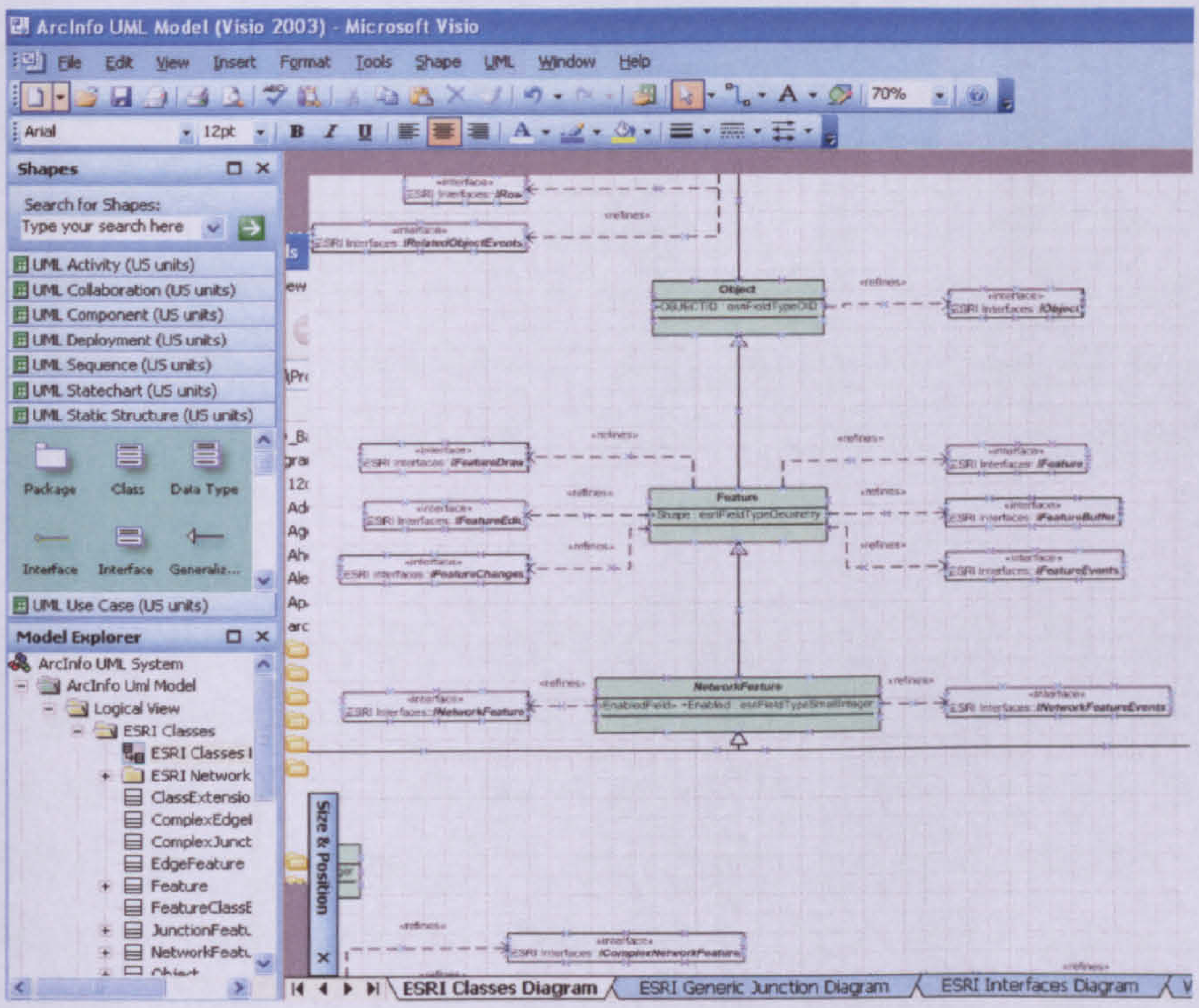


Figure 5.6: ArcInfo UML Model's ESRI Classes Diagram in Visio

UML model standard diagrams are displayed within the Shapes and Model Explorer of the Visio program view as in Figure 5.6.

The ArcInfo UML Model diagram holds the object model required for using UML to model the database. The object model has four packages (See Figure 5.7):

- Logical View.
- ESRI Classes.
- ESRI Interfaces.
- Workspace.

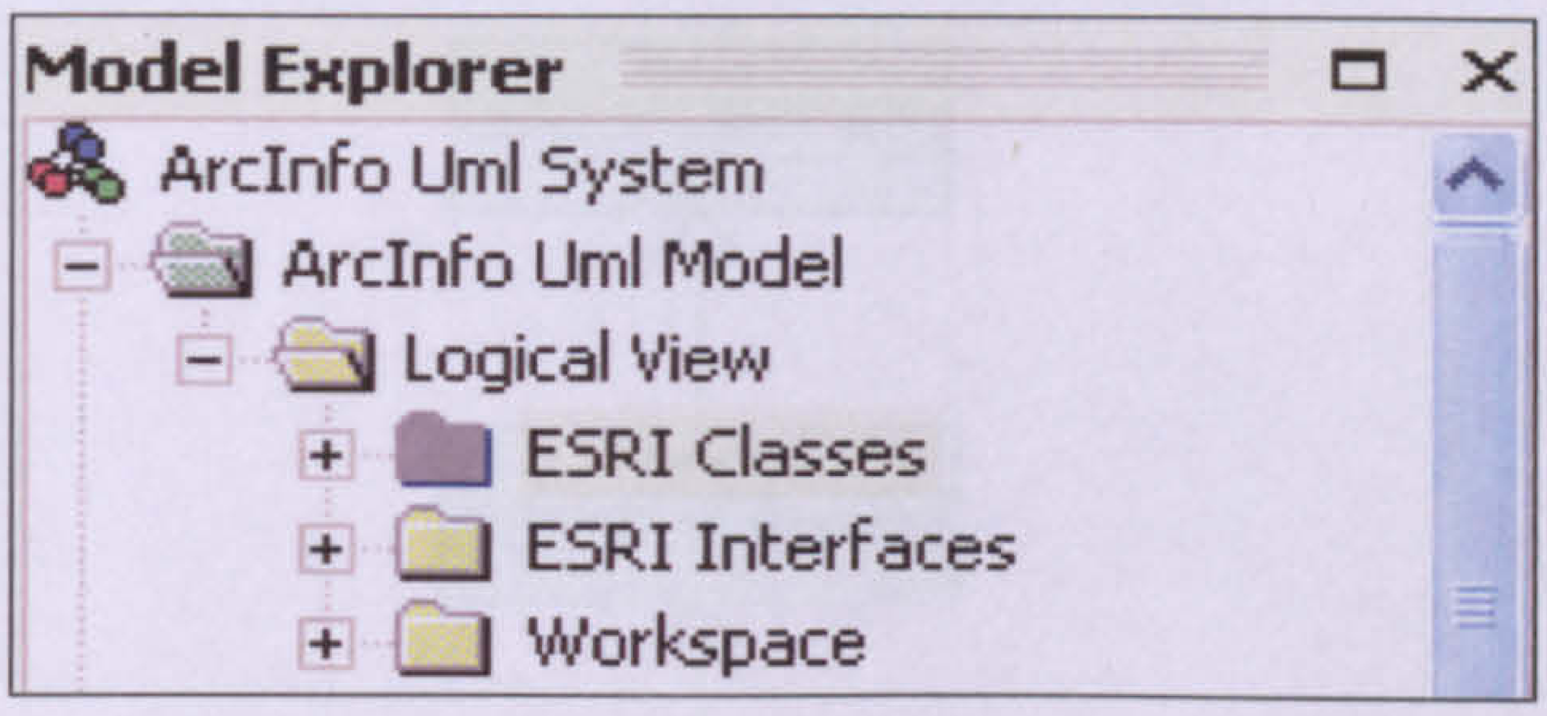


Figure 5.7: The four packages within UML Model Explorer in Visio

These UML packages perform as directories where different components of the entire logical model are maintained. The Logical View package is the root level and contains the other three packages. A package contains any number of UML elements such as interfaces, classes, other packages and diagrams. This package represents the database. Database designers and developers use this package to create objects and database designs. A complex database requires that more packages be created. There is no limit to the number of packages an object may contain but all must be created under the Workspace package.

The ESRI Classes package contains existing ESRI COM classes' diagrams that can be reused by either aggregation or containment. Figure 5.8 shows the ESRI Classes diagrams in the ArcInfo UML Model template. The UML classes in this diagram represent COM classes that belong to the geodata access components of ArcInfo necessary to create an object model. ESRI Classes in this package are used to access spatial data sources, including the geodatabase. Feature classes and object classes in the object model inherit from these classes. As mentioned in Chapter 3, classes inherit from other classes. As in Figure 5.8, *Feature* inherits from *Object*, meaning Feature 'is a kind of' *Object*. *NetworkFeature* implements all the interfaces Feature does, including those implemented by *Object*.

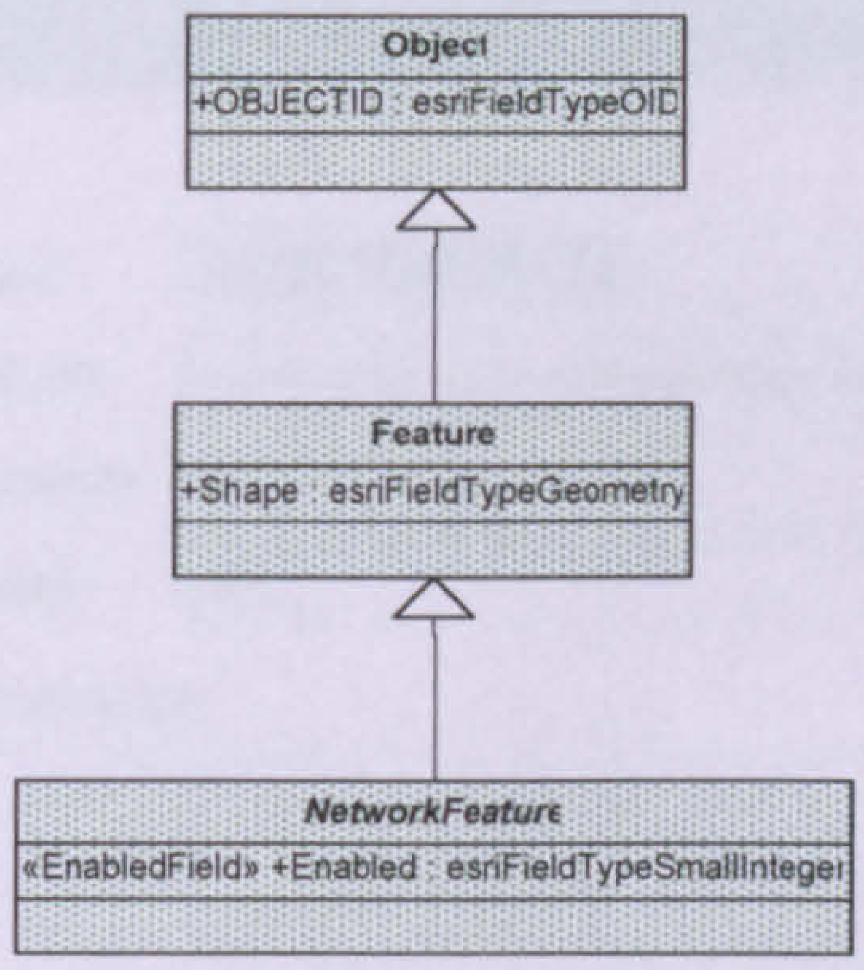


Figure 5.8: Inheritance of classes

The ESRI Interfaces package contains the definition of the interfaces implemented by the components shown in the ESRI Classes package. The UML classes in the ESRI Classes package thus represent COM classes which provide services through interfaces in the ESRI *Interfaces* package (Figure 5.9). For example, *Feature* implements the interface *IFeatureBuffer*. This means that in ArcGIS software, ArcMap can be asked to implement a buffer operation on a feature using the attribute *Shape*. The feature is displayed when asked in ArcMap using the interface *IGeometry*. A UML refinement is used to express the relationship between the class and the implemented interface.

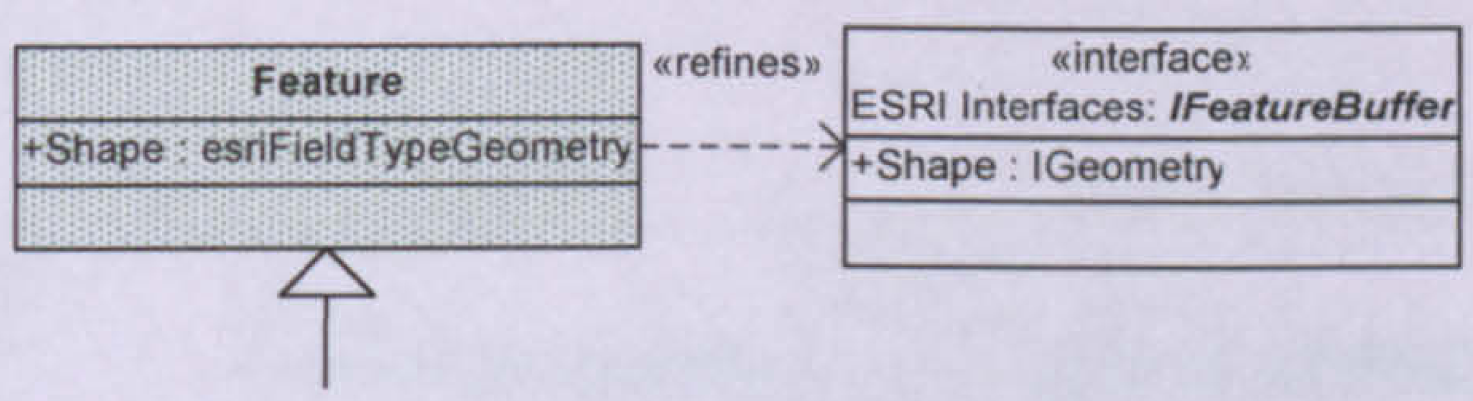


Figure 5.9: Sample of a COM class implements interfaces

The creating of the logical model of the database for the research is described in the following steps. Using the Visio template tools, UML Classes are dragged from the Shapes menu to the Class diagram window and double clicking the class object enables properties to be specified as in Figure 5.10. An object UML class called *GeospatialData_capture* is named and classified.

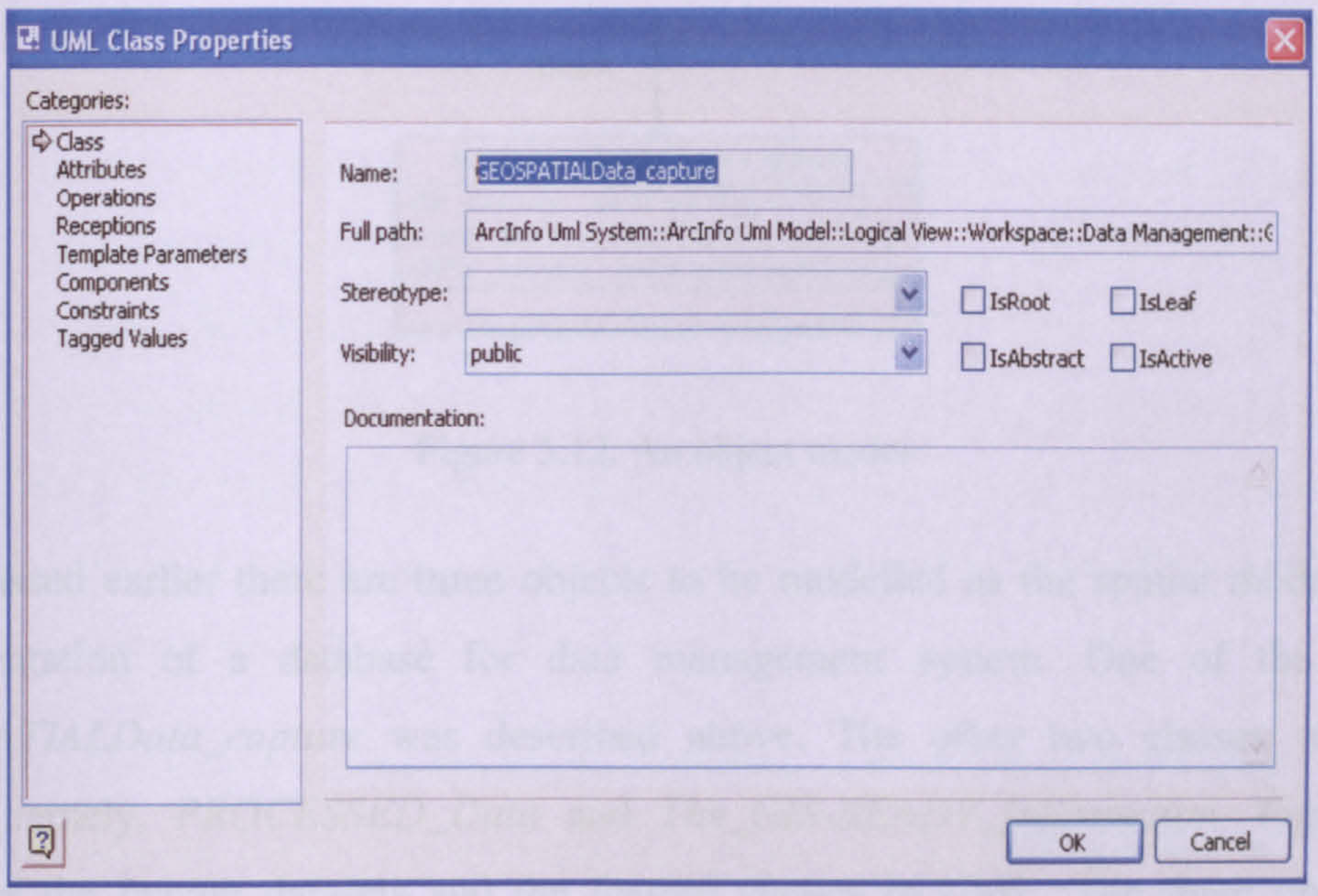


Figure 5.10: UML Class Properties window to specify object model

The attributes of the feature datasets *GEOSPATIALData_capture* are given as *Air_survey, Field_survey* and *GPS_survey,* features that are actually feature class sub-objects. The *Type* is chosen for the field entry (Figure 5.11).

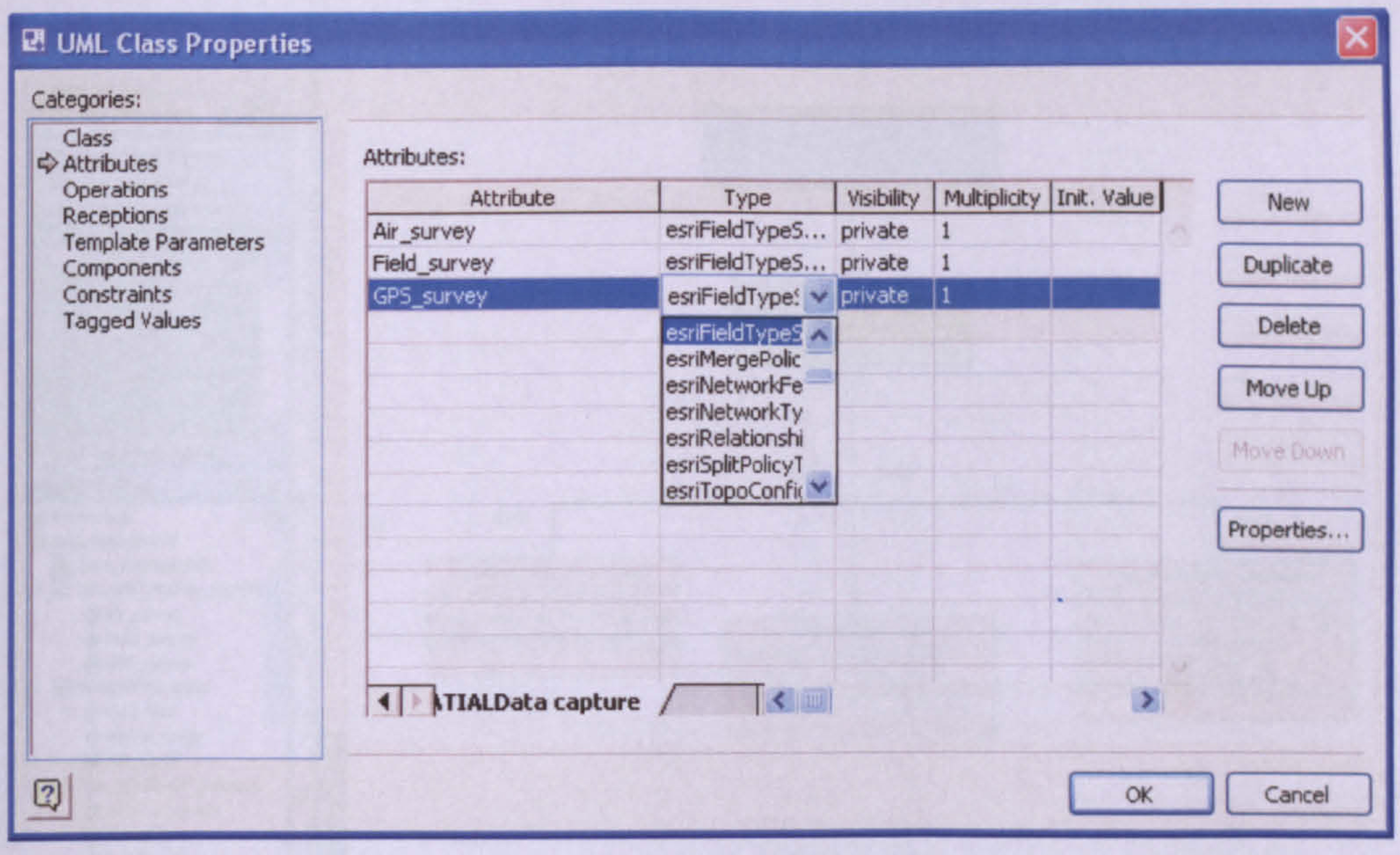


Figure 5.11: Attributes for the feature dataset

When the class object has been specified the object model is displayed as in Figure 5.12 after clicking the OK tab.

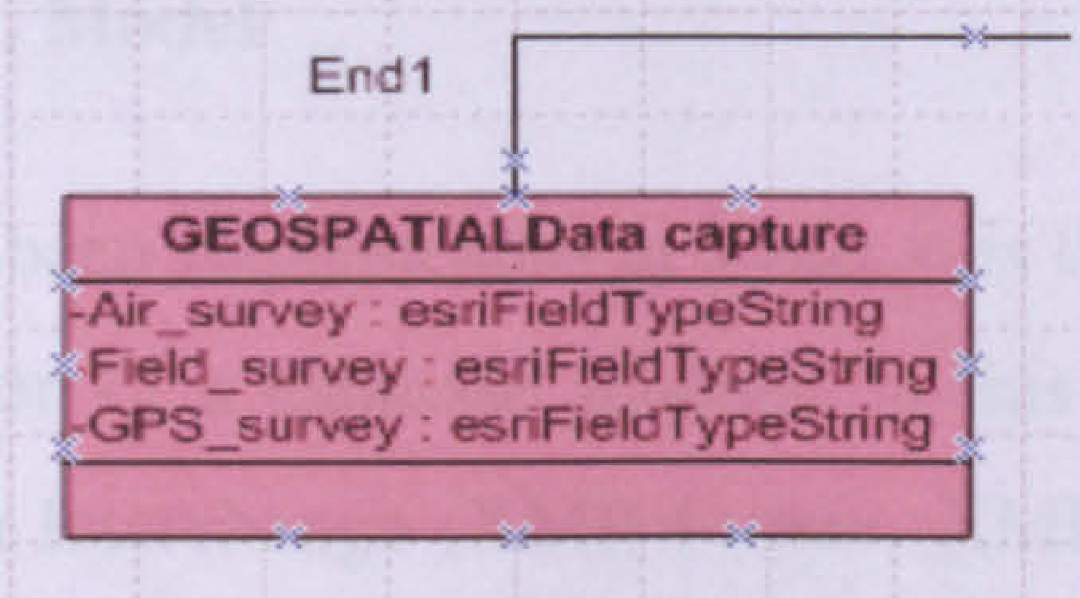


Figure 5.12: An object model

As discussed earlier there are three objects to be modelled as the spatial model for the implementation of a database for data management system. One of the objects, *GEOSPATIALData_capture* was described above. The other two classes were also created namely, *PROCESSED_Data* and *The_GIS-READY_Information*. Figure 5.13 illustrates the feature datasets and the feature classes in UML. The three objects are representatives of many more objects of similar types. Basically, from raw capture data to the point of GIS-ready information production there are many more objects that can be modelled as feature classes under the three feature datasets. Once the logical model has been set up and stored in Visio, the next task is to export the model to a file that can be used to create a physical database schema.

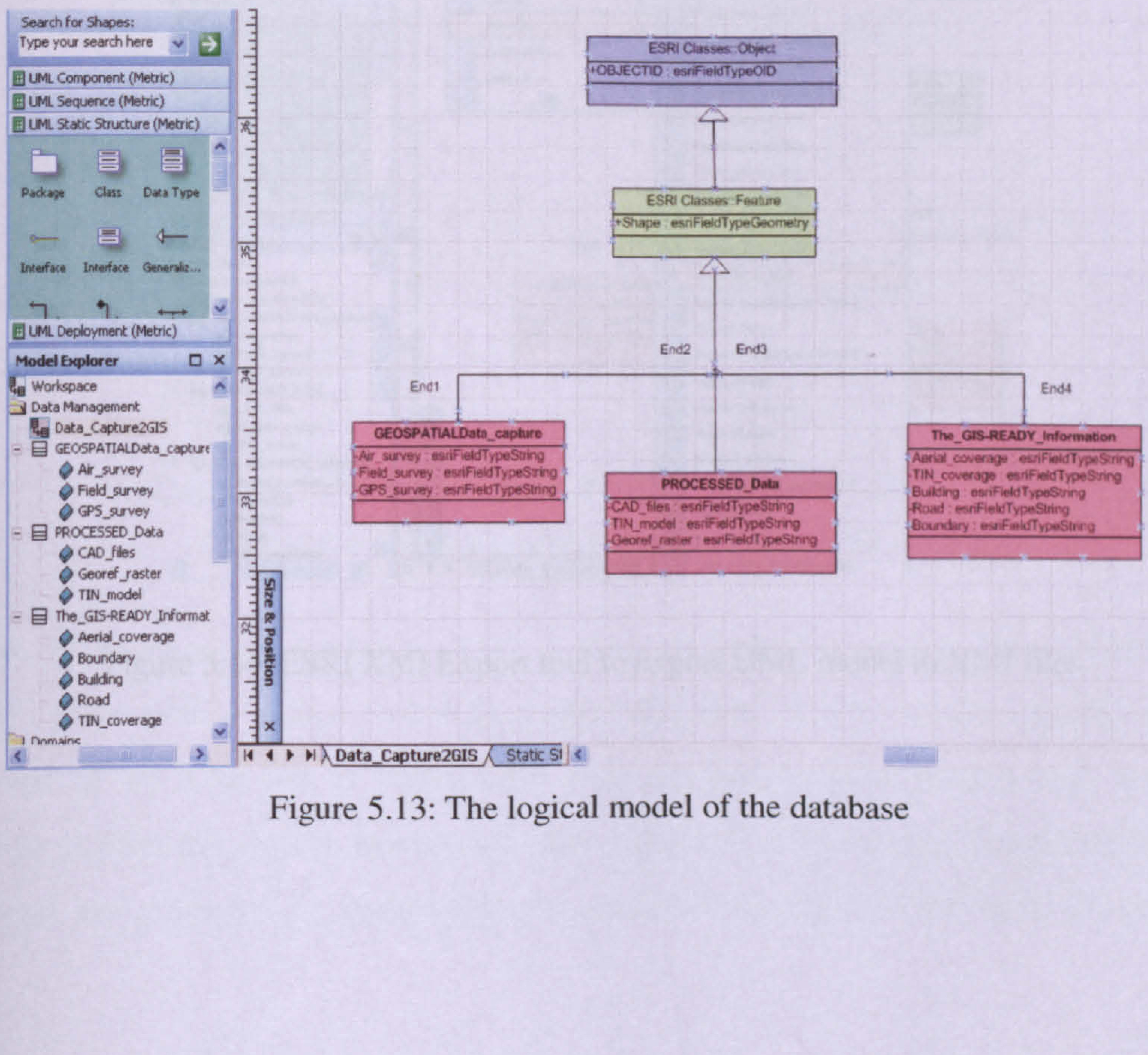


Figure 5.13: The logical model of the database

5.5.2 Exporting the UML Model

Once the logical model has been implemented in Visio, it is then converted into a format that can allow the generation of a database schema. In this case the UML model was exported in XML Metadata Interchange (XMI) format. XMI is an Object Management Group (OMG) standard that specifies how to store an UML model in an XML file (OMG, 2004). The latest GIS software with the upgraded CASE tools provides support for the generation of schema and code from models stored in XMI. The tools to export the UML model to XMI are contained within Visio.

ESRI XMI Export option from the Tools>Add-Ons is used to export the UML model to XMI format (Figure 5.14). A dialog box was shown to specify the directory and folder where the XMI file to be saved. Figure 5.15 shows the process of the exporting to database schema. Figure 5.16 indicates the exporting has been successful.

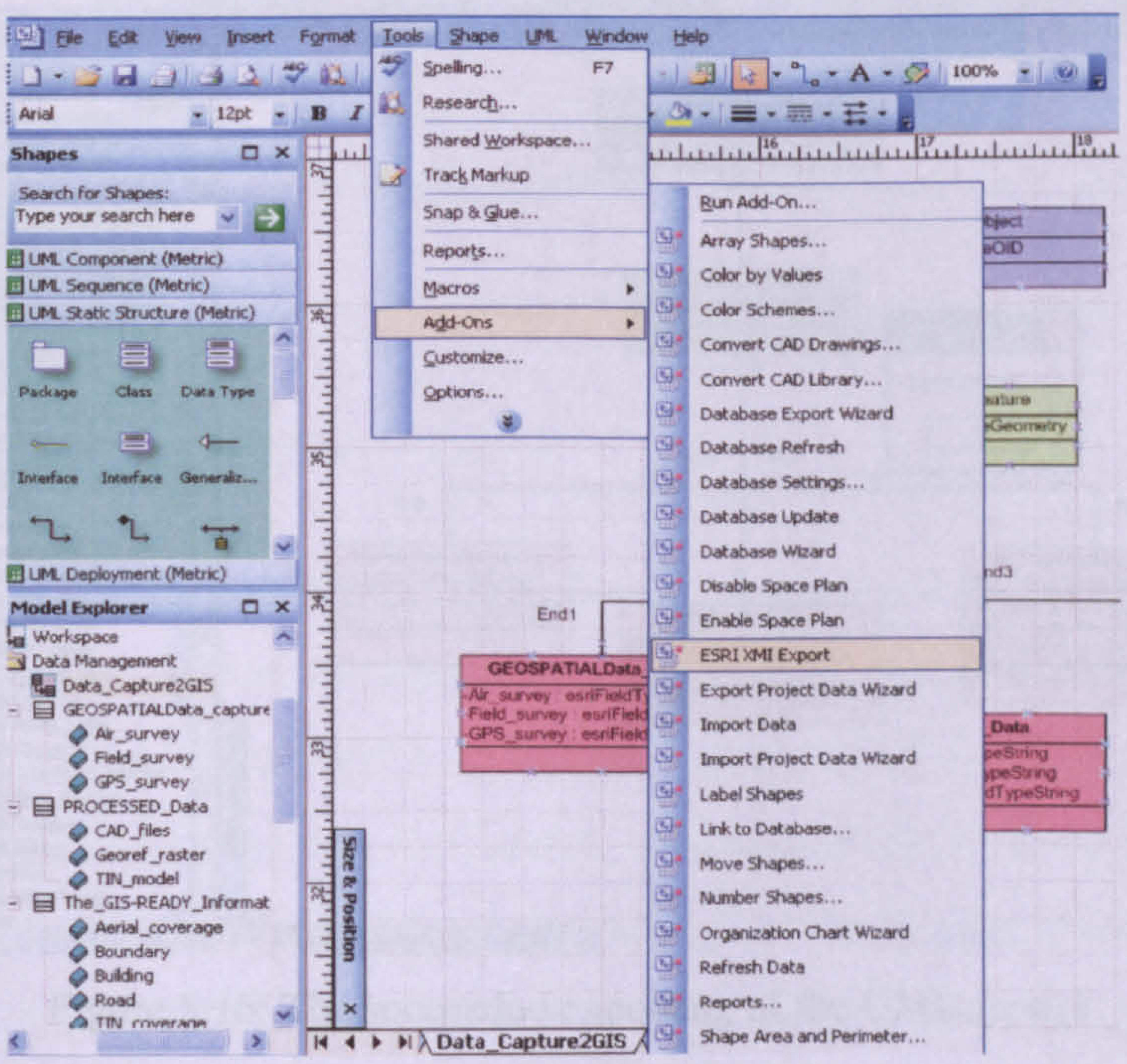


Figure 5.14: ESRI XMI Export tool to export UML model to XMI files

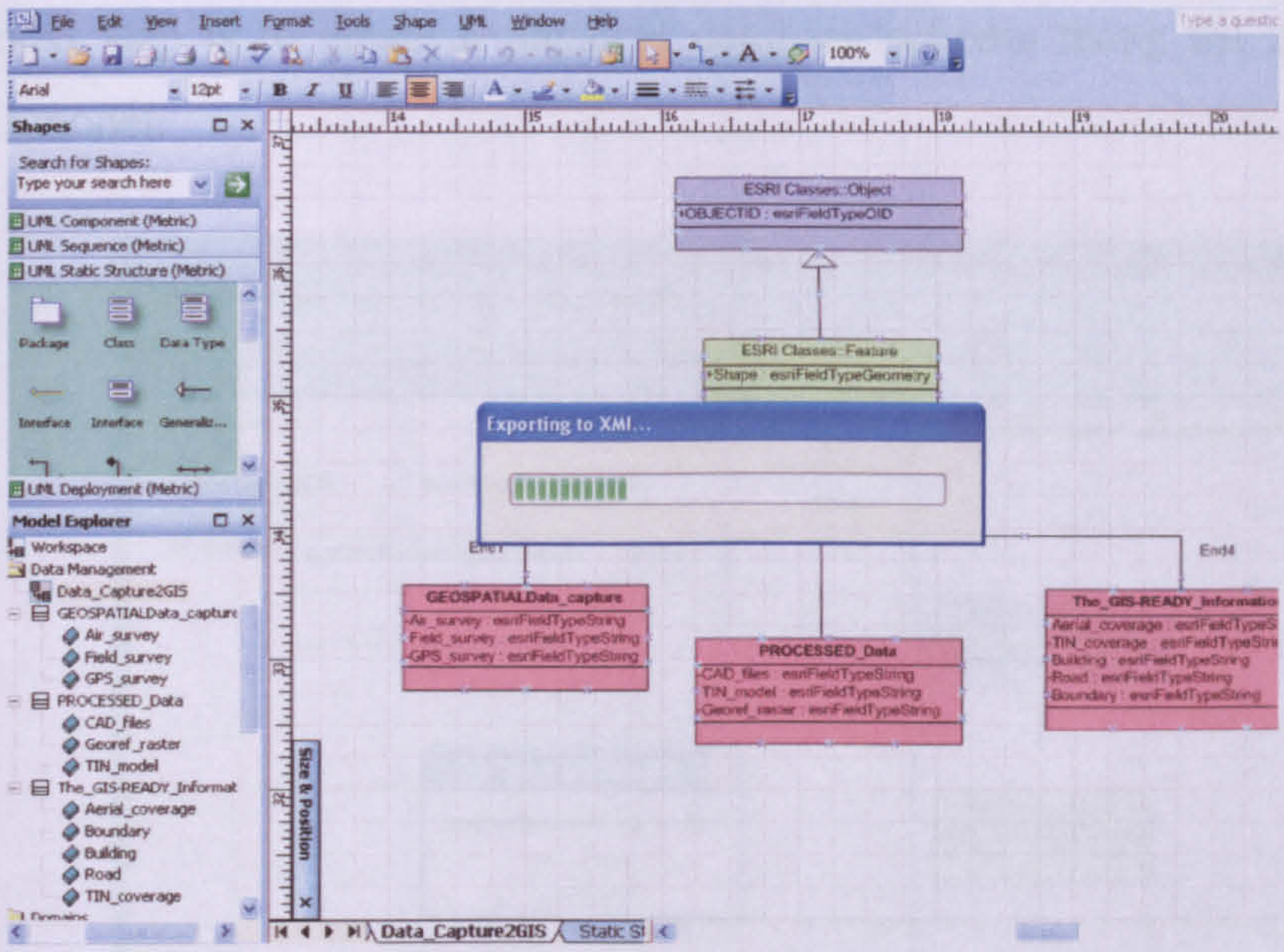


Figure 5.15: The exporting of UML as XMI file

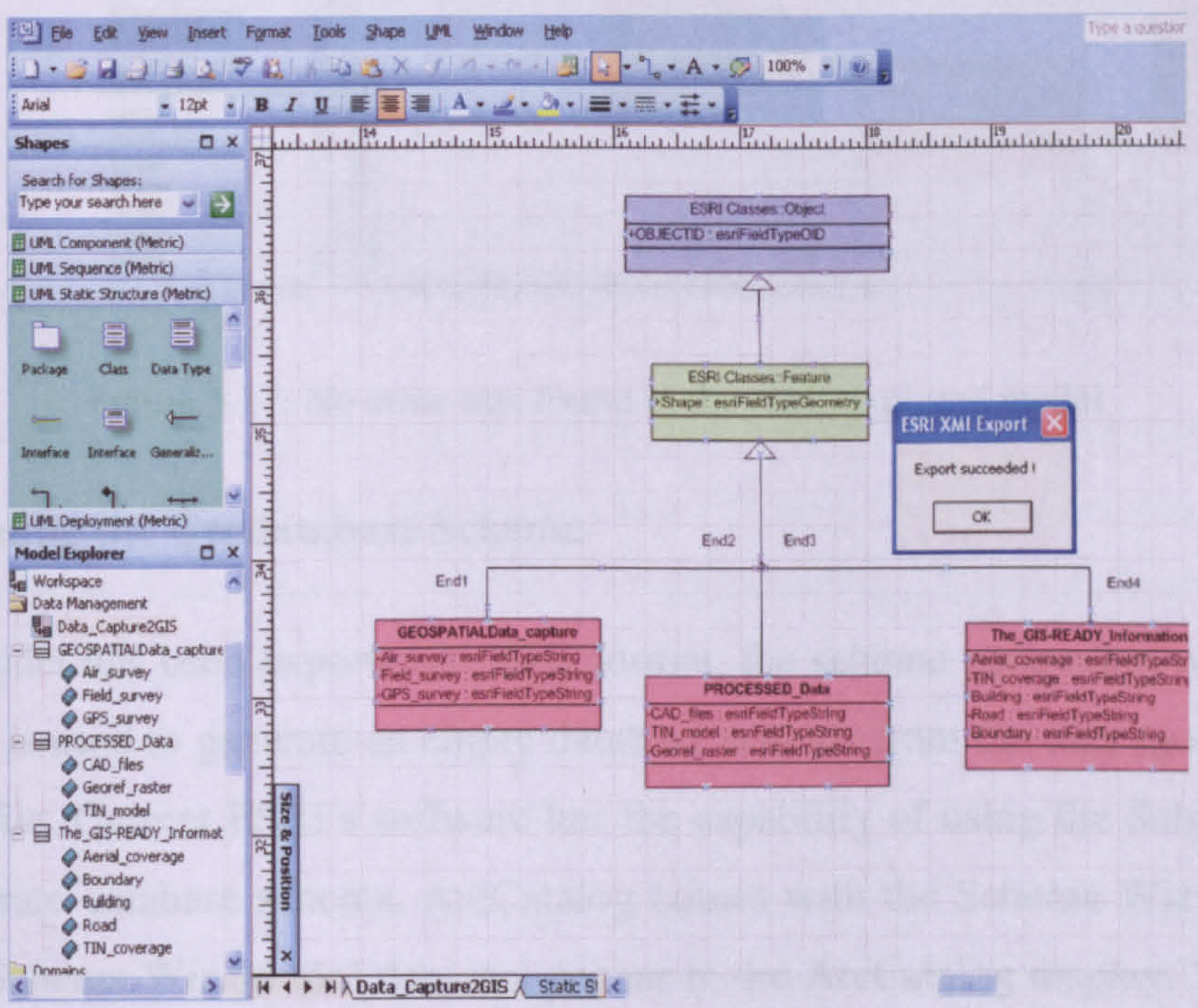


Figure 5.16: The succeeded exporting of the UML model

The resultant exported model in XMI is then checked for errors. The ArcInfo UML model comes with a macro called Semantics Checker used to verify whether a model enclosed in an XMI file is compatible with the geodatabase model. The Semantics Checker is then run on the exported model. Figure 5.17 illustrates that the XMI model designed and exported is compatible for geodatabase type model with no particular errors. The next step is to create a physical database schema using the exported and checked XMI model.

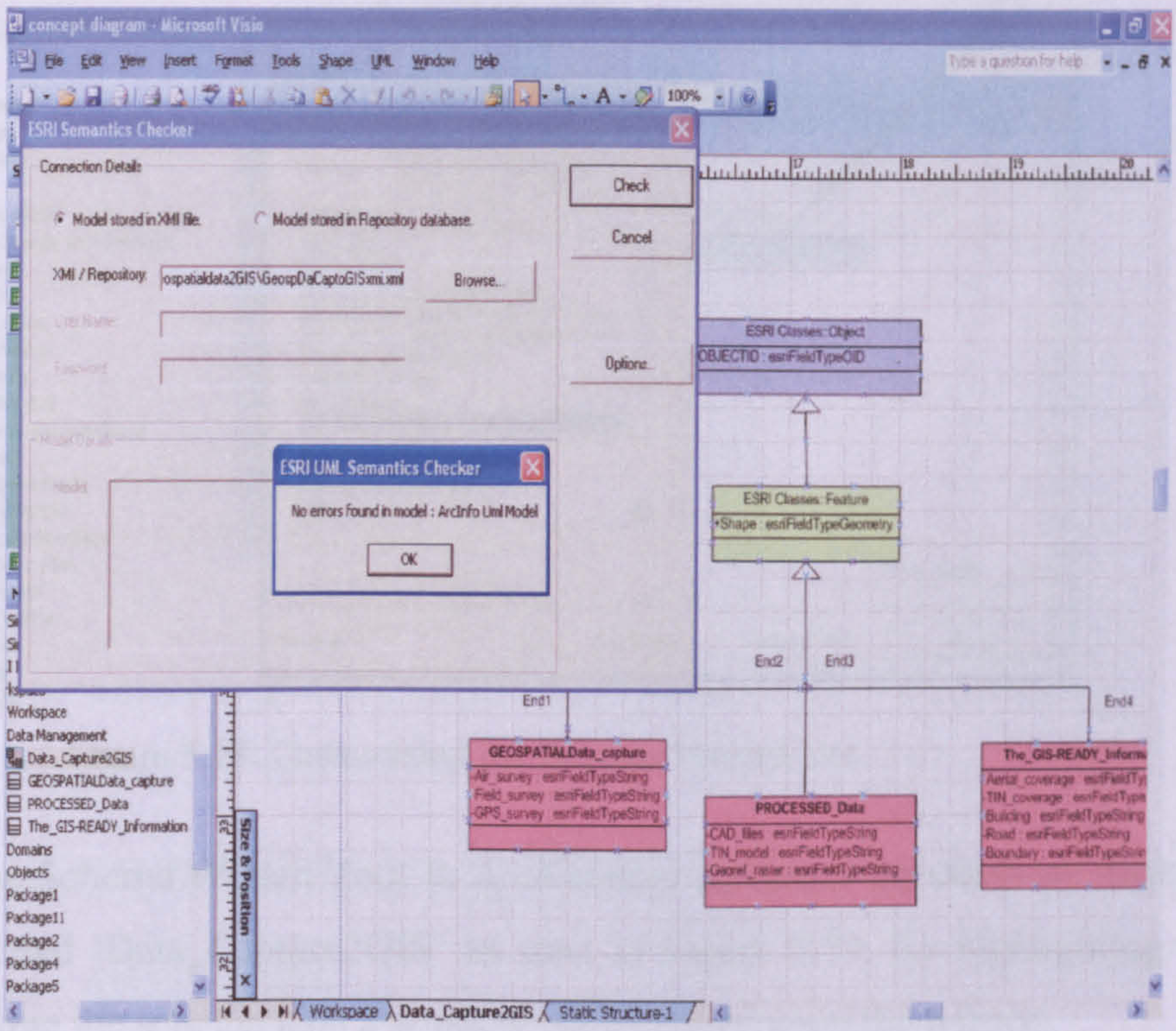
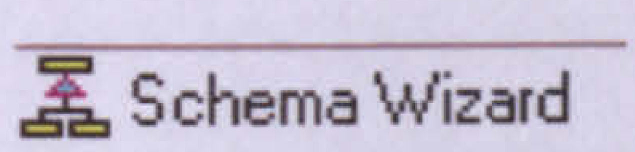


Figure 5.17: No error was found in the XMI exported model

5.5.3 Creating the Geodatabase Schema

Once the model has been exported to XMI format, the schema generation wizard of the CASE tools is used to generate an empty database schema from the data model template of an XMI file. Current ESRI's software has the capability of using the Schema Wizard tool to generate database schema. ArcCatalog comes with the Schema Wizard tool. By default the Schema Wizard tool does not appear in the ArcCatalog display. So, by using the Tools menu, the ArcCatalog toolbar was customised by displaying new icon Schema Wizard as the diagram below, in the toolbar menu.



This is done by dragging the icon in the Customise dialog box of ArcCatalog. The icon is displayed in the toolbar at the top (Figure 5.18).

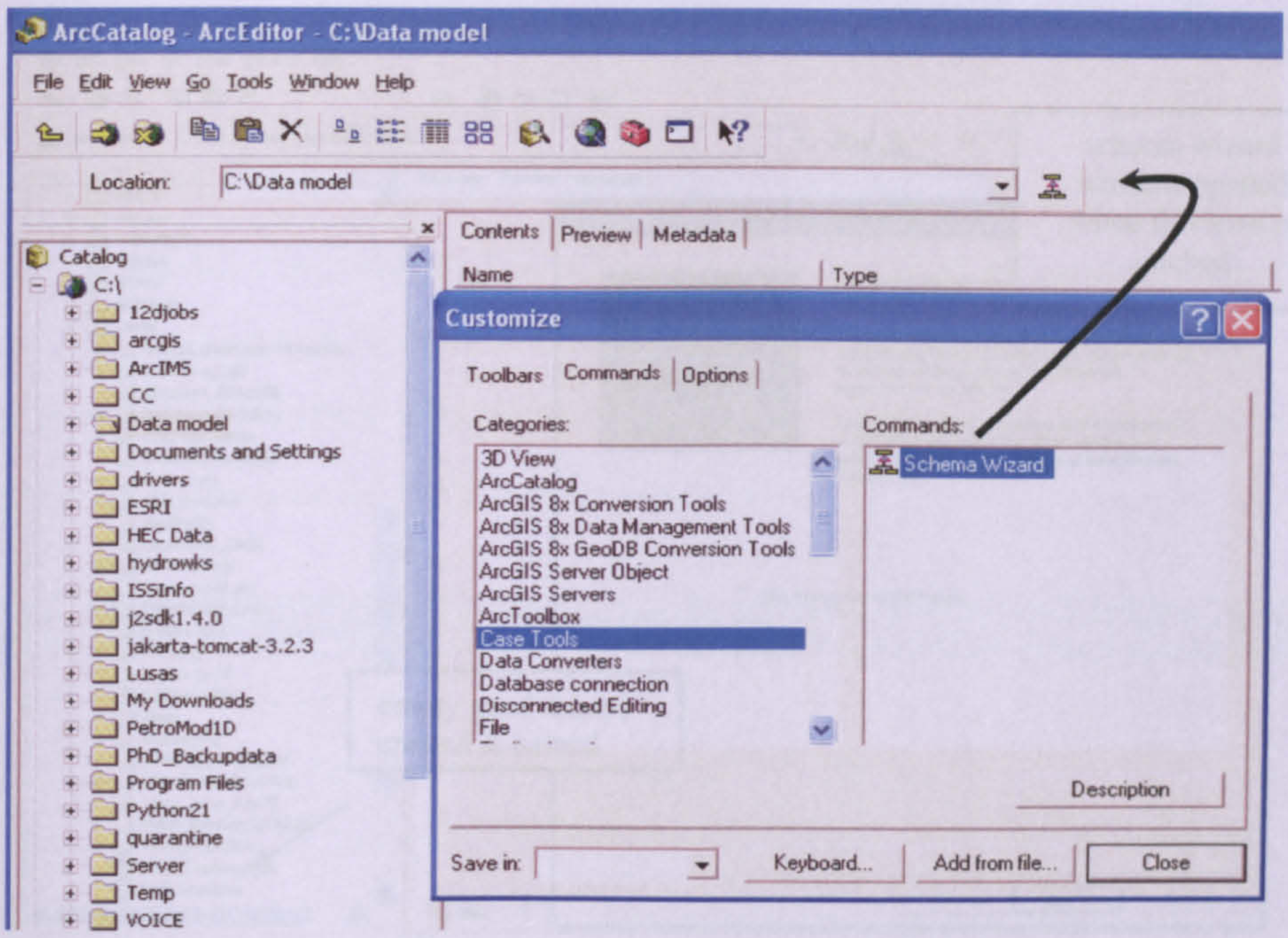


Figure 5.18: Customising the Schema Wizard tool

Before running the Schema Wizard Tool in ArcCatalog, an empty geodatabase must be created with a named 'Data_Capture2GIS' as seen in Figure 5.19. By highlighting the database and clicking the schema icon a window opens up showing the wizard.

A dialog box as in Figure 5.20 is used to specify the XMI file that was converted from UML and saved in a directory. The schema wizard reads the XMI file and translates it into a geodatabase feature dataset and feature classes which are automatically created in the empty geodatabase formed and named earlier.

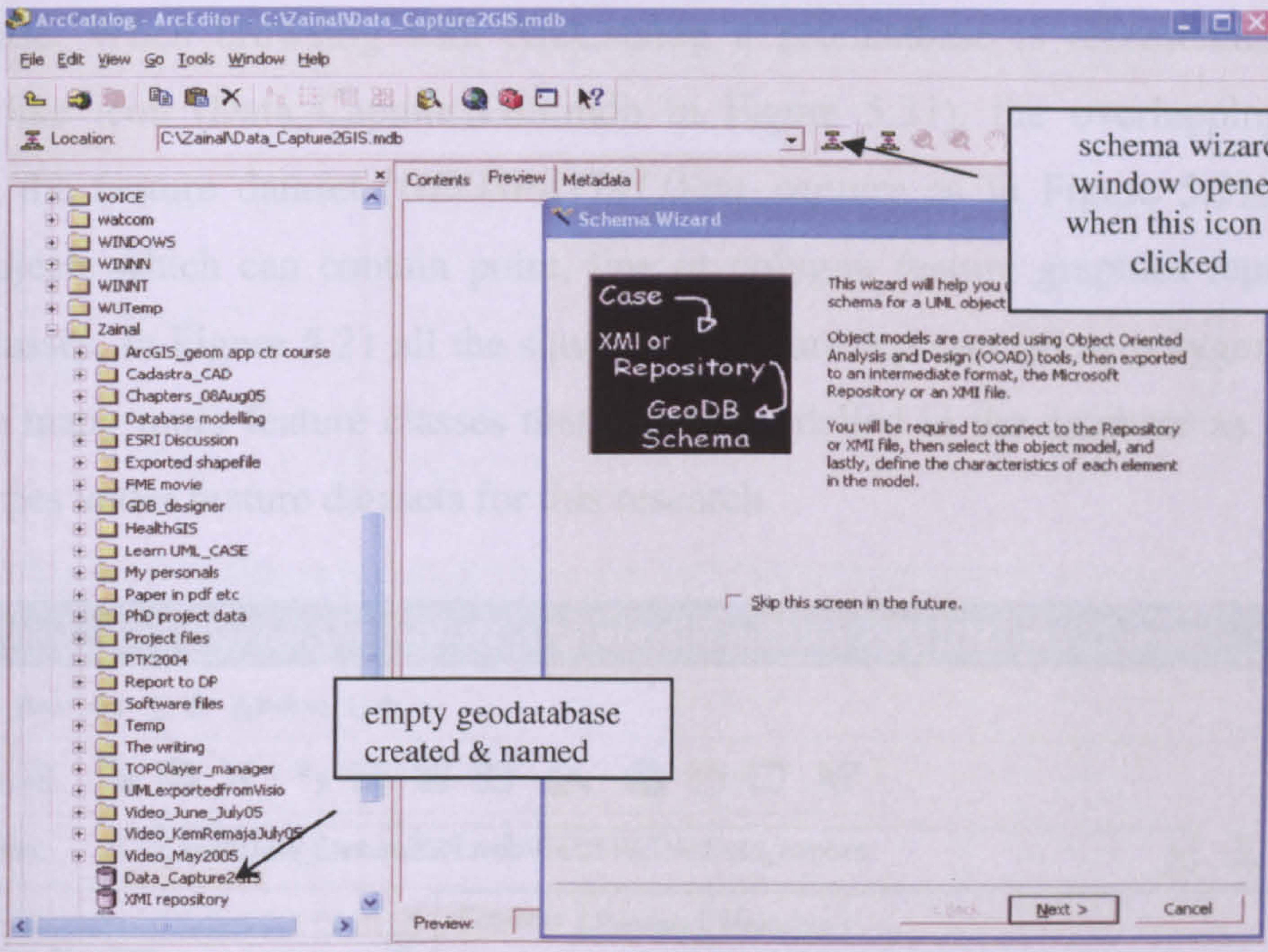


Figure 5.19: Creating geodatabase schema from UML model

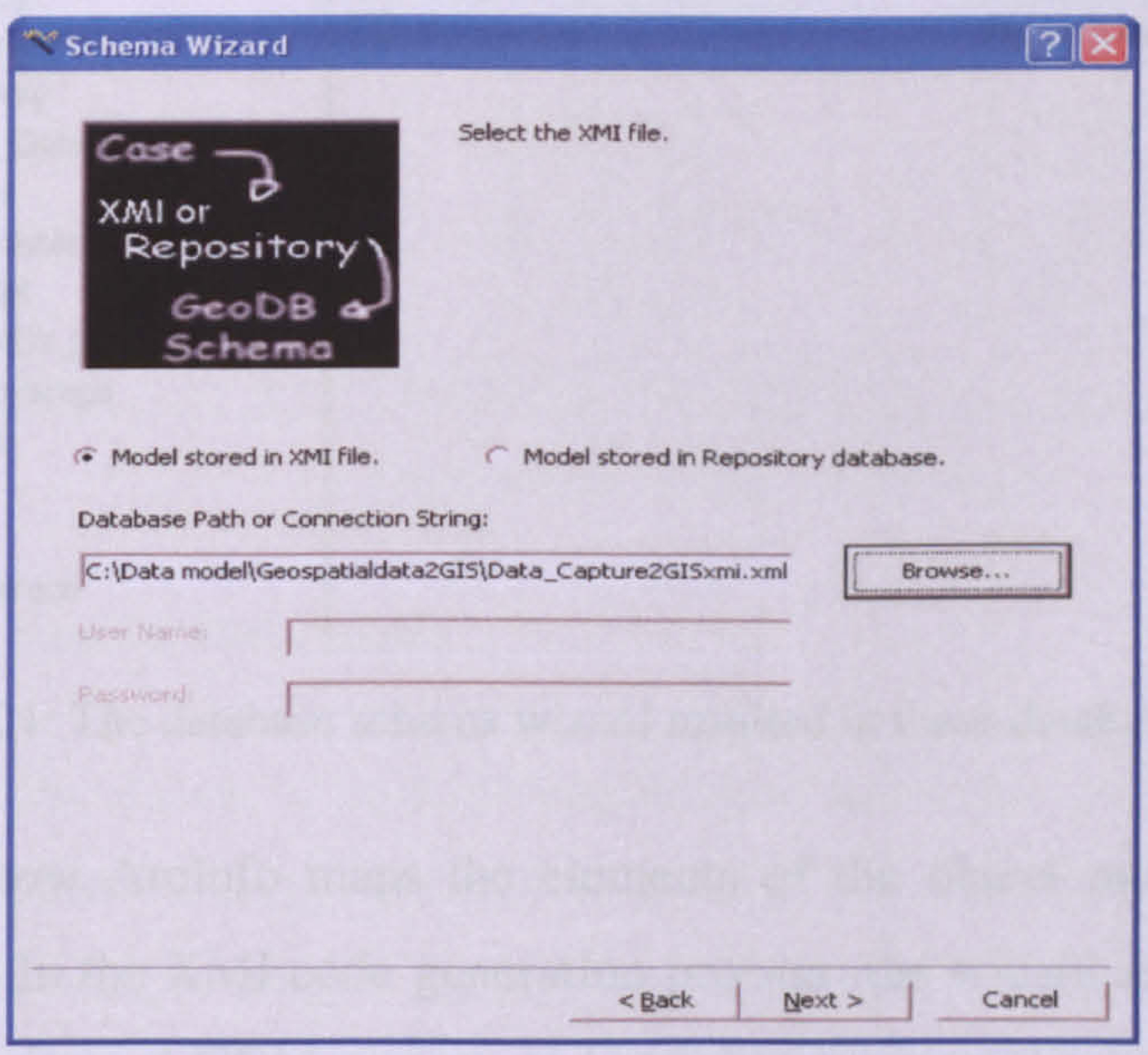


Figure 5.20: Specifying the XMI files path or connection string

The schema wizard of the CASE tools has been used to generate an empty geodatabase, after which data can be loaded into that schema. Thus, a new object-oriented data model, geodatabase model has been created using ArcInfo 9 (Figure 5.21). This geodatabase model provides a standardised framework into which various types of data can be stored. Once created, the geodatabase is a Microsoft Access file called an ArcGIS personal geodatabase. When browsing with ArcCatalog a geodatabase is represented by as a cylinder like icon (Data_Capture2GIS.mdb in Figure 5.21), the overlapping squares represent the feature dataset (*GEOSPATIALData_capture* as in Figure 5.21), and the square objects which can contain point, line or polygon feature graphics represent the feature classes. In Figure 5.21 all the square objects are represented as polygon features. There are many more feature classes that can be modelled in the database as objects of similar types under feature datasets for this research.

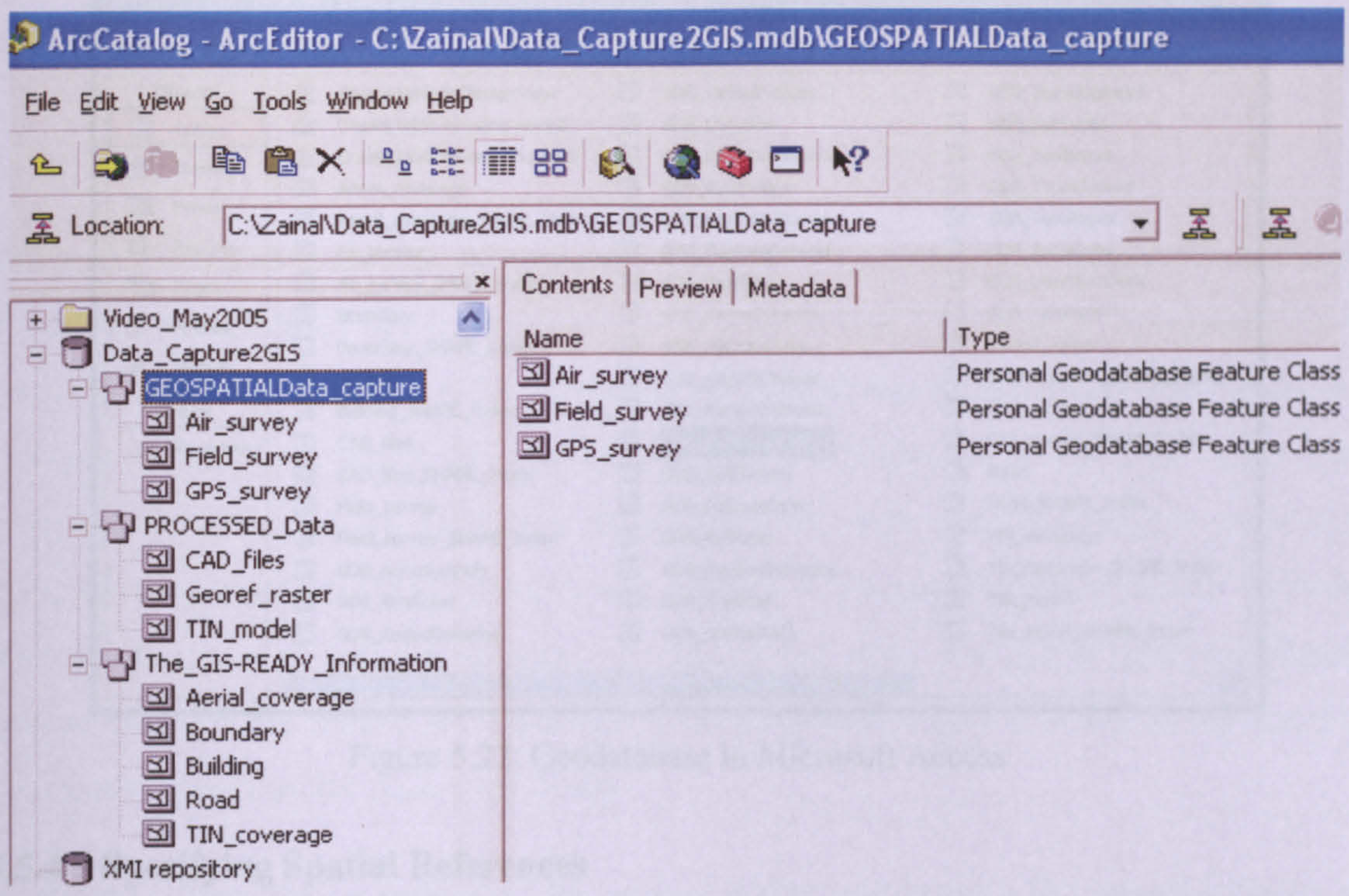


Figure 5.21: The database schema wizard resulted in these database objects

Table 5.1 shows how ArcInfo maps the elements of the object model into relational database elements. In the XMI code generation process, the wizard searched the system registry for any registered COM classes and attaches them to the geodatabase features classes. As the geodatabase holds a Microsoft Access file format, it can be opened to view the features created (Figure 5.22).

The schema generation wizard has generated a table for the feature classes that can be manipulated and managed using ArcCatalog. Other tables and feature classes can be added in ArcCatalog which may hold the same properties as the features created in UML.

Object model components	**Relational database components**
Object	Row
Attribute	Field, Column
Class	Table

Table 5.1: Mapping of object model into relational database

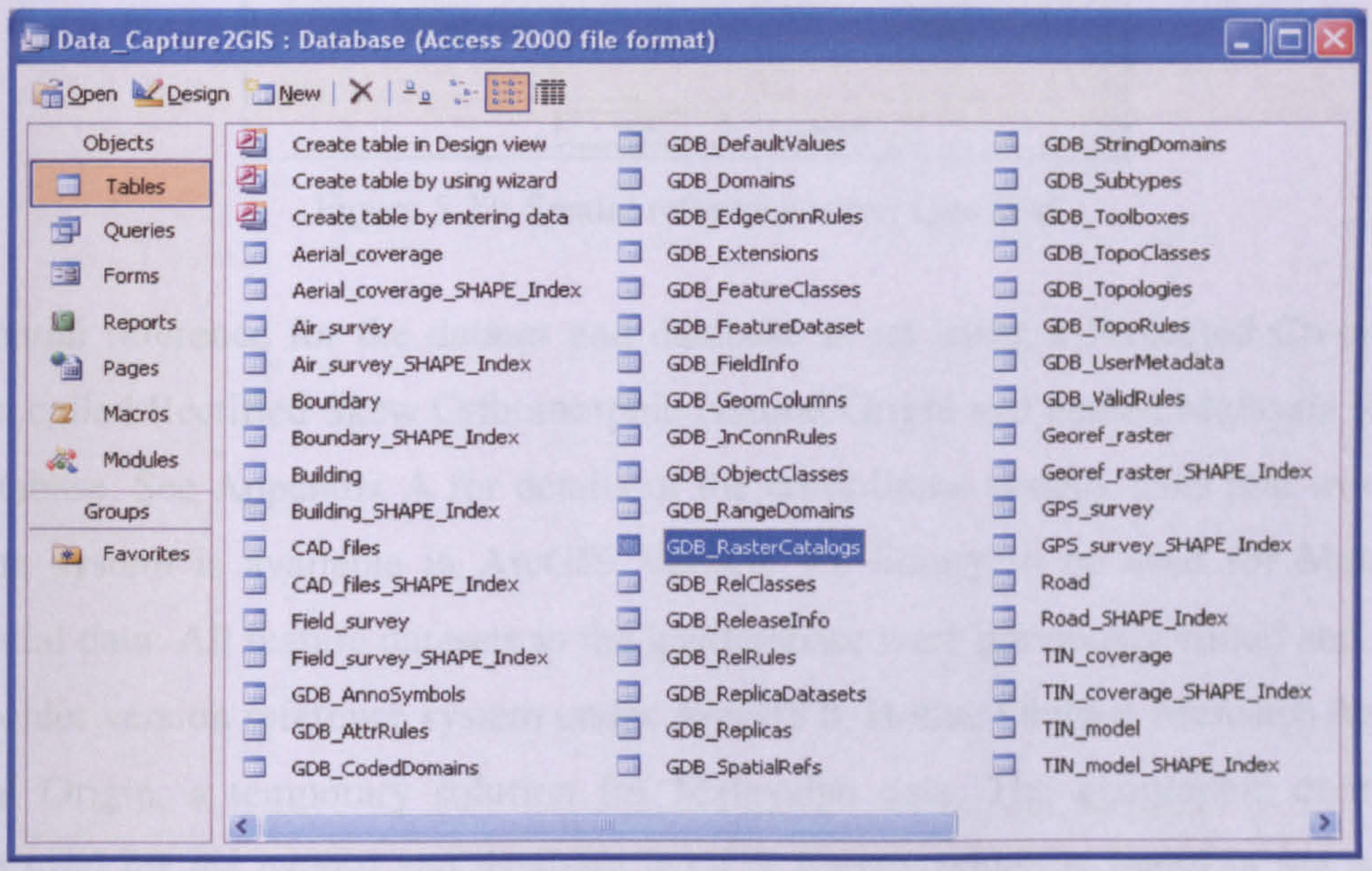


Figure 5.22: Geodatabase in Microsoft Access

5.5.4 Specifying Spatial References

When we point to the feature dataset or feature class and click on the properties, it is shown that the spatial reference is unknown and the X/Y Domains may not be representative of the dataset (Figure 5.23). Hence it is essential that spatial reference be specified to represent the local co-ordinate systems and spatial extent of the data. By clicking on the Edit button, there are options that allow specifying a new spatial reference and X/Y Domains.

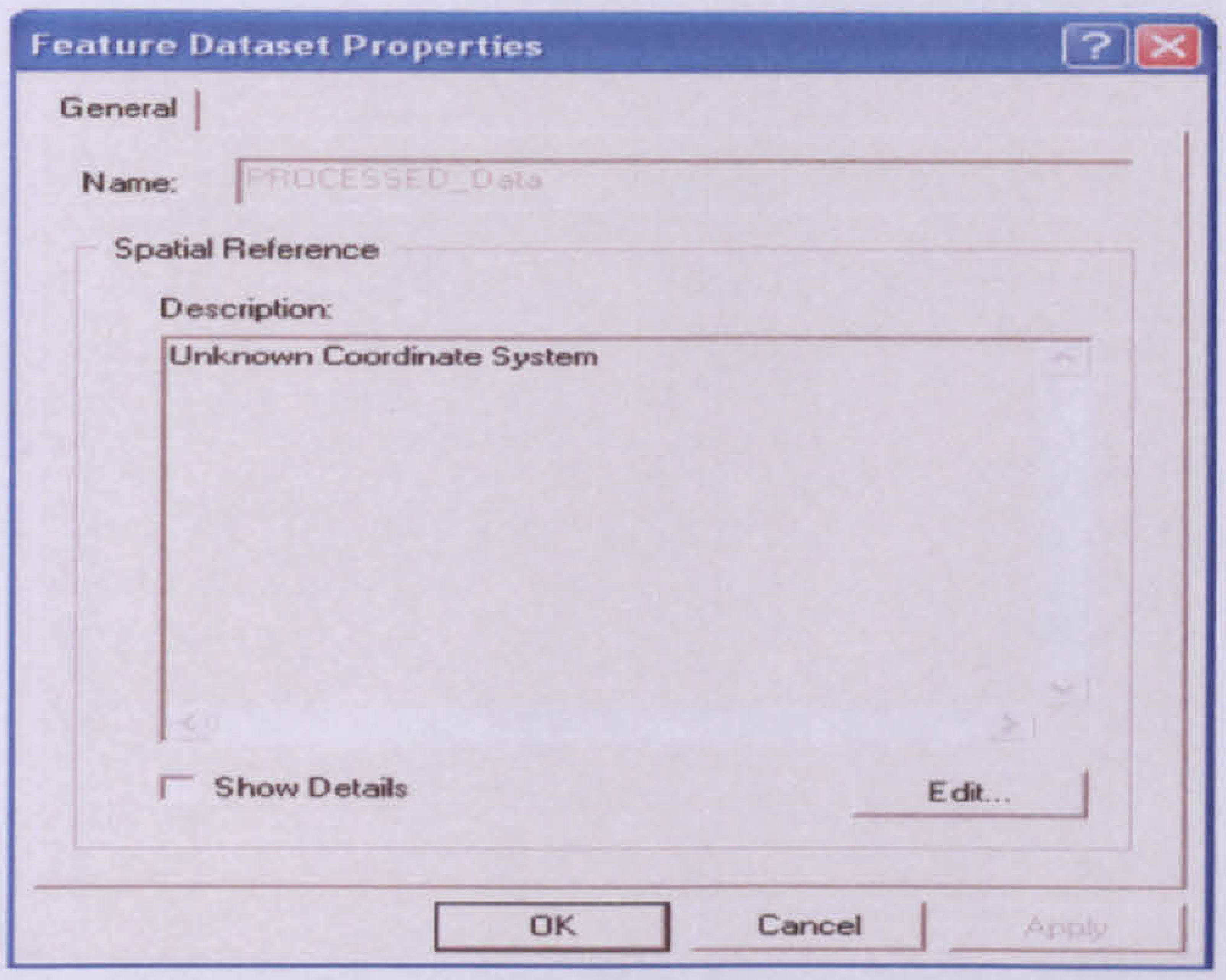


Figure 5.23: Spatial reference is not specified

The spatial reference for the dataset and database is set using a Projected Co-ordinate System called Rectified Skew Orthomorphic Natural Origin and named Malaysia RSO in the database. See Appendix A for details of the co-ordinate system. This real world co-ordinate system is available in ArcGIS Version 9's library to be used for Malaysian geospatial data. All feature datasets in the geodatabase were previously edited and stored in the older version reference system under ArcGIS 8, Hotine Oblique Mercator Azimuth Natural Origin, a temporary solution for Malaysian data. The geographic coordinate system used for the dataset and database is GCS Kertau, which is found in the ArcGIS coordinate library under Asia zone folder.

While specifying the co-ordinates, the spatial extent within with the features are to be managed is also specified. This is carried out using the X/Y Domain tab in the Spatial Reference Properties window. A MinX value of 374850, a MaxX value of 443590, a MinY value of 318910 and MaxY value of 387650 were entered as the coordinate extents of the area of interest (Figure 5.24). Z and M domains need not be used because all datasets are two-dimensional only.

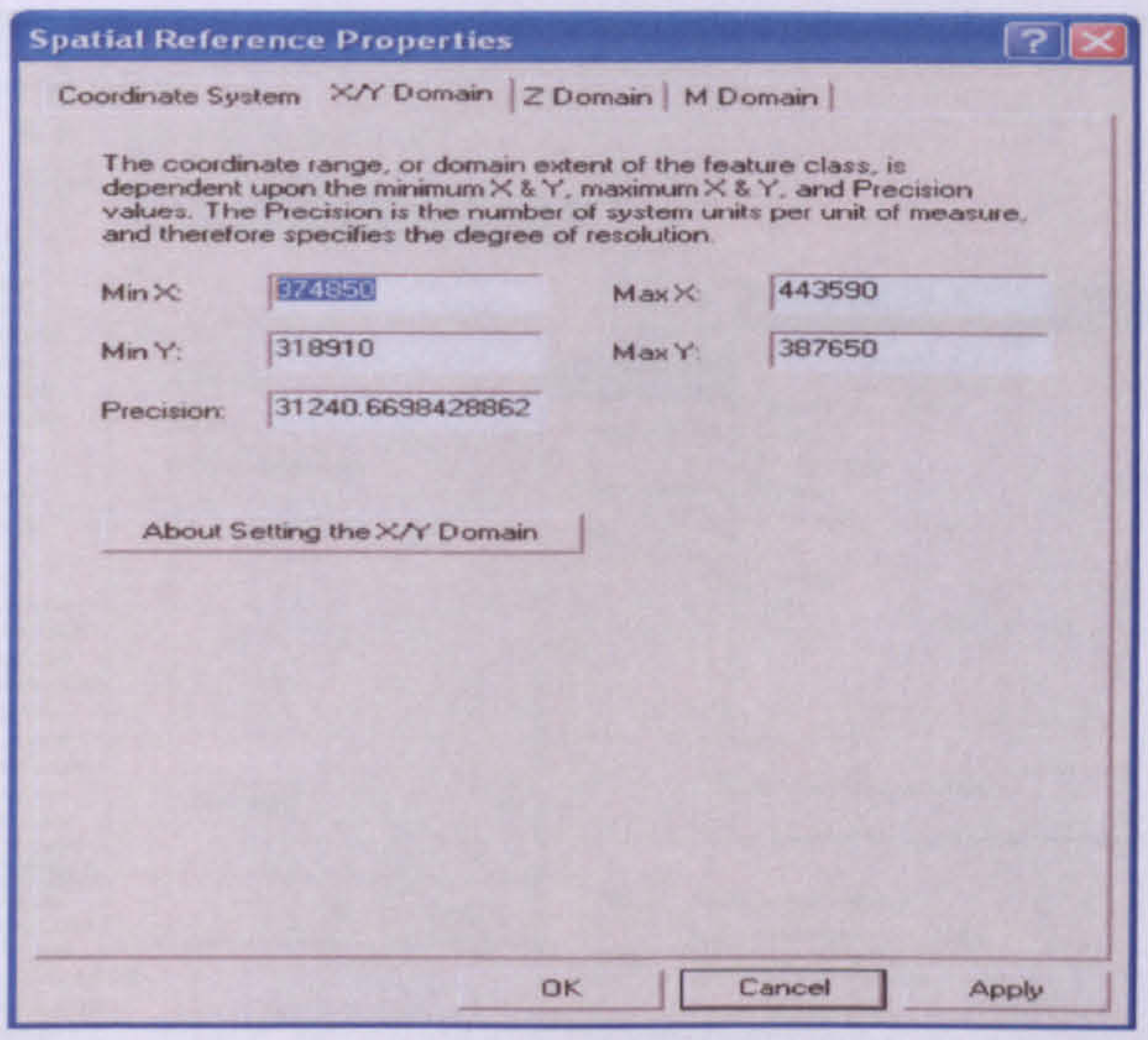


Figure 5.24: The coordinate spatial extent for the area of interest

There are two ways of specifying a coordinate system for a dataset, that is, by importing a co-ordinate system that has been specified in a dataset previously stored, and secondly by creating a new co-ordinate system if we know the spatial reference system for a region or country. Figure 5.25 shows the co-ordinate system being imported to the feature dataset in the geodatabase from a previously co-ordinate-specified dataset. Figure 5.26 shows the detail of the co-ordinate system being displayed in the wizard (left) and on the right is the name of the projected and geographic co-ordinate system being specified. By clicking 'apply' the dataset has then the spatial reference properties specified. As in Figure 5.26 there is an option to select a predefined co-ordinate system which is actually selecting a system in the co-ordinate system library similar to the option tab for creating a new co-ordinate system for the dataset area of interest.

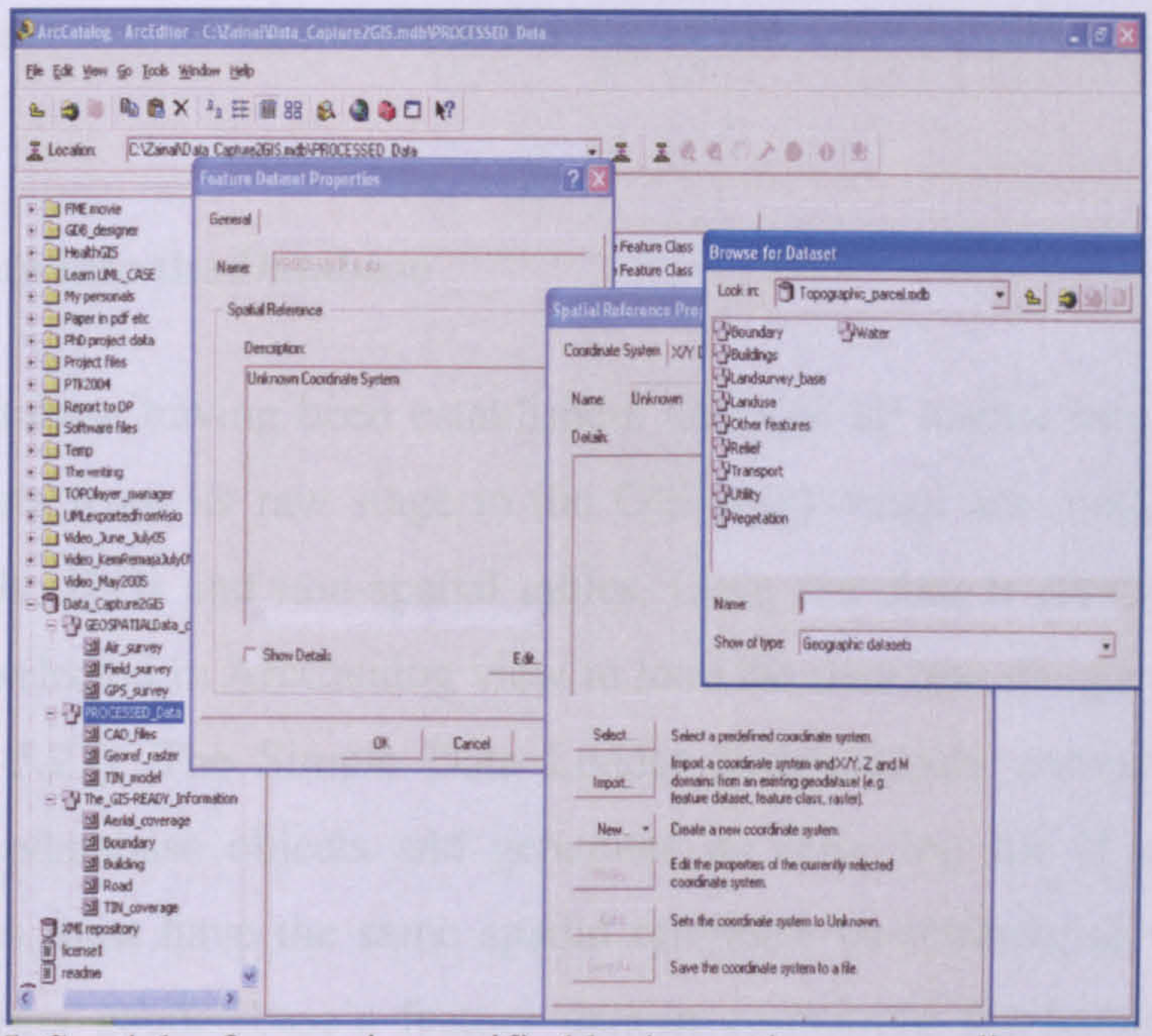


Figure 5.25: Spatial reference is specified by importing a co-ordinates system from an existing dataset (from directory folder) that already has a co-ordinates system

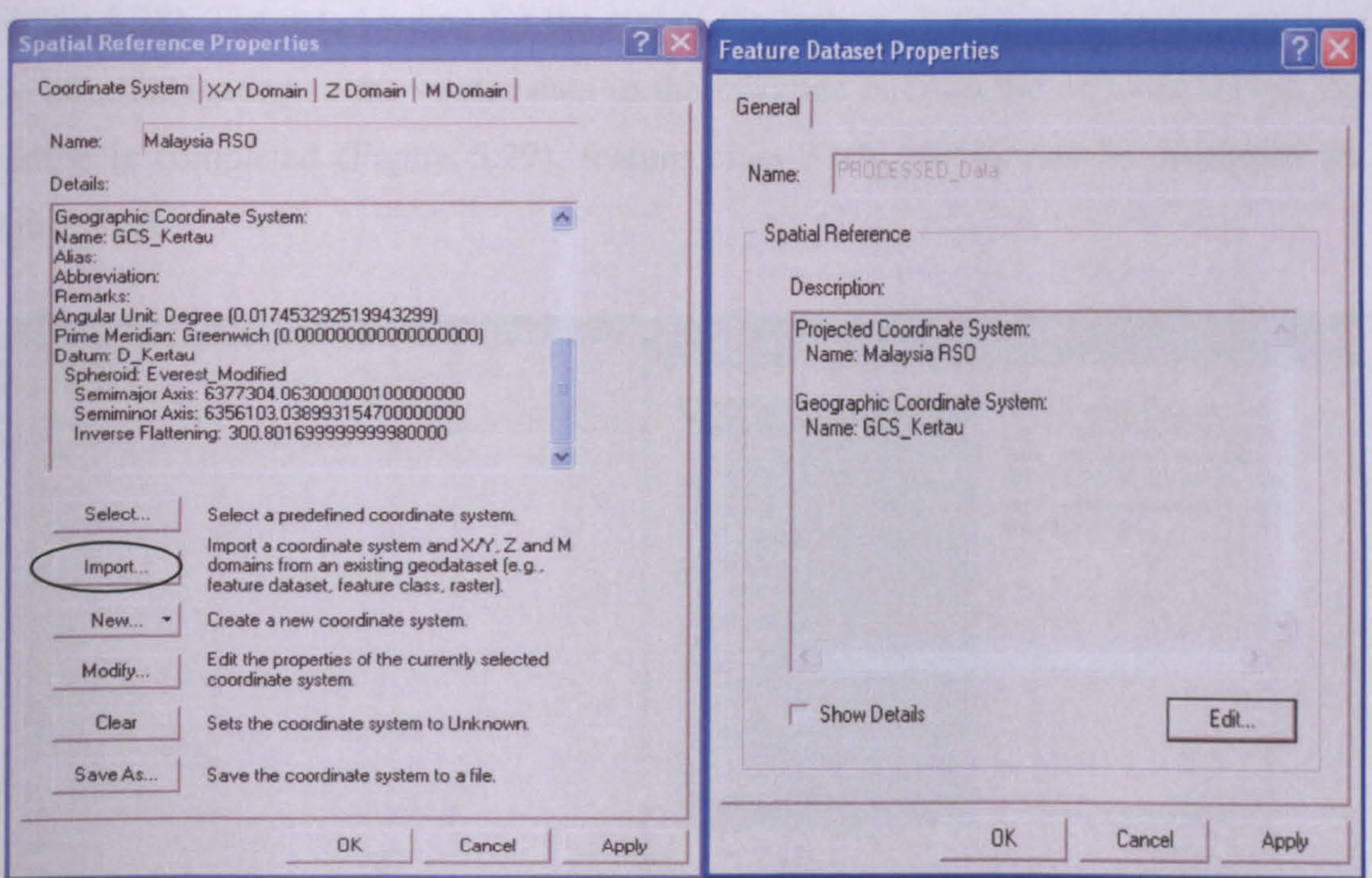


Figure 5.26: The imported co-ordinate system from a co-ordinated dataset is displayed (left) and then applied to the feature dataset (right) in the geodatabase

A geodatabase has been established in this section where objects of geospatial data from the raw capture stage to the point where it become GIS-ready information can be managed. The next section describes the loading of raw data and processed data in the geodatabase.

5.6 Loading Data into the Database

With this geodatabase having been established, data can be loaded into that schema. In this research, data from its raw stage to the GIS-ready stage are mostly in shapefiles, DXF files, IMGs, JPGs and non-spatial tables. Once the data is identified, the *Simple Data Loader* is selected in ArcCatalog view to load the data into the geodatabase feature classes (Figure 5.27). The Simple Data Loader automatically converts the shapefile features into geodatabase objects and generates an error log file if appropriate. The source shapefiles must have the same spatial reference co-ordinate as the geodatabase schema. The target geodatabase's feature dataset containing the feature classes to be loaded with data was chosen, and a summary of the data load operation was presented (Figure 5.28). The data loading for the rest of the feature classes were carried out one by one after navigating to the source data in the machine or from the network server. After loading is completed (Figure 5.29), feature class *Field_survey* can be displayed as in Figure 5.30.

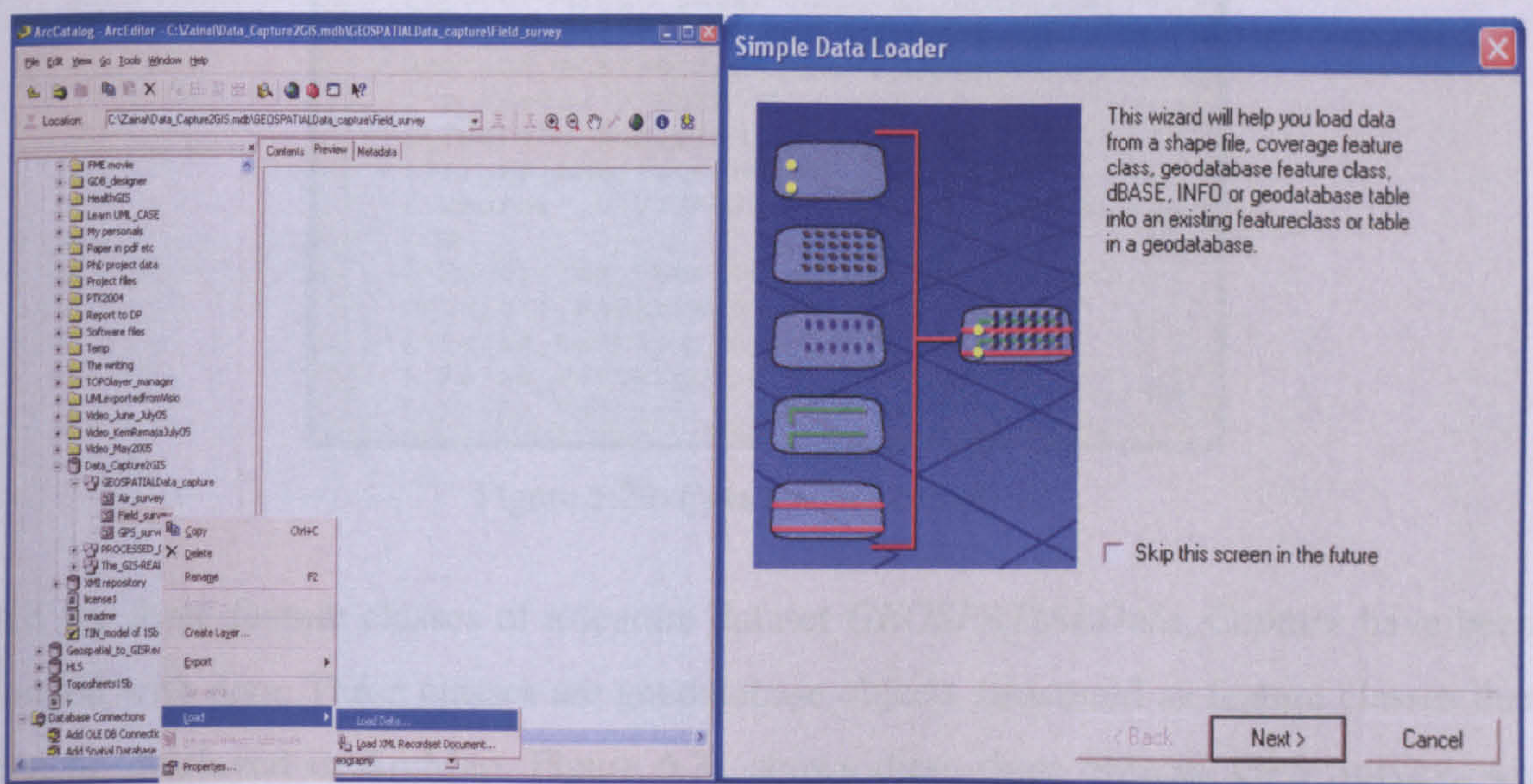


Figure 5.27: Data loading into *Field_survey* using Simple Data Loader in ArcGIS

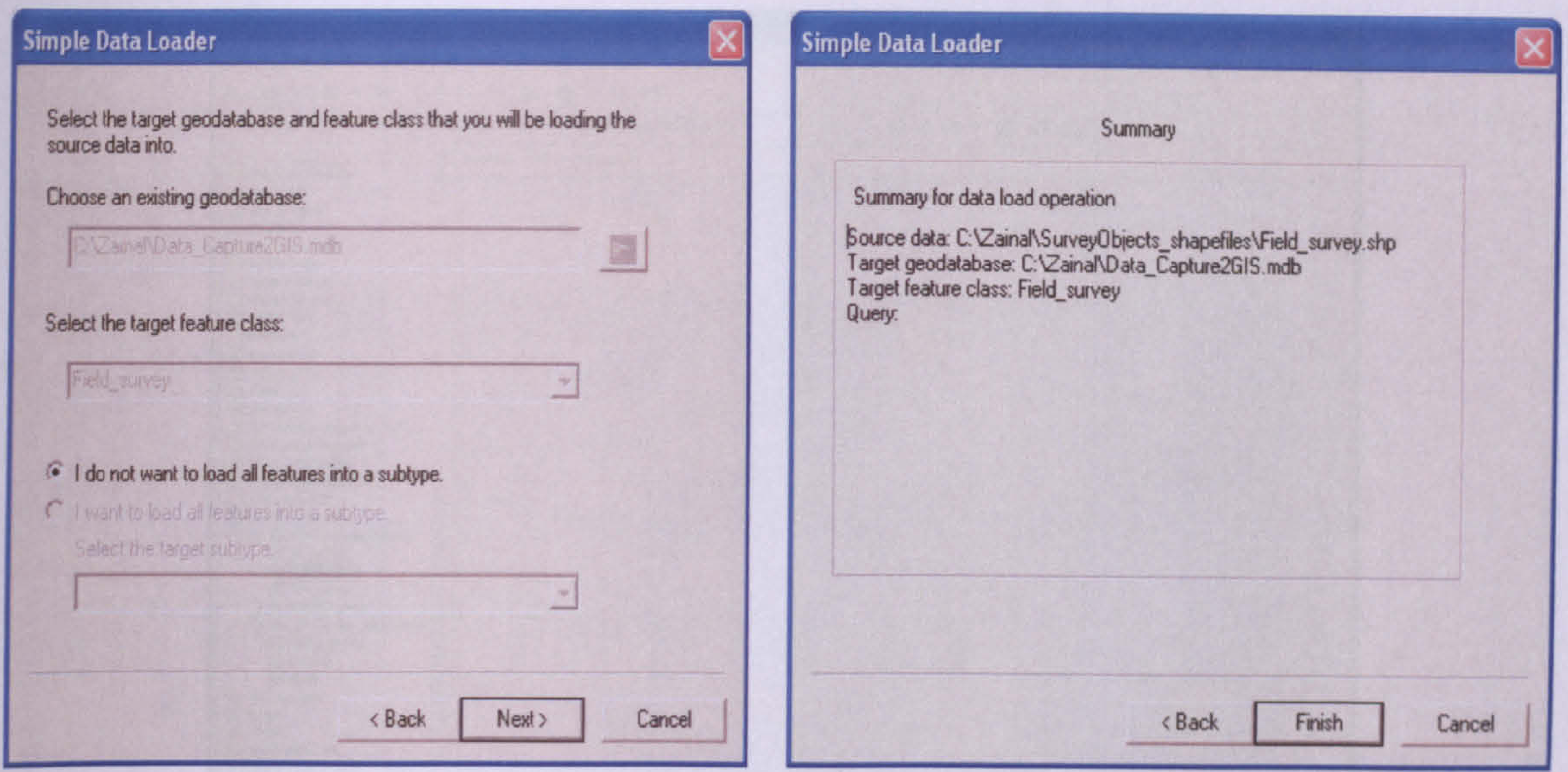


Figure 5.28: The target geodatabase is specified and summary for the operation is displayed

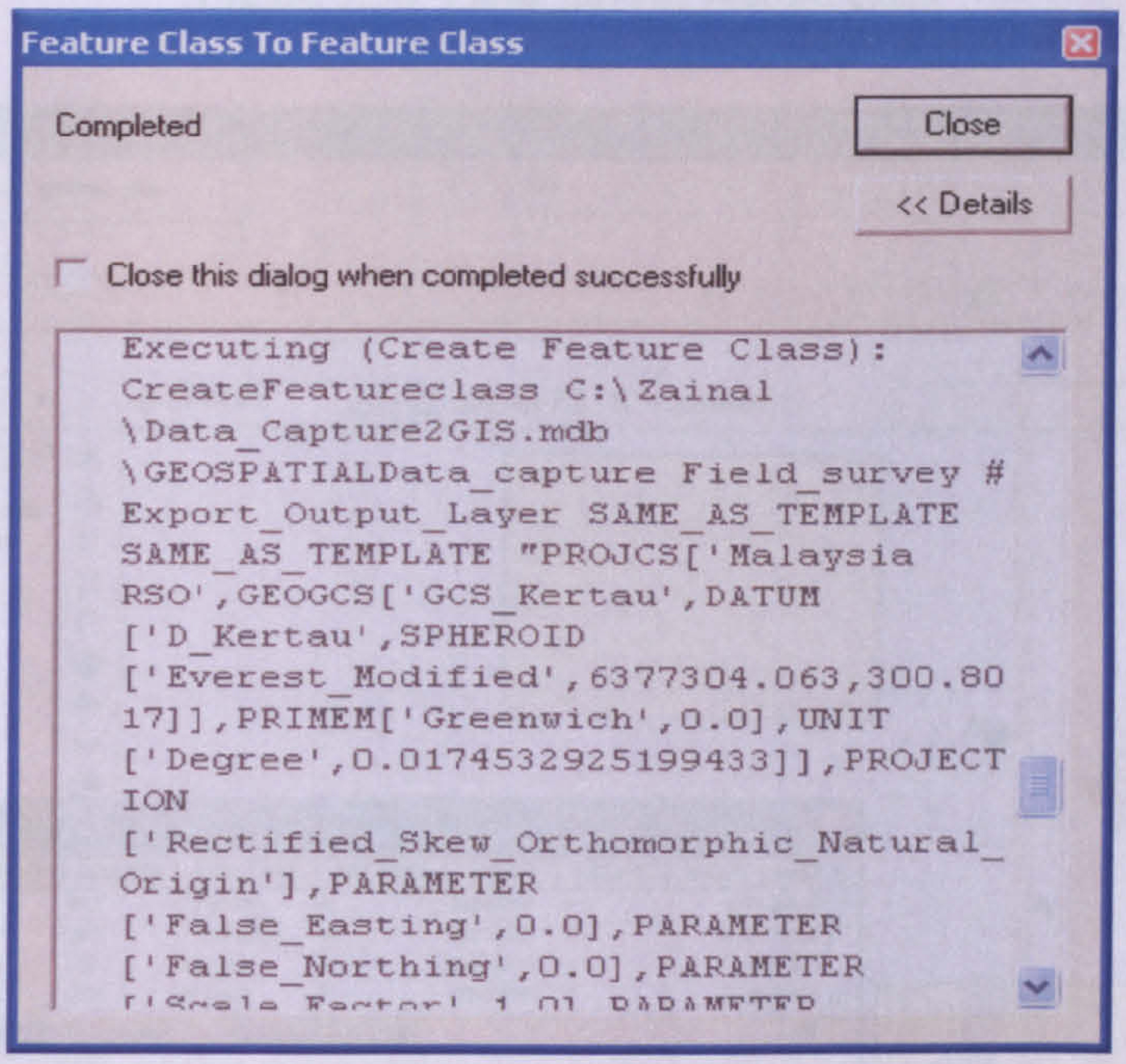


Figure 5.29: Data loading completed

All the three feature classes of a feature dataset *GEOSPATIALData_Capture* have been loaded with data. These classes are geodatabase objects structured as feature classes that can be displayed in ArcMap. Figure 5.31 shows these three objects, GPS Survey, Air Survey and Field Survey model in the geodatabase feature dataset. The attribute table of object Field Survey is presented as the properties of the object.

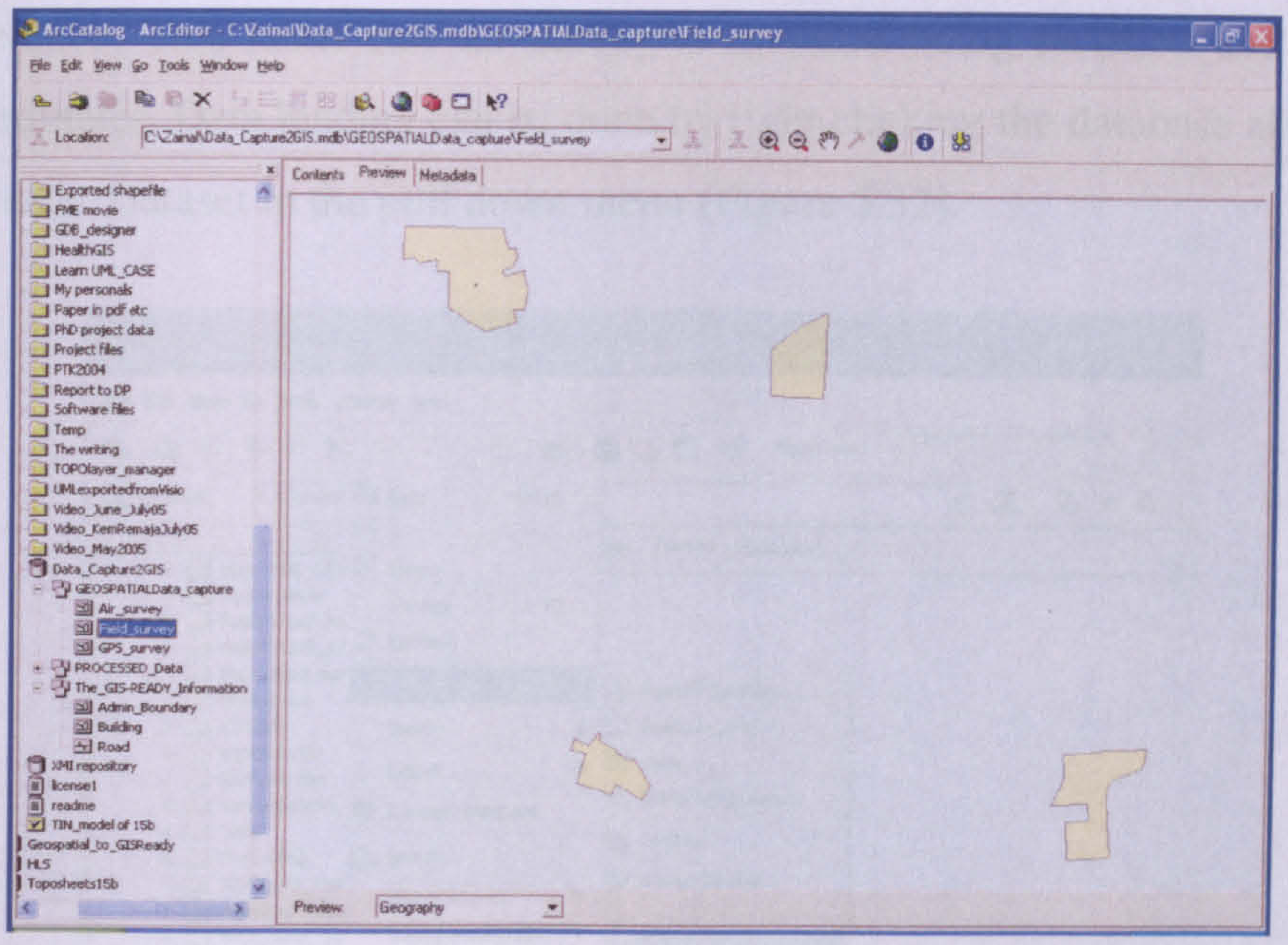


Figure 5.30: *Field_survey* feature class

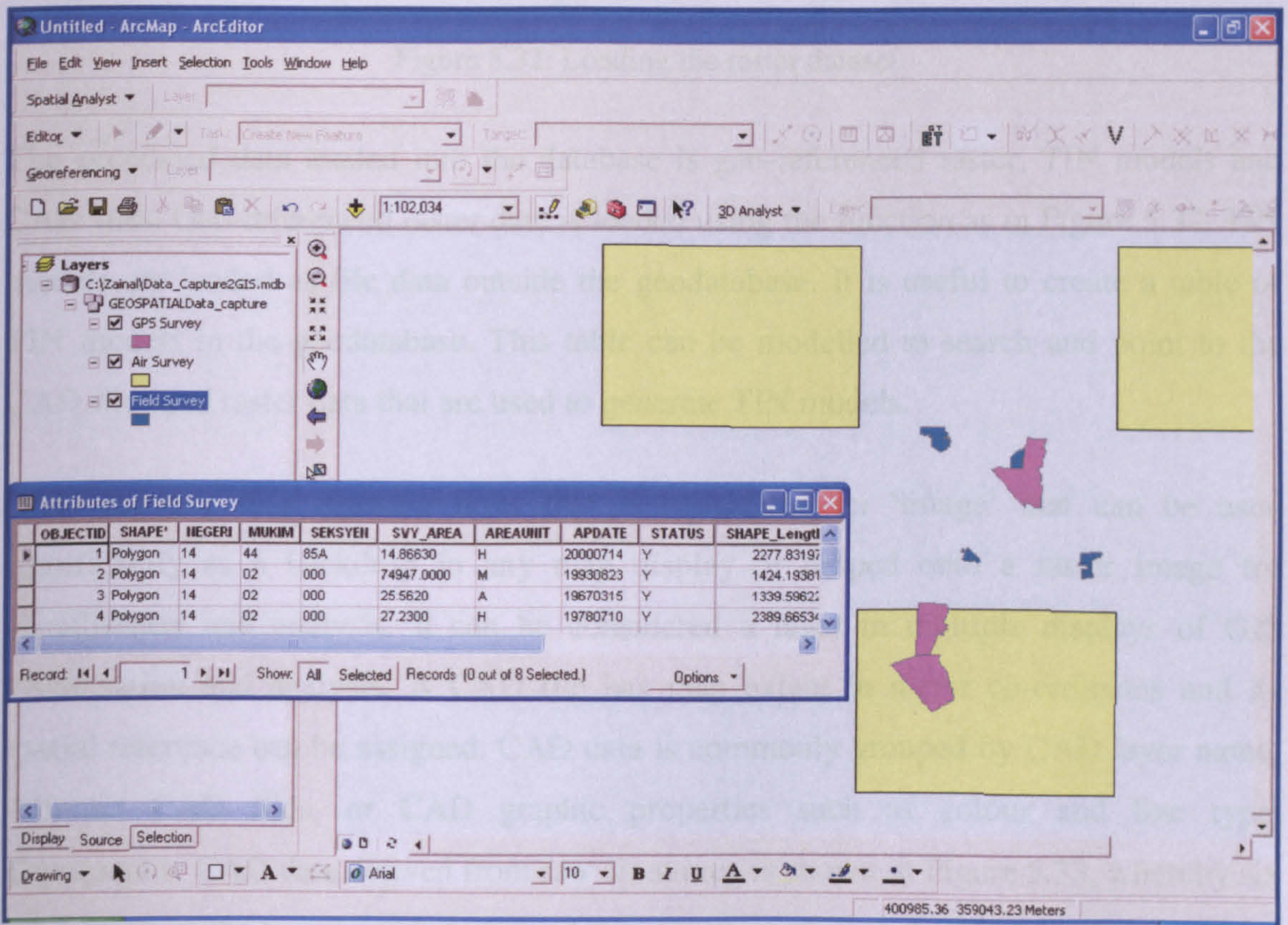


Figure 5.31: The feature dataset, feature classes and the attribute of a feature class

Raw air survey data is loaded as raster data in JPG and IMG format. Raster data is loaded into the geodatabase using the ArcCatalog's raster dataset elements after specifying spatial reference. This raster raw air survey is assumed being rectified and given real world co-ordinates. Data loading can be done by right clicking the database and choosing New and Raster Dataset in the pull down menu (Figure 5.32).

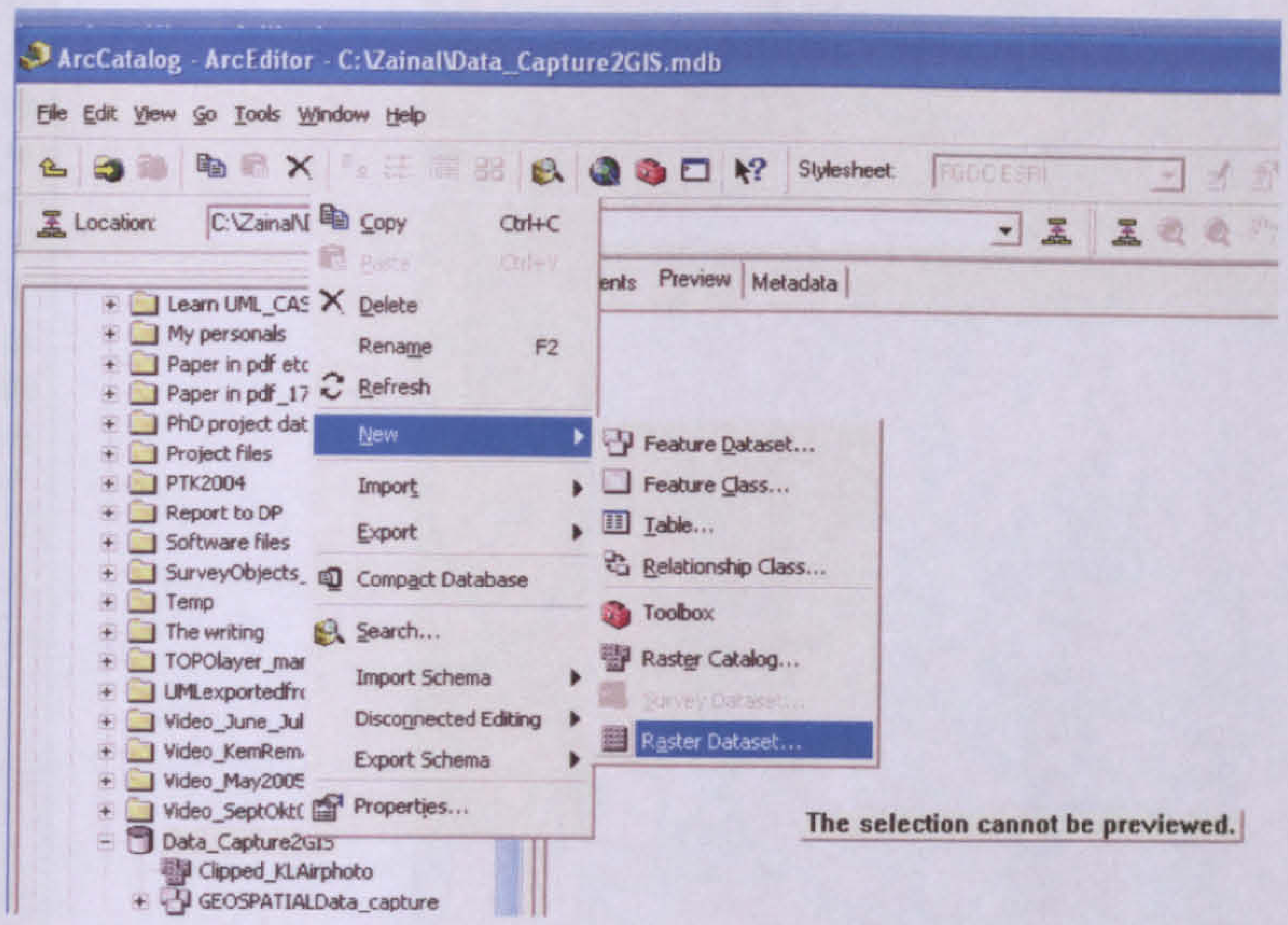


Figure 5.32: Loading the raster dataset

The processed data loaded into the database is geo-referenced raster, TIN models and CAD files. Geo-referenced raster data is loaded using the function as in Figure 5.32. TIN models are loaded as file data outside the geodatabase. It is useful to create a table of TIN models in the geodatabase. This table can be modelled to search and point to the CAD files and raster data that are used to generate TIN models.

CAD drawing is a drawing layer that is considered an 'image' that can be used significantly as a backdrop in any map display or draped onto a raster image for visualisation and analysis. It can be considered a layer in multiple displays of GIS visualisation and analysis. A CAD file has map extent in meter co-ordinates and its spatial reference can be assigned. CAD data is commonly grouped by CAD layer name, different CAD files, or CAD graphic properties such as colour and line type. Topographic CAD data derived from raw air survey is shown in Figure 5.33, whereby six map sheets of the topographic processed data are merged in an ArcMap display.

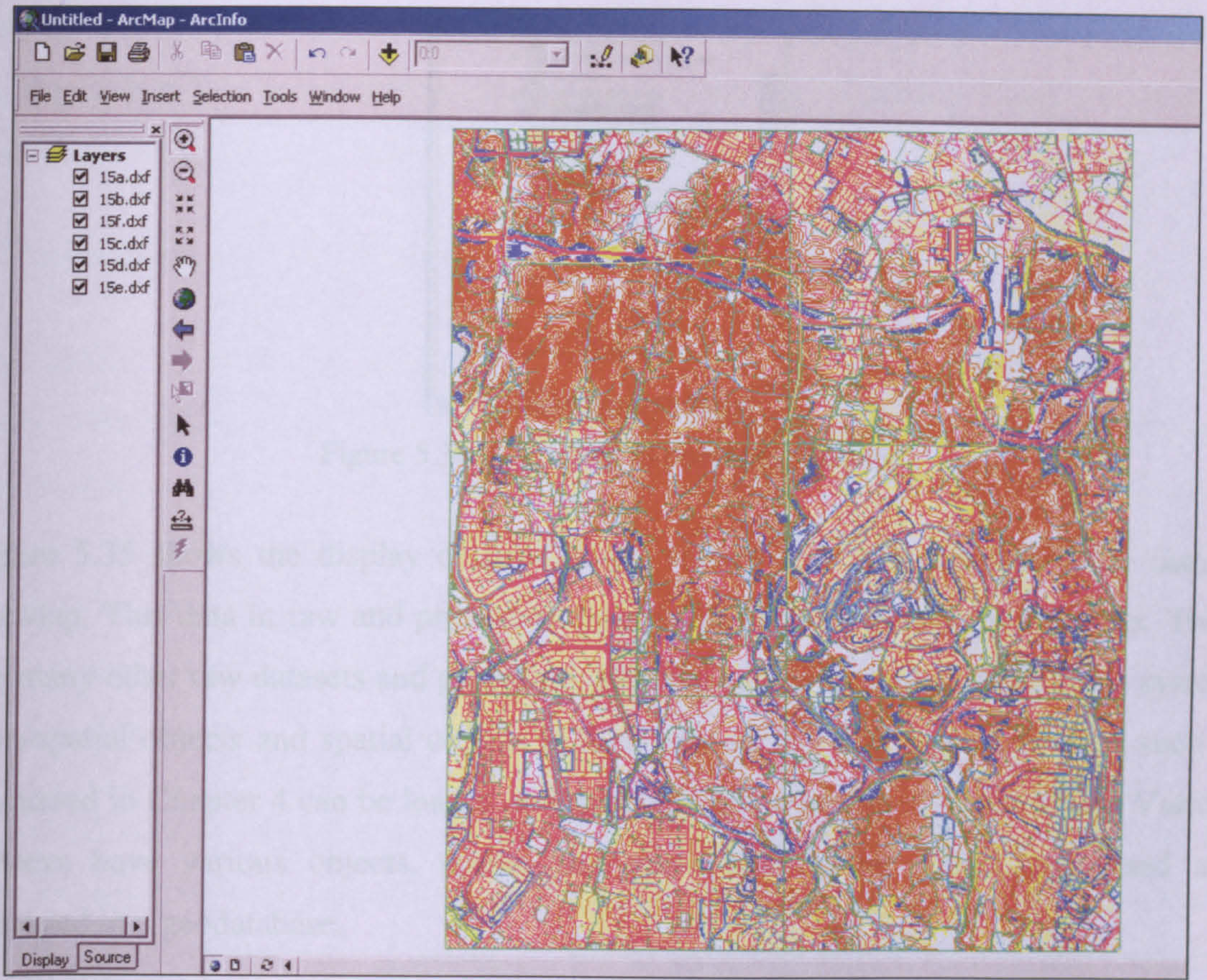


Figure 5.33: The display of six map sheets of topographic CAD data

CAD files are processed and produced by digitising and photogrammetry methods. ArcGIS has the capability of storing CAD data into five types of feature classes: polygon, polyline, point, multipath and annotation (Figure 5.34). Therefore in ArcCatalog, CAD files, in any directory or folder, are seen as feature classes. Hence, the task is to load them into a geodatabase. This is carried out by creating new CAD feature classes within the feature dataset and then importing CAD data from the folder or directory where CAD data is situated.

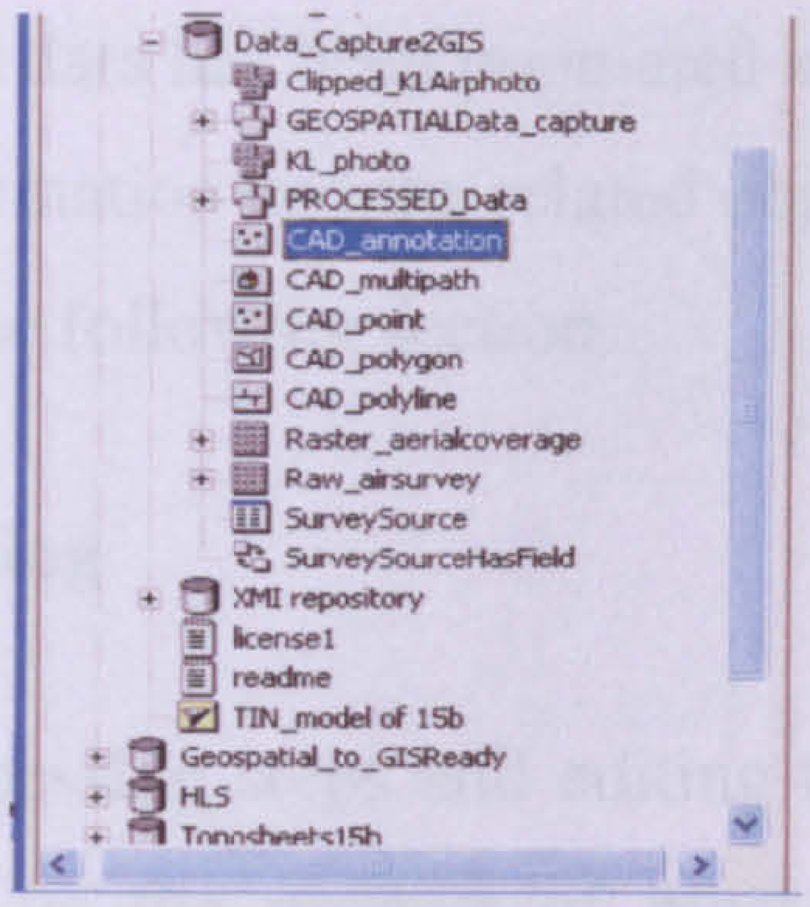


Figure 5.34: CAD data as processed dataset

Figure 5.35 shows the display of the raw air survey and its processed CAD data in ArcMap. This data in raw and processed types is displayed together in GIS tools. There are many other raw datasets and processed data that can be managed in the same system. Non-spatial objects and spatial objects in their raw survey and processed stage such as discussed in Chapter 4 can be loaded and managed in the designed geodatabase. Various surveys have various objects, spatial and non-spatial to be stored, processed and managed in a geodatabase.

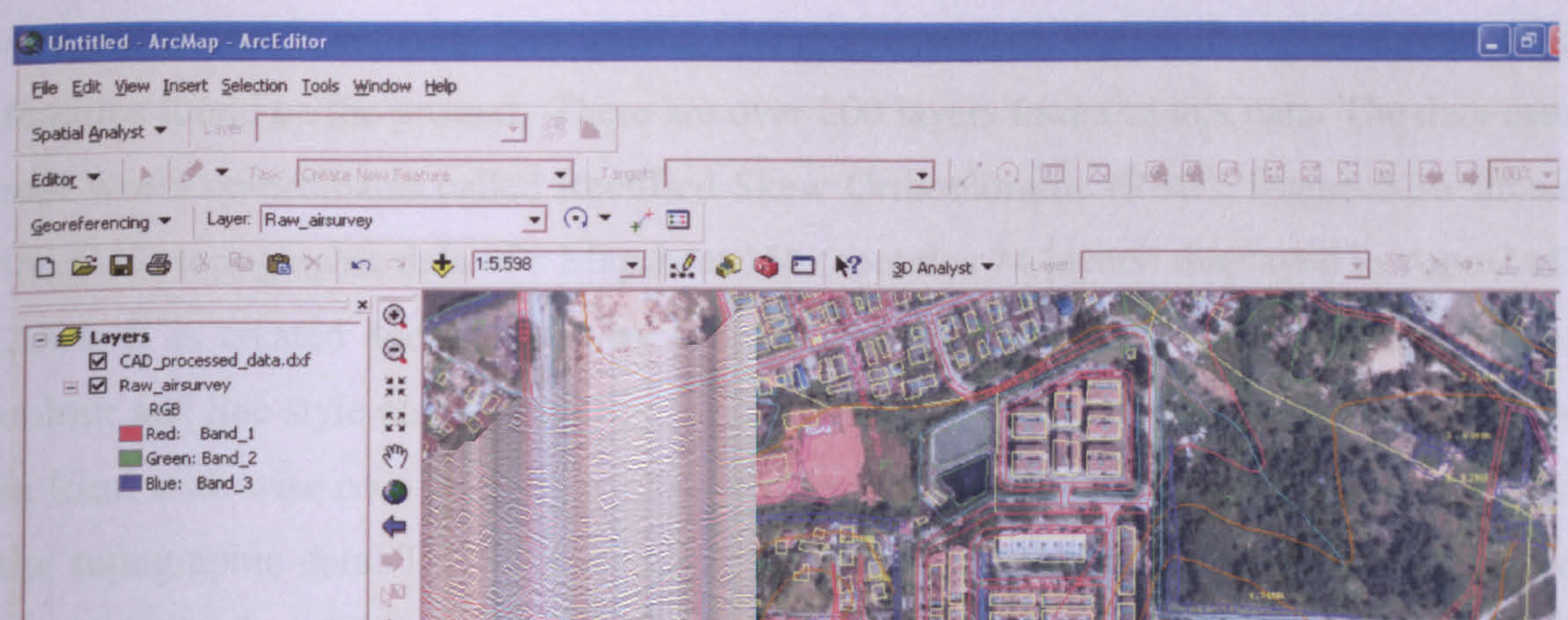


Figure 5.35: CAD processed data and raw air survey in the same application

Raw survey data and processed data has been populated into the geodatabase. The third feature dataset, GIS-ready information and the related objects are processed and created in the next steps described in the following section.

5.7 Processing Steps and Editing

This section describes the processing steps and editing towards the production of GIS-ready information. CAD data is the product of most surveys, such as air surveys, cadastral surveys and field topographic surveys, carried out in JUPEM. They are used for national mapping, cadastral title plan and population into digital cadastral and topographic map sheets. JUPEM uses the CAD format to sell digital topographic maps derived from aerial photographs and field surveys. This data is not GIS-ready data but processed surveys.

The CAD data for the project was provided by JUPEM. They comprise CAD drawing files produced for each map sheet in the topographic map sheet. Each map sheet consists of a large-scale digital map of detailed surfaces on the landscape, with relative positions and elevations. It contains information in annotations and inserts about the topographic features found on the ground. There are over 200 layers found in this data. The data uses real world co-ordinates called Rectified Skew Orthomorphic (RSO). Figure 5.36 shows the CAD topographic data of a Mapsheet 15b (contains 74 layers) displayed in AutoCAD 2002. It is created and grouped by CAD layer name and CAD graphic properties of colour and line style combined in one single CAD file. CAD data for cadastral surveys is in local state-wise co-ordinates. It was transformed into real world co-ordinates similar to the topographic data. This is needed so that cadastral data and topographic data can be superimposed on the same platform of spatial reference.

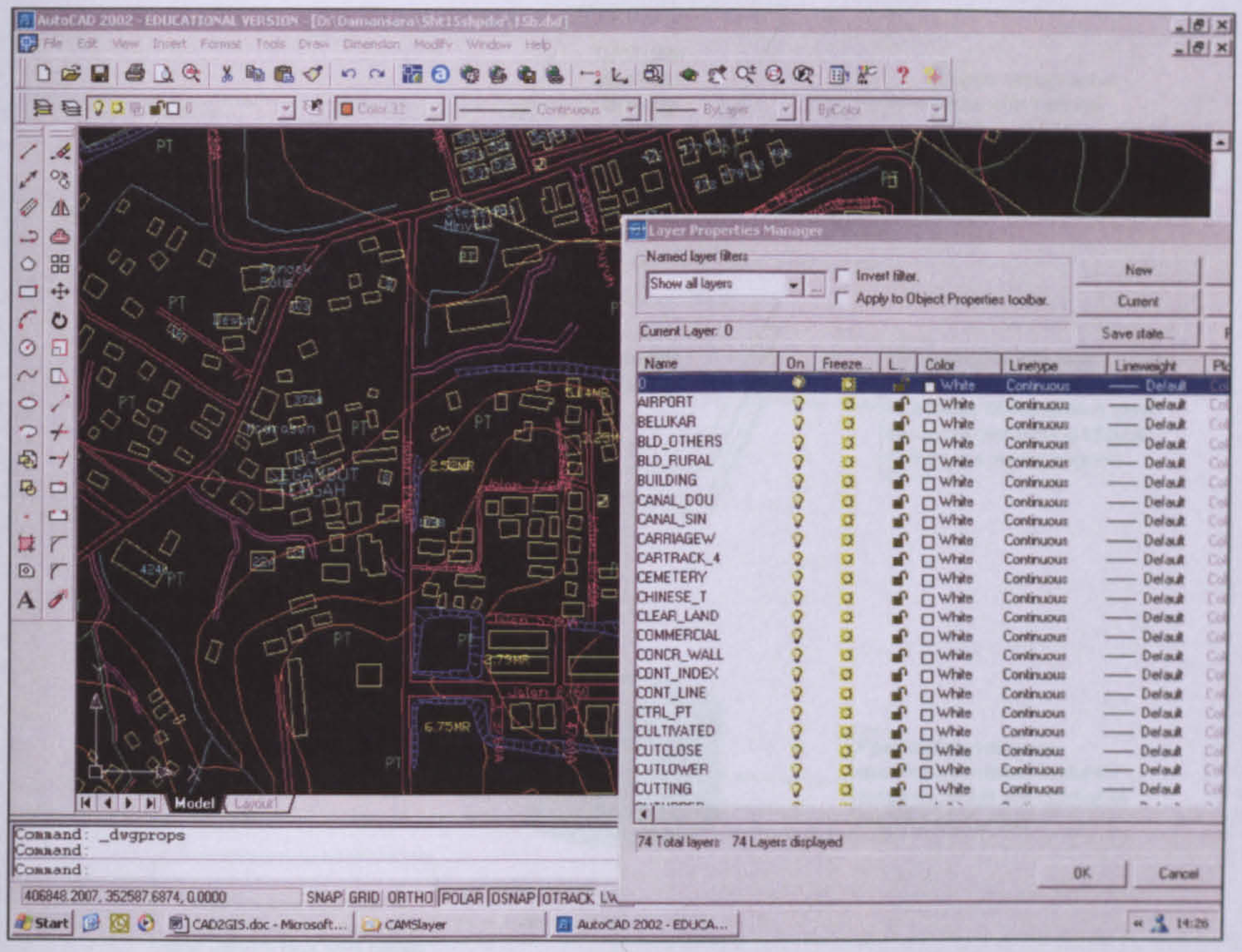


Figure 5.36: CAD data that has 74 layers is displayed in AutoCAD 2002

The transformation of the CAD digital to GIS-ready information is carried out as illustrated in the flow chart of Figure 5.37. Editing was also done to ensure the fully functionality of GIS-ready datasets. The following explains the processing and editing steps undertaken:

- The CAD drawing file was translated into layers of shapefiles using Feature Manipulation Engine (FME) software. FME software converted the CAD file into shapefile format before it could be stored and edited. FME Workbench interface was used to translate them to form object-oriented features. The translated data became feature classes for GIS display covering contiguous map coverage. The source datasets were accessed from the geodatabase and transferred into the same database directly during the translation. Figure 5.38 shows the diagrammatic flow of the translation in FME.
- The co-ordinates system was read by FME as RSO and the transformation resulted in the same co-ordinate set for the transformed data.

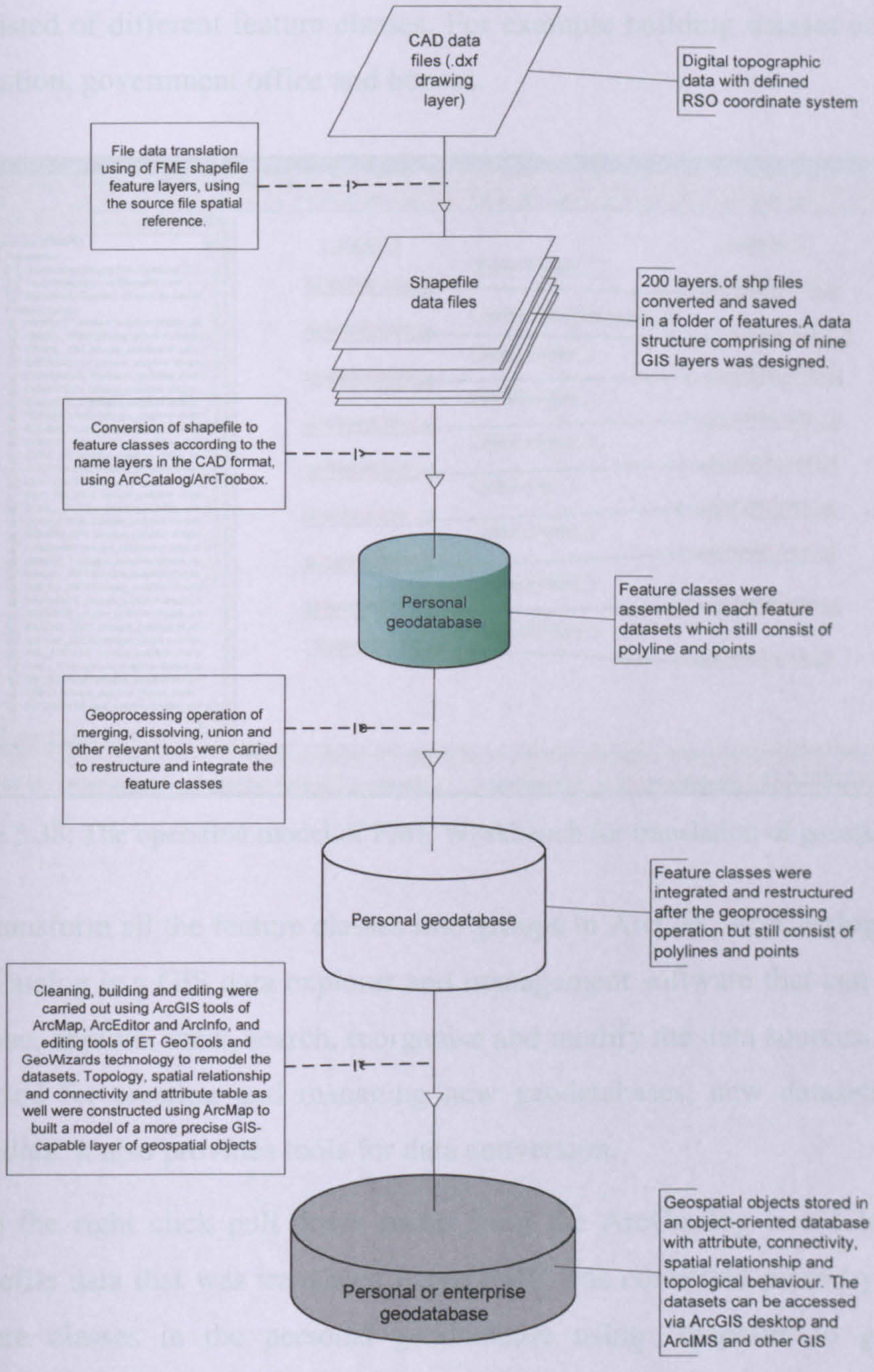


Figure 5.37: The processing flow of the CAD data to create GIS layers

- The translated datasets which consisted of many layers of features from the CAD files were saved in a folder for editing processes. For easy search and manipulation in later stages, it is crucial to group them. These layers were viewed and analysed and then a data structure was analysed and constructed. These features were collapsed into nine layers. They are building, boundary, transport, relief, water, vegetation, utility, zoning and miscellaneous. Each dataset consisted of different feature classes. For example building dataset comprises of education, government office and houses.

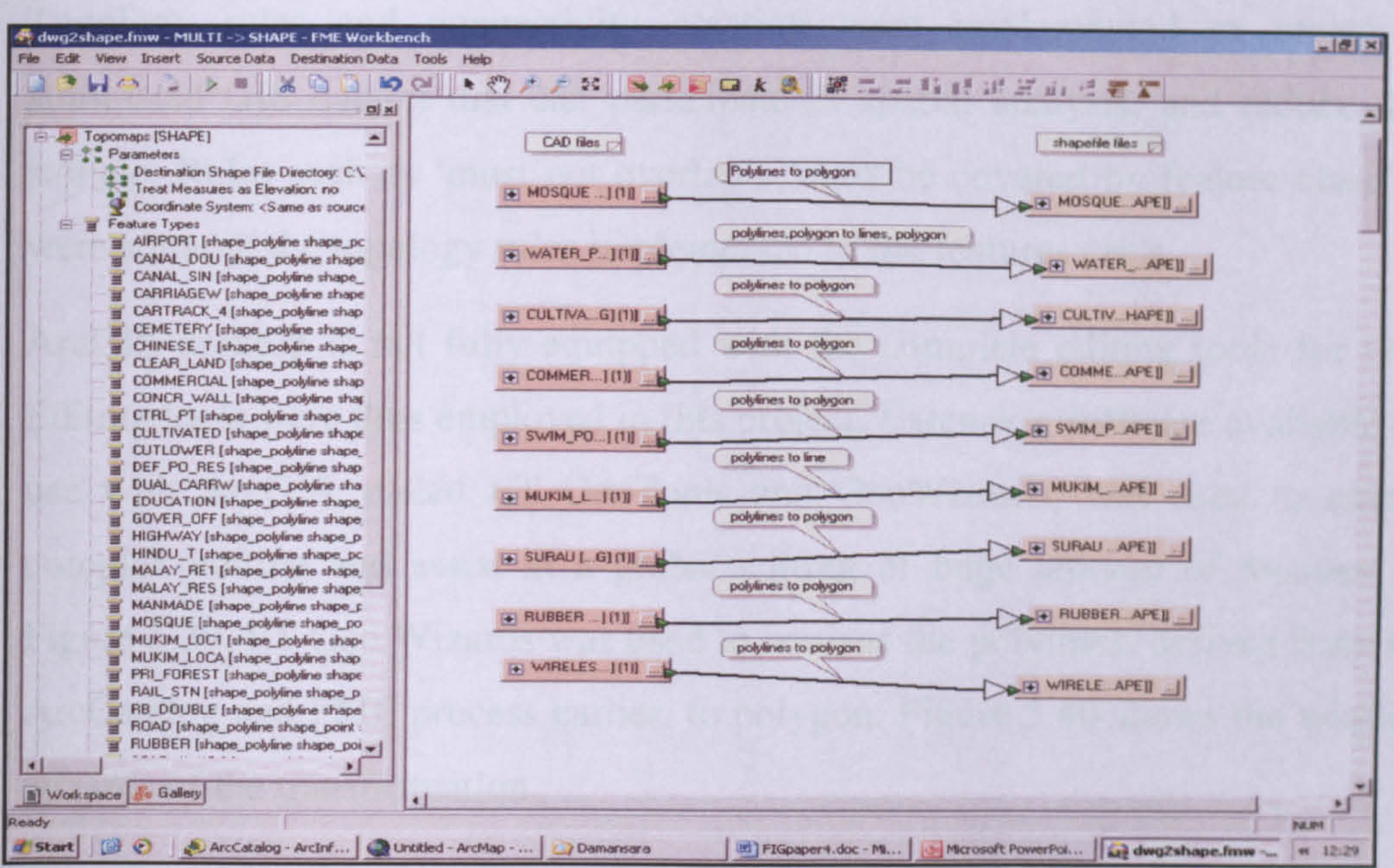


Figure 5.38: The operation model of FME Workbench for translation of geospatial data

- To transform all the feature classes into groups in ArcGIS, ArcCatalog was used. ArcCatalog is a GIS data explorer and management software that can be used to browse, explore, view, search, reorganise and modify the data sources. It can also be used for creating and managing new geodatabases, new datasets and new metadata. It also provides tools for data conversion.
- With the right click pull down menu from the ArcCatalog tree folder, all the shapefile data that was translated using FME was converted piece by piece into feature classes in the personal geodatabase using 'shapefile to geodatabase wizard'.
- All the features were displayed and studied for the possibility of geo-processing operations in ArcMap. Geo-processing procedures such as merging, dissolving, and union were utilised to integrate, separate and manipulate the feature classes accordingly for appropriate GIS layer usage and management. Relevant land parcels, for example need to be within administrative boundaries.

- As all the data were in polyline, cleaning, building and editing have to be done to create topology and relationship between feature classes.
- Topology rules and connectivity creation were implemented to create an automated GIS dataset that can participate in spatial analysis, and reduce data storage. Rules such as 'must not overlap', 'must be covered by feature class of' were some of the topology rules implemented to the features class.
- ArcGIS desktop is not fully equipped with the complete editing tools for GIS. Editing tools were thus employed in this project. Extension software available for use with ArcGIS, called ET GeoTools and GeoWizards, was used to extend complex editing and assist in a global editing of huge amount of features. In Figure 5.39, ET Geo Wizards was used to convert the polylines, derived from the ArcCatalog and FME process earlier, to polygon. Figure 5.40 shows the window wizards of the transformation.

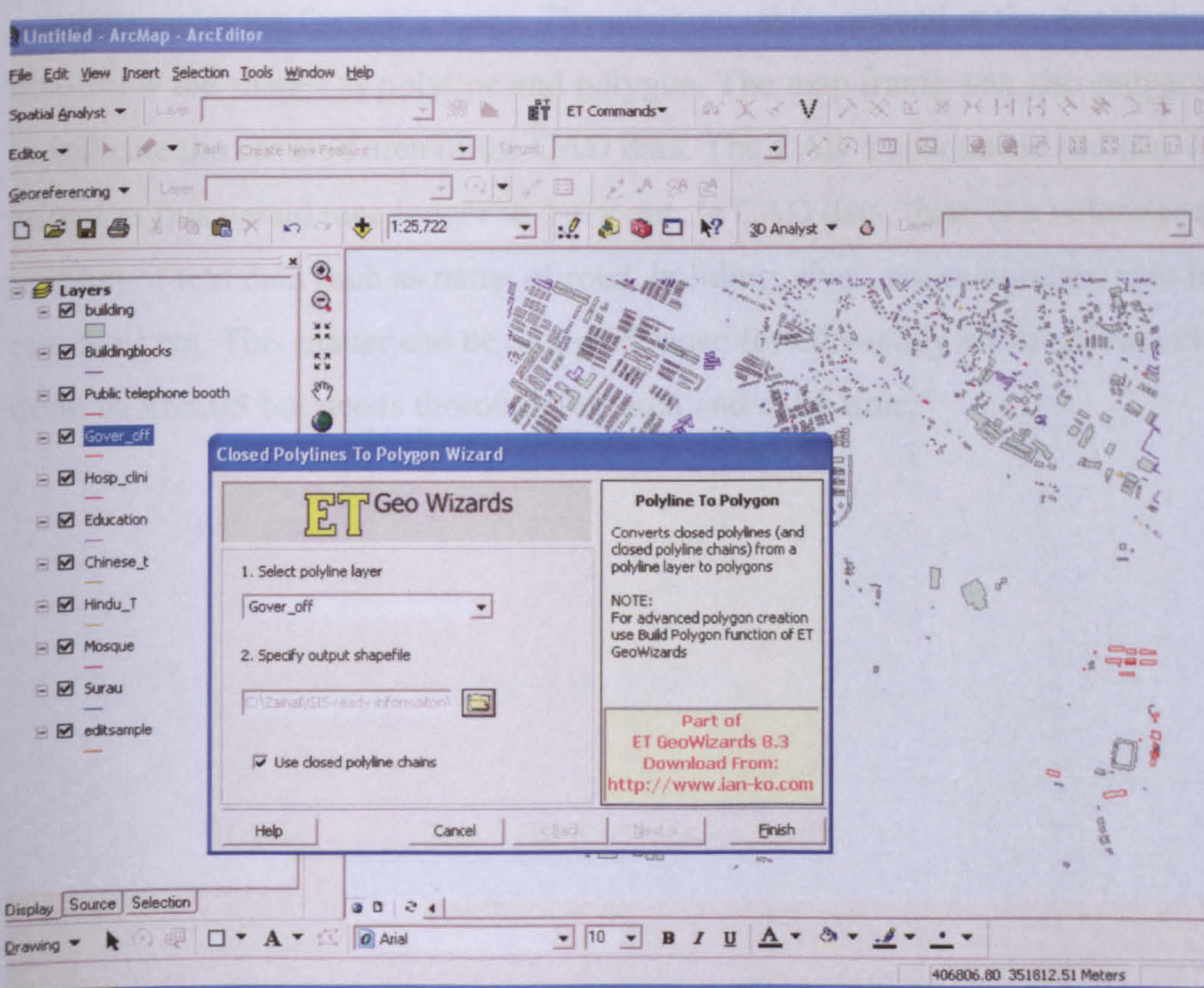


Figure 5.39: By clicking on an icon the software is able to convert feature in the layers that are displayed on the left

- Editing of the geospatial objects results in layers of GIS that are clearly featured in point, lines and polygon and stored in a personal geodatabase that can be accessed from any ArcGIS software and ArcIMS.

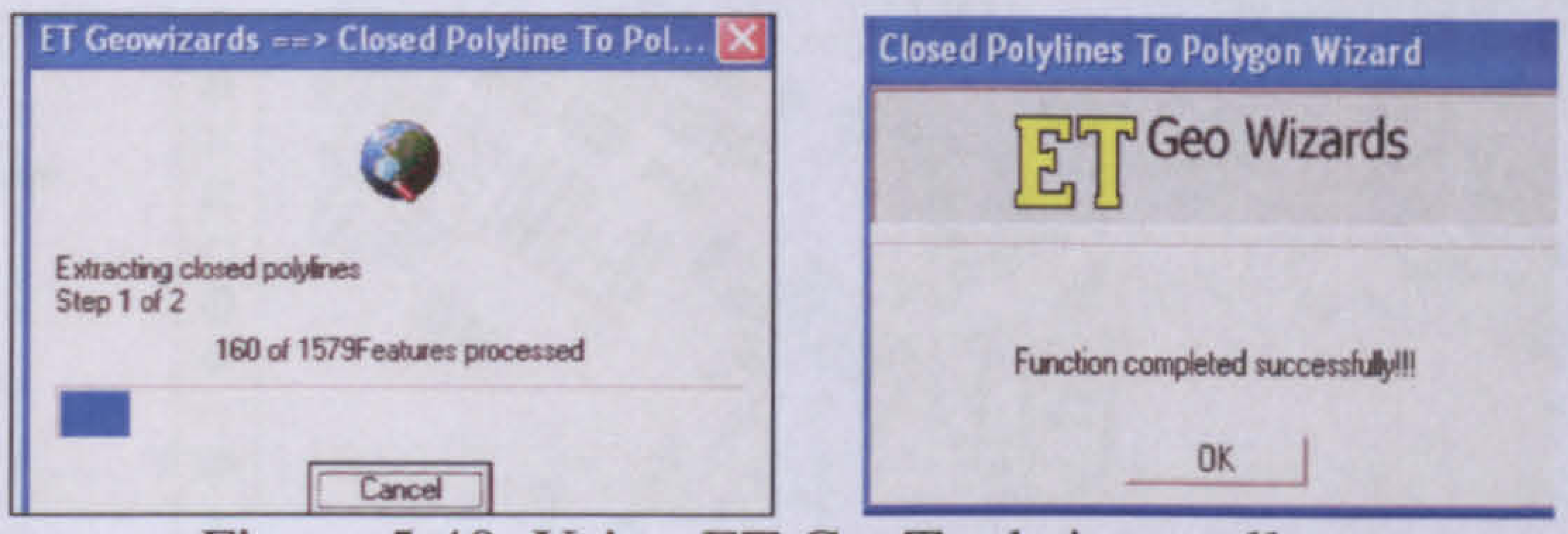


Figure 5.40: Using ET GeoTools is excellent for the huge amount of features in CAD data

Figure 5.41 illustrates the feature classes that are in the form of polylines, points, and closed polylines that need to be edited. They are polylines as shown under the feature's name in the table of content, on the left side in ArcMap. Figure 5.42 shows the edited version which is depicted by the graphic polygon as shown in the table of contents under the feature's name. The attribute tables opened in the ArcMap view also show the shapes as polyline and polygon. The map frame was also polygonised to indicate the covered area of the CAD data. The CAD's annotation and text was not edited as this would take longer to carry out. In CAD data, there is a voluminous amount of text data such as name of road, building, river, house number, spot height, reservoir etc. This matter can be of significance for GIS-ready information and can be done in ArcGIS but needs thorough research and extra time.

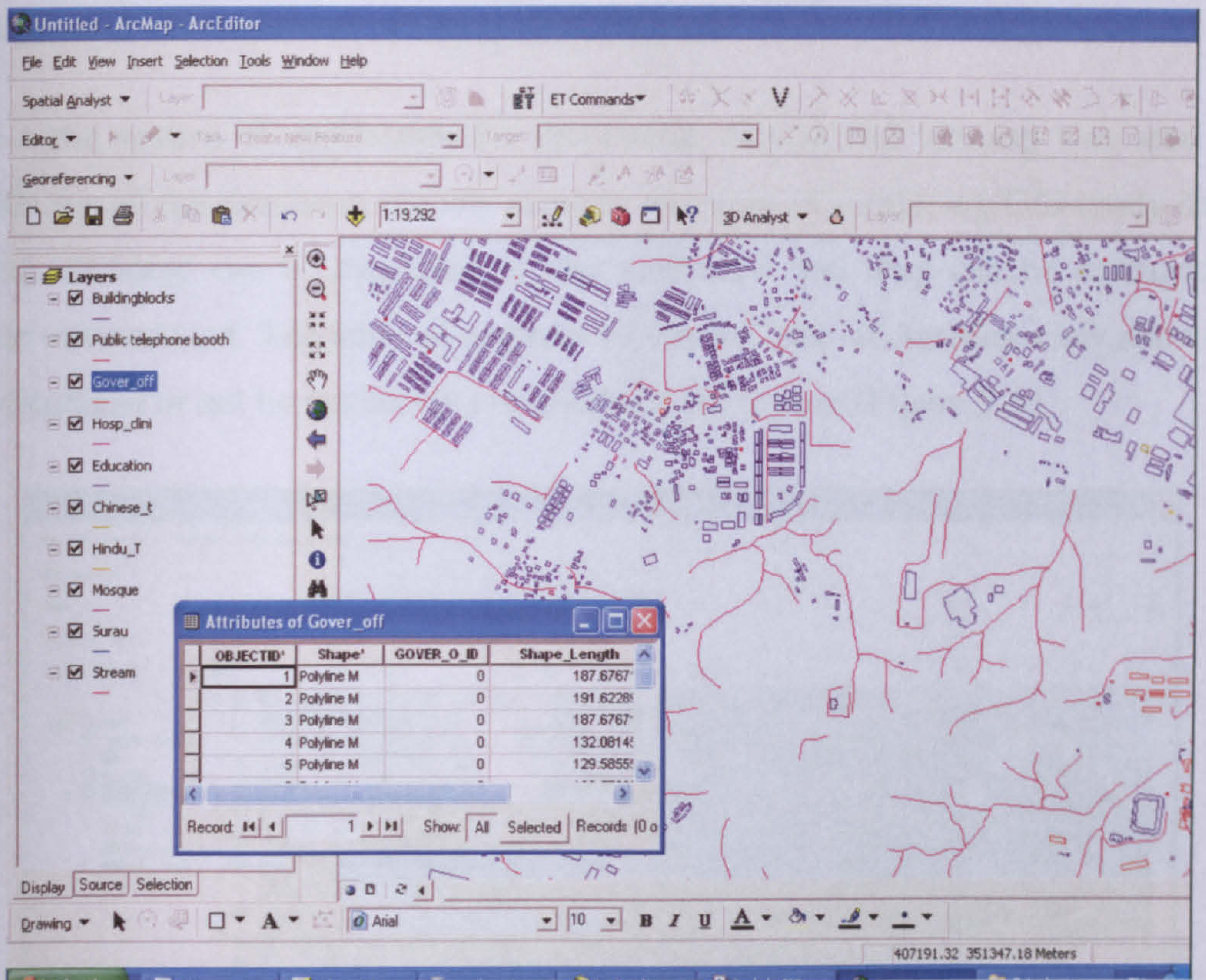


Figure 5.41: Feature classes that are still in point, polylines and closed polylines

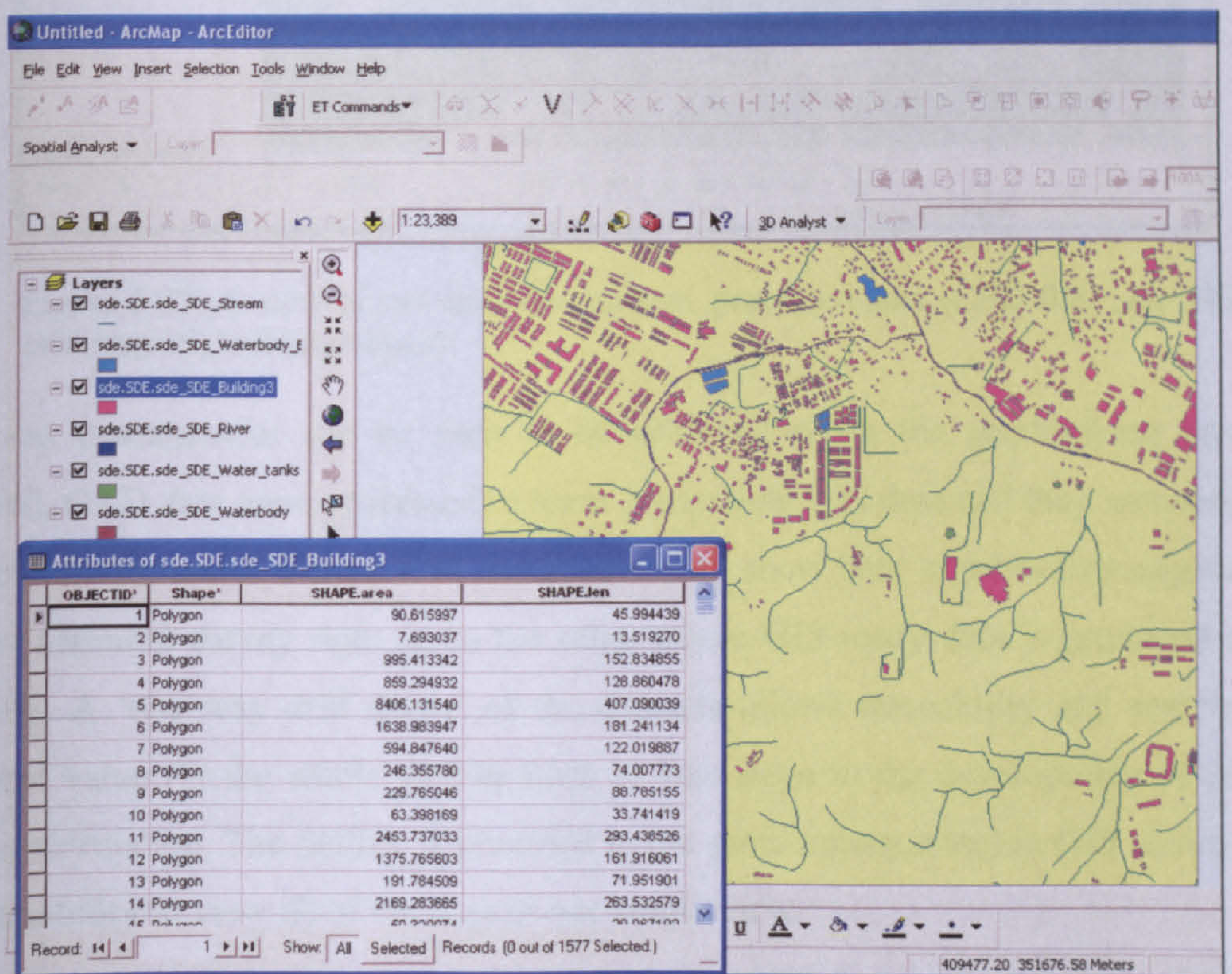


Figure 5.42: Feature classes that have been processed and edited to form GIS-ready information

5.8 Improved Management of Geospatial Datasets

In the same system of application the processing, editing and viewing was seen to be possible for all the data from the raw stage to the stage of producing GIS-ready data. In ArcMap all these can be displayed in one map view and they can be managed in a flexible environment. The table of contents on the left side of ArcMap view allows data to be displayed or not by unchecking or checking the feature (Figure 5.43).

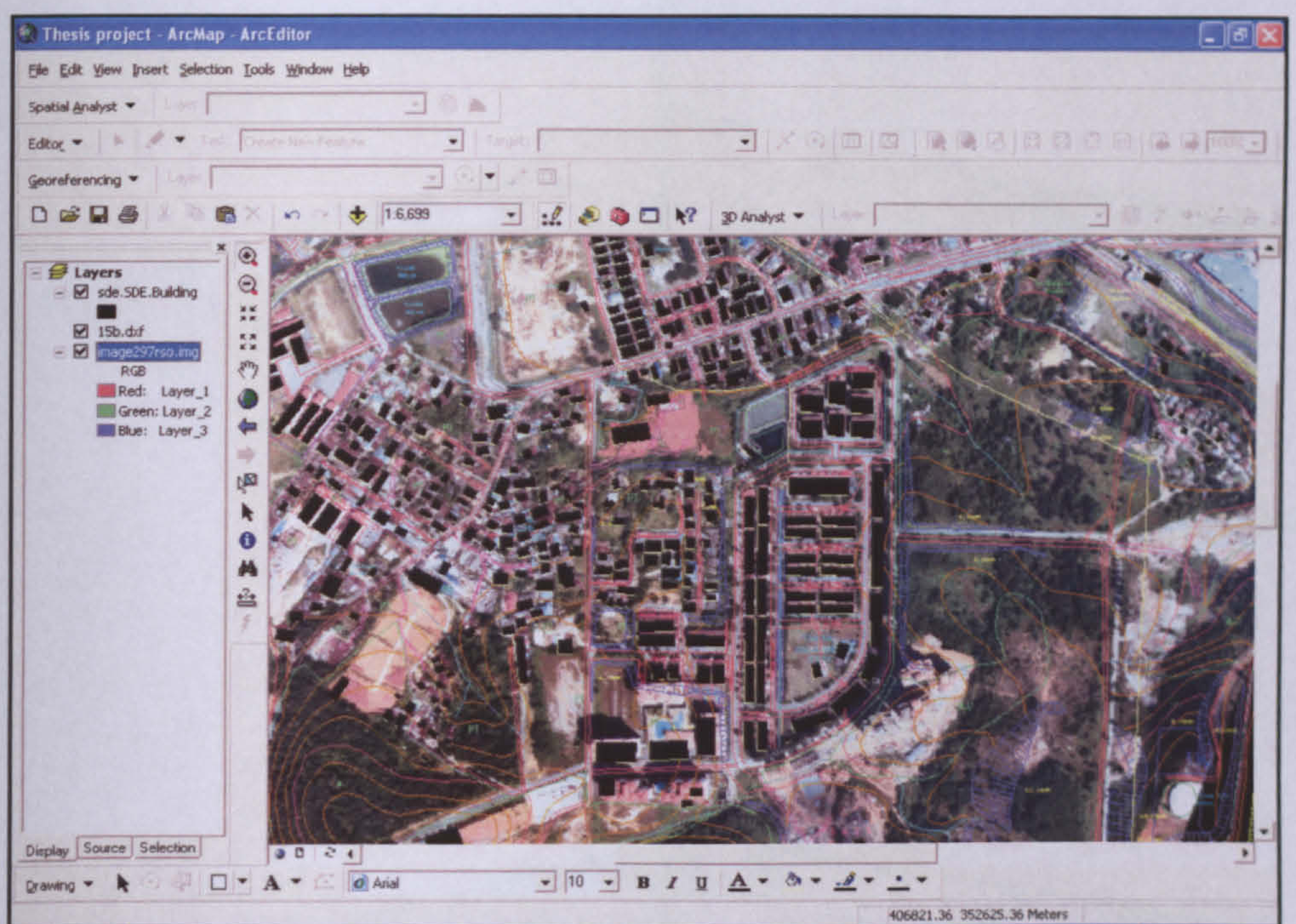


Figure 5.43: A view of raw data (air survey), processed data (CAD file) and GIS-ready information (building object)

Improved management can be seen to be achieved when the geodatabase has been designed, CAD data were processed to form compatible GIS data and they were edited in the same environment. Figure 5.44 and Figure 5.45 show how improved management of the raw captured survey right up to the stage where GIS-ready data is produced can be achieved. A 'one way drill down' of the datasets allows traceability and searching of historical value for the whole dataset from its raw stage to the development of the GIS ready- information. The facility is provided in the same management in GIS environment with capability to view all of the data in one single view.

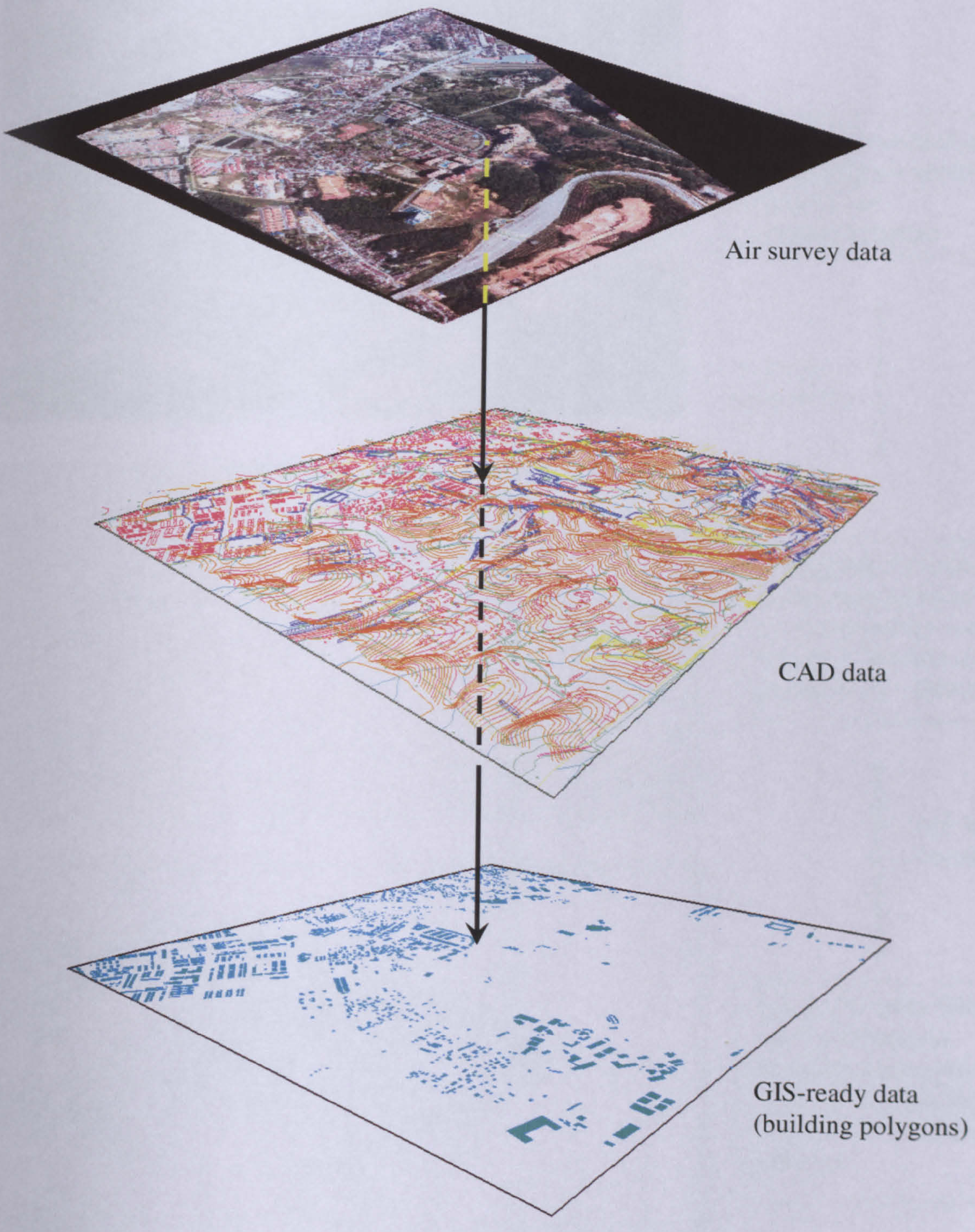


Figure 5.44: Figure shows the 'one-way drill down' that could trace back to the original survey from air survey to CAD (derived from photogrammetry) to a GIS-ready object (building), and vice versa

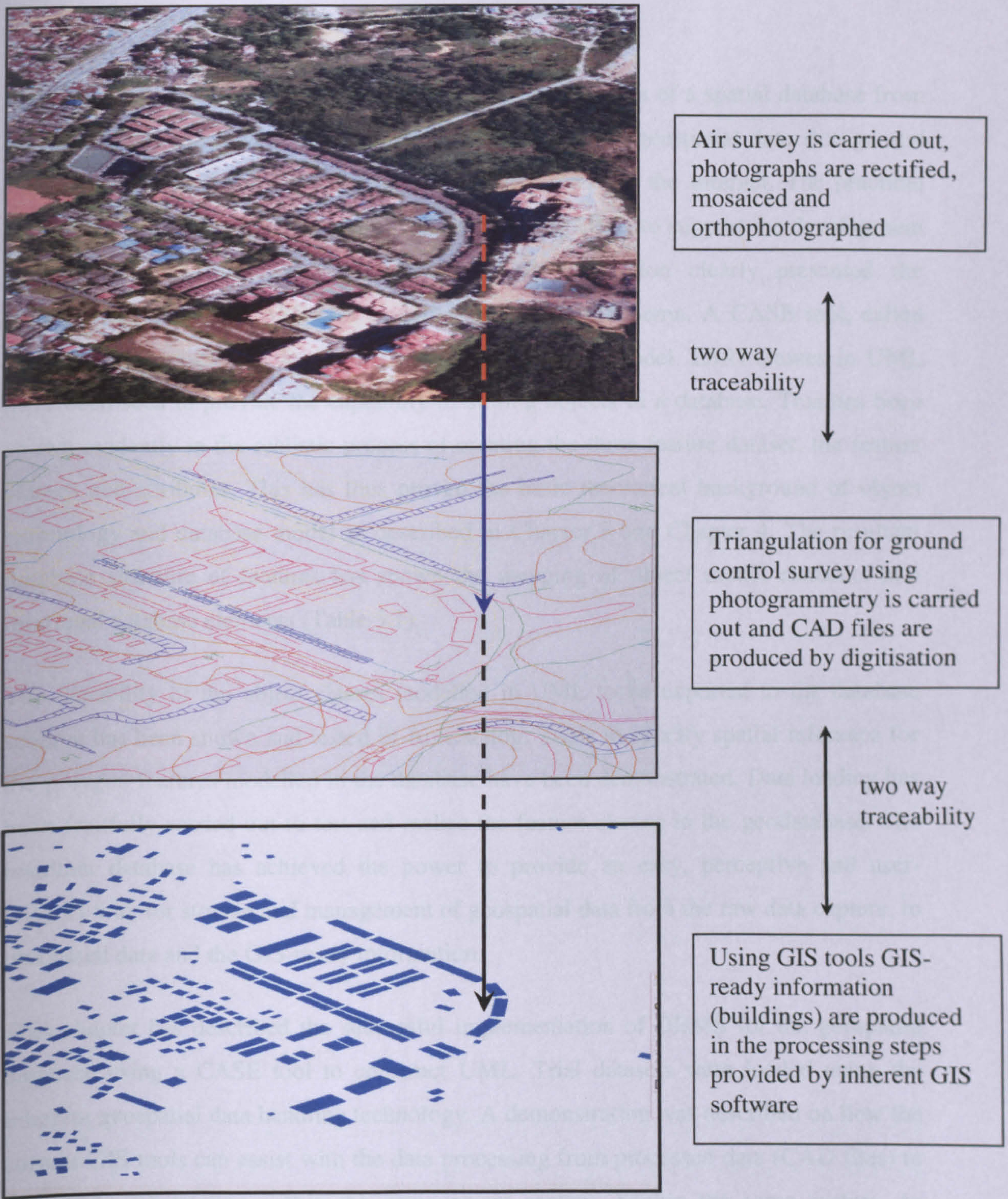


Figure 5.45: Figure shows the close-up of Figure 5.44 of the improved management by GIS giving capability for maintaining complete traceability of the processing steps and the origin of the survey data. Two ways tracking from the raw to GIS-ready information and vice versa can be carried out

5.9 Chapter Summary

This chapter has revealed the successes of the implementation of a spatial database from the conceptual and logical model using object technology, geospatial data storage and management, and sharing and access of geospatial data via the Internet. The practical demonstration of creating a database data model using UML to support the development of a physical database was a success. The demonstration clearly presented the achievement of using CASE tool to derive the database schema. A CASE tool, called Visio has been useful in developing the conceptual data model. ESRI classes in UML have been seen to provide the capability of storing objects in a database. This has been shown evidently in the realistic process of creating the three feature dataset, the feature classes and attributes. This has thus proved the basic theoretical background of object technology and database model as described in Chapter 3 and Chapter 4. The resultant database structure of features has shown the mapping of object model elements into relational database elements (Table 5.1).

The capability of the object classes modelled in UML to be exported to the database schema has been shown and tested in ArcCatalog. Tools to specify spatial reference for the polygon features modelled in the database have been demonstrated. Data loading has been fruitfully carried out to test and realise the feature classes in the geodatabase. The resultant database has achieved the power to provide an easy, perceptive and user-friendly way for storage and management of geospatial data from the raw data capture, to processed data and the GIS-ready information.

This chapter has described the successful implementation of DBMS for the geospatial database using a CASE tool to construct UML. Trial datasets were loaded using the inherent geospatial data handling technology. A demonstration was described on how the current GIS tools can assist with the data processing from processed data (CAD files) to GIS-ready information within the management system. Within the same system, an improved management has been illustrated showing the 'one-way drill down', single interface and integration of all data and information in the management package.

The following chapter gives the prototype delivery of the data across the Internet.

Chapter 6

Delivery of GIS-Ready Information over the Internet

6.1 Introduction

In this chapter, GIS-ready information datasets are tried for online delivery implementation. A development of online delivery of the spatial objects or GIS-ready information via the ArcSDE and ArcIMS, and open source GIS using geospatial standard and technologies is presented. Comparison is made between these two technologies.

6.2 GIS-Ready Information Delivery

Feature dataset *The_GIS-ready_Information* has a substantial number of feature classes such as boundaries, road, building and raster coverage. As defined in Chapter 2 and 4, GIS-ready information should have the capability to be accessed online through a sophisticated graphical interface. GIS-ready information is not fully functional if it cannot be accessed by outside communities. How can we make this data sharable and easily accessible? One way to deliver data is to use the Internet. Coupled with GIS, Internet GIS is an application that uses the Internet or inter-networking systems to facilitate the access and dissemination of geographic information and spatial analysis knowledge (Peng and Tsou, 2003). The following sub-sections describe two ways of developing Internet GIS: the use of current vendor specific software technology (ArcIMS) and Open Source Technology.

6.2.1 Delivery Using ArcIMS

To build an application using ArcIMS, ArcIMS Author was opened to add in data from a network. In this case the network was connected to ArcSDE server, where most features that have been edited were stored and managed. Connection was tested and all the data available in the server was displayed in the Catalogue window (Figure 6.1). This window displayed all the data that can participate in the design of a GIS website (Figure 6.2). Apart from adding data, ArcIMS Author provides tools such as properties display, scale setting, labelling, rendering and setting geo-coding capabilities.

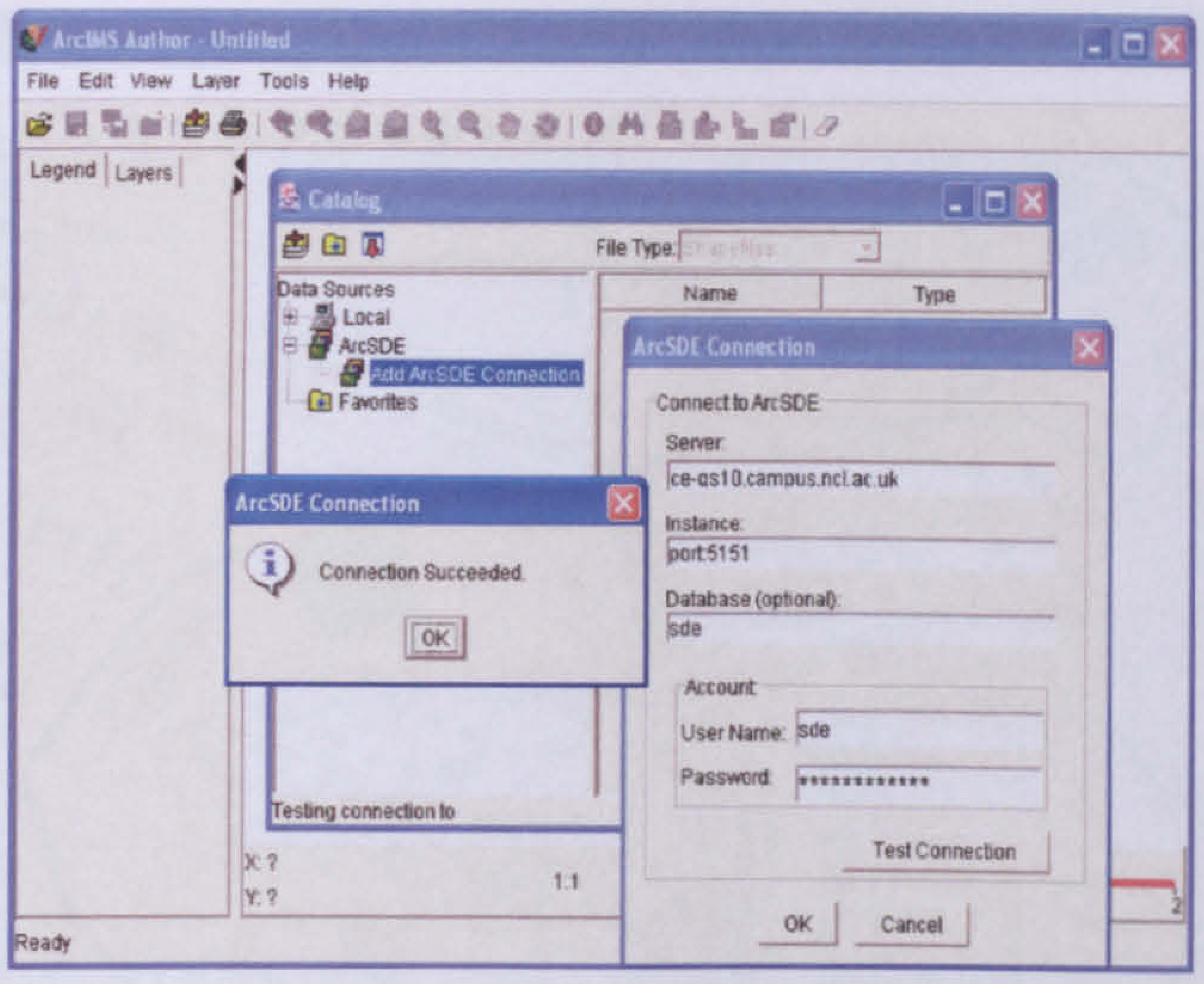


Figure 6.1: Connection to ArcSDE server is tested to load data

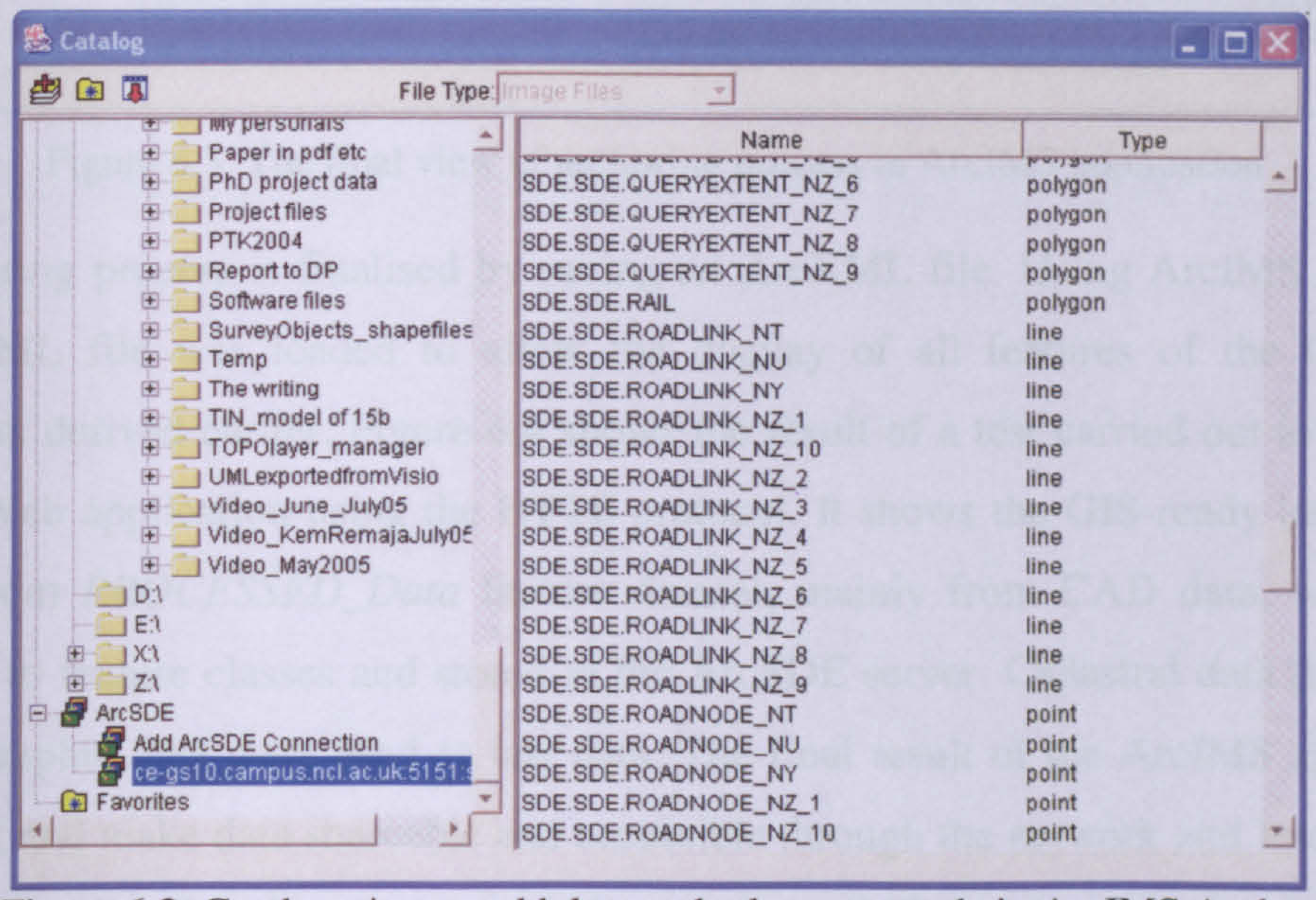


Figure 6.2: Catalog view to add data and other map tools in ArcIMS Author

Figure 6.3 illustrates the view after authoring was completed. In this view, data can be customised and edited to suit the map view set for the Internet GIS. This data was produced and edited as GIS-ready information as in Figure 5.42 in the last chapter.

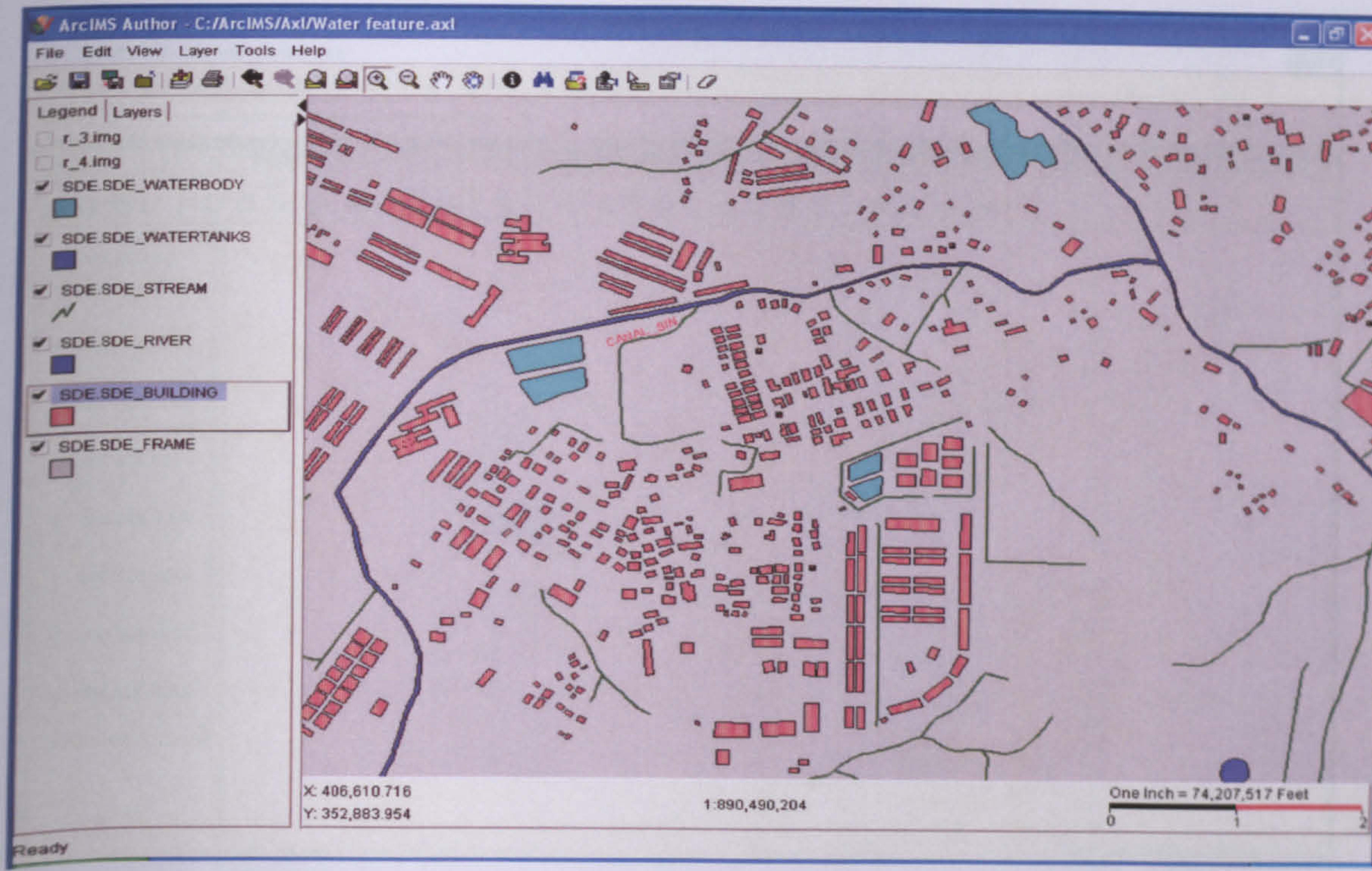


Figure 6.3: The final view of authoring process in ArcIMS application

The authoring process is finalised by saving an ArcXML file. Using ArcIMS Designer, the ArcXML file was loaded to allow the display of all features of the GIS-ready information derived earlier. Figure 6.4 shows the result of a test carried out to create an ArcIMS Web application using the HTTP protocol. It shows the GIS-ready information derived from *PROCESSED_Data* feature dataset, mainly from CAD data, which was converted to feature classes and stored in the ArcSDE server. Cadastral data (boundary) and topographic data were used as test data. The final result of the ArcIMS application was to test and make data shareable and accessible through the network and Internet. The design of a well-tailored corporate website using ArcIMS can be carried out but is not intended in this research. ArcIMS is capable of providing the geospatial data provider with an integrated approach for creating and maintaining geography-based Web sites. A test was made to access the application from different computers and it was successful enabling multiple users of the ArcIMS application.

Figure 6.4: ArcIMS window showing cadastral and topographic feature via ArcSDE

6.2.2 Delivery Using Open Geospatial Standards, Protocols and Technologies

This section concentrates on the use of open source technology comprising existing geospatial interoperable standards, protocols and technologies to disseminate geospatial data from multiple heterogeneous sources. OGC standards and specifications (GML, WFS, WMS), HTTP, XML, Common Gateway Interface (CGI), Spatial Relational Database (SRDB), are some of the available freeware. The development and implementation of Internet GIS using open source technology emphasizes the access, delivery and use of geospatial data over the Web. The development and test was being carried out and written in a research paper in co-operation with Chunithipaisan et al., (2003). It is described briefly in the next subsection.

The methodologies for the application use a 3-tier web service architecture, namely a client, middleware and server.

6.2.2.1 Server

The server was based on OGC Web Feature Server (WFS) and Web Map Server (WMS) specifications. This specification allows the serving of raster (WMS) and vector (WFS) geospatial data from multiple sources via the Web. WMS produces the requested geospatial data in raster format (e.g. GIF, JPEG, PNG), so is called the image server, whilst WFS produces vector data, typically in GML and so is called the GML server.

For the image server, a free map server called Minnesota MapServer was utilised to serve raster data. Minnesota MapServer is an open source development environment created by the University Of Minnesota, USA for constructing spatially enabled Internet-web applications (Neteler and Mitasova, 2002). This map server supports the ESRI Shapefile and image data format. Figure 6.5 illustrates the use of the Shapefile format to serve images of Kuala Lumpur (KL), Malaysia on the Web. CGI was used to request data, however it could be set to request in a manner compatible to OGC WMS specification.

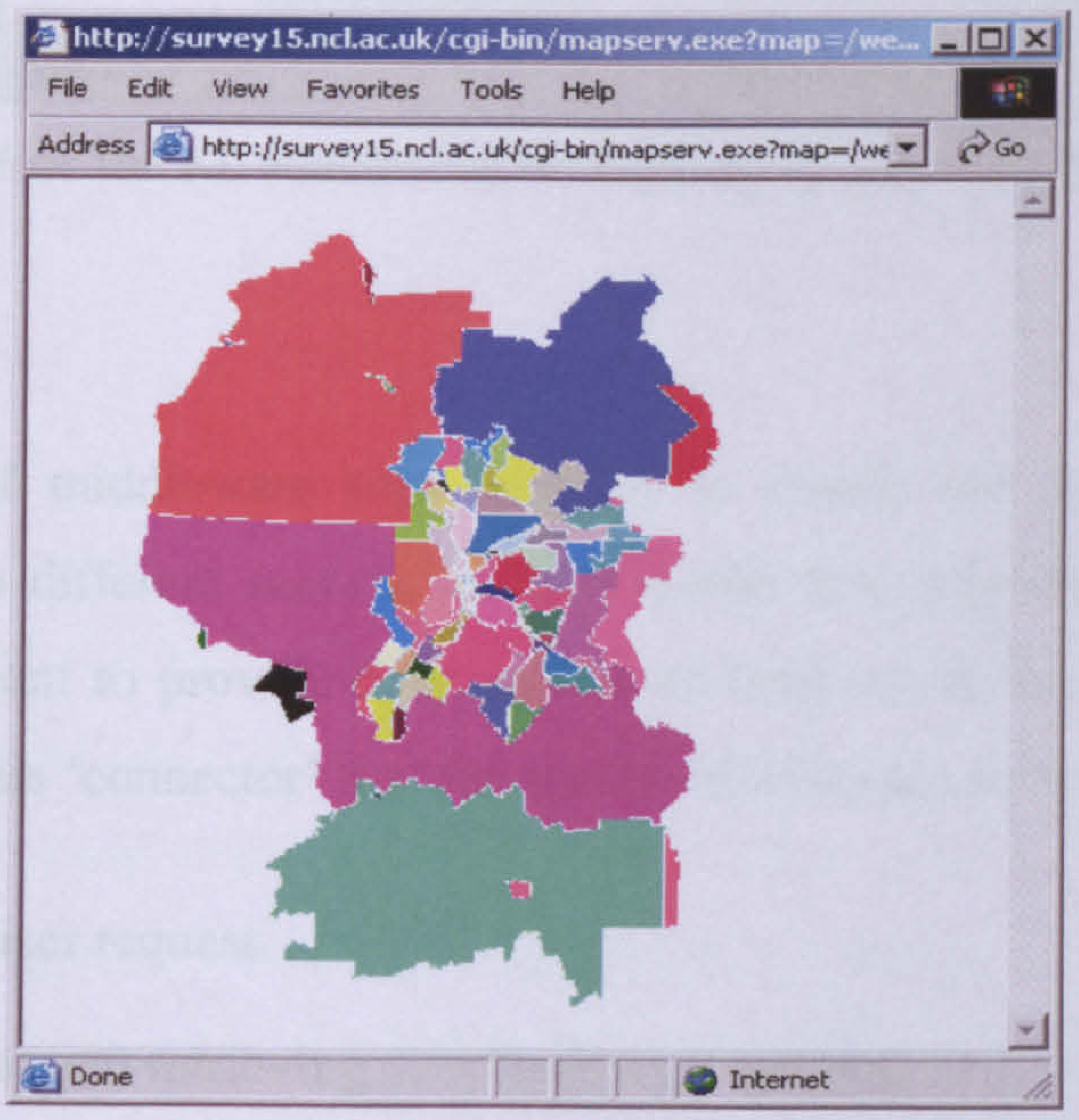


Figure 6.5: Kuala Lumpur administrative map using the Web MapServer

The GML server was developed to provide vector data in GML format. GML technology provides transport of geographic information by XML encoding. The use of GML thus enables one to facilitate geospatial data interoperability. The OGC WFS specification was implemented to develop the GML server. It was developed using Java servlet technology. The two main WFS operations utilised were *GetCapabilites* and *GetFeature*. The formats supported by the GML server include data stored in the Spatial Relational Database (SRDB) and Shapefile. Figure 6.6 shows the created GML from the server which was an administrative boundary feature of Kuala Lumpur similar geographically to the image map in Figure 6.5.

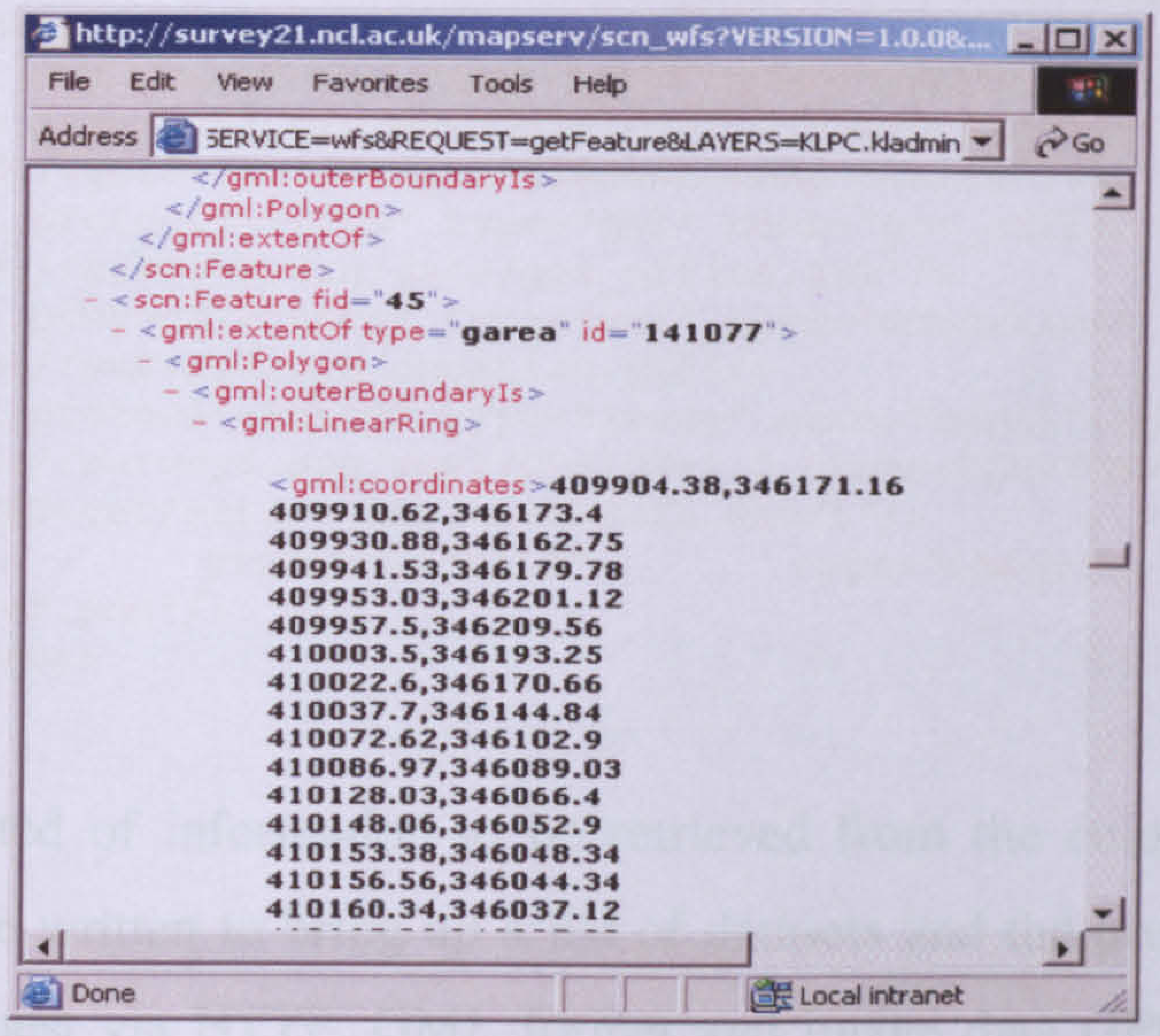


Figure 6.6: The KL administrative boundary feature in GML format

6.2.2.2 Middleware

The development of middleware was intended to enable the application to retrieve geospatial data from different servers. A Java servlet that supported CGI was utilised. This servlet was meant to provide communication between server and client via URL. The capabilities of this 'connector' that the servlet middleware performed include:

- getting the user request.
- connecting to and retrieving data from map servers.
- combining the requested data.
- delivering the resultant data back to the client application.

The data type requested by the middleware can be raster or vector. The procedure to retrieve data can be from HTTP protocol or FILE protocol. HTTP was used for datasets located on a different server and FILE protocol was used for datasets that are stored in the same server as the middleware connector. For data retrieving from different servers and combining that data by sending only one request, a project request was made. The project request was encoded using XML and a sample is illustrated as follows.

```
<project xmlns="http://geomatics.ncl.ac.uk/webgis/scn/">
  <box minx="402250" miny="345972.47" maxx="408970.0" maxy="352473.62"/>
  <rwo_dataset>
    <dataset protocol="http" type="gml" name="hospital_clinic" address=
       "http://survey1.org/wfs?VERSION=1.0.0&amp;SERVICE=wfs&amp;
       REQUEST=getFeature&amp;LAYERS=KL.hosp_clinic"/>
     <dataset protocol="http" type="gml" name="gov_office" address=
       "http://survey2.org/data/gov_office.gml"/>
     <dataset protocol="http" type="gml" name="education" address=
       "http://survey3.org/ducation.gml"/>
     <dataset protocol="http" type="image" name="canal" address=
       "http://survey4.org/wms?VERSION=1.0.0&amp;SERVICE=wfs&amp;
       REQUEST=getFeature&amp;LAYERS=KL.canel"/>
     <dataset         protocol="file"         type="image"         name="river"
address="river.png"/>
  </rwo_dataset>
</project>
```

The project consisted of information to be retrieved from the connector. The project request code can be written to bring up a list of datasets and the project dataset can be opened in one request via HTTP. GML format and image data (JPG, GIF, JPEG) are supported for the request.

6.2.2.3 Client

In the client side, a map viewer is required to render the data requested by the servlet. This application is mainly for the use of GIS tools such as rendering, querying, pan and zoom. An applet application, which should be able to run on a web browser, was developed using Java technology. It allows users to open data locally and from different sources remotely. Figure 6.7 shows the application Graphic User Interface (GUI) with some functions of zoom, pan, rendering and query. Figure 6.8 displays the dialog window to open data locally and remotely. Selections of data format are supported including image files, GML and Shapefile. As long as the project request code contains information about the list of datasets, they can be displayed with one request via HTTP.

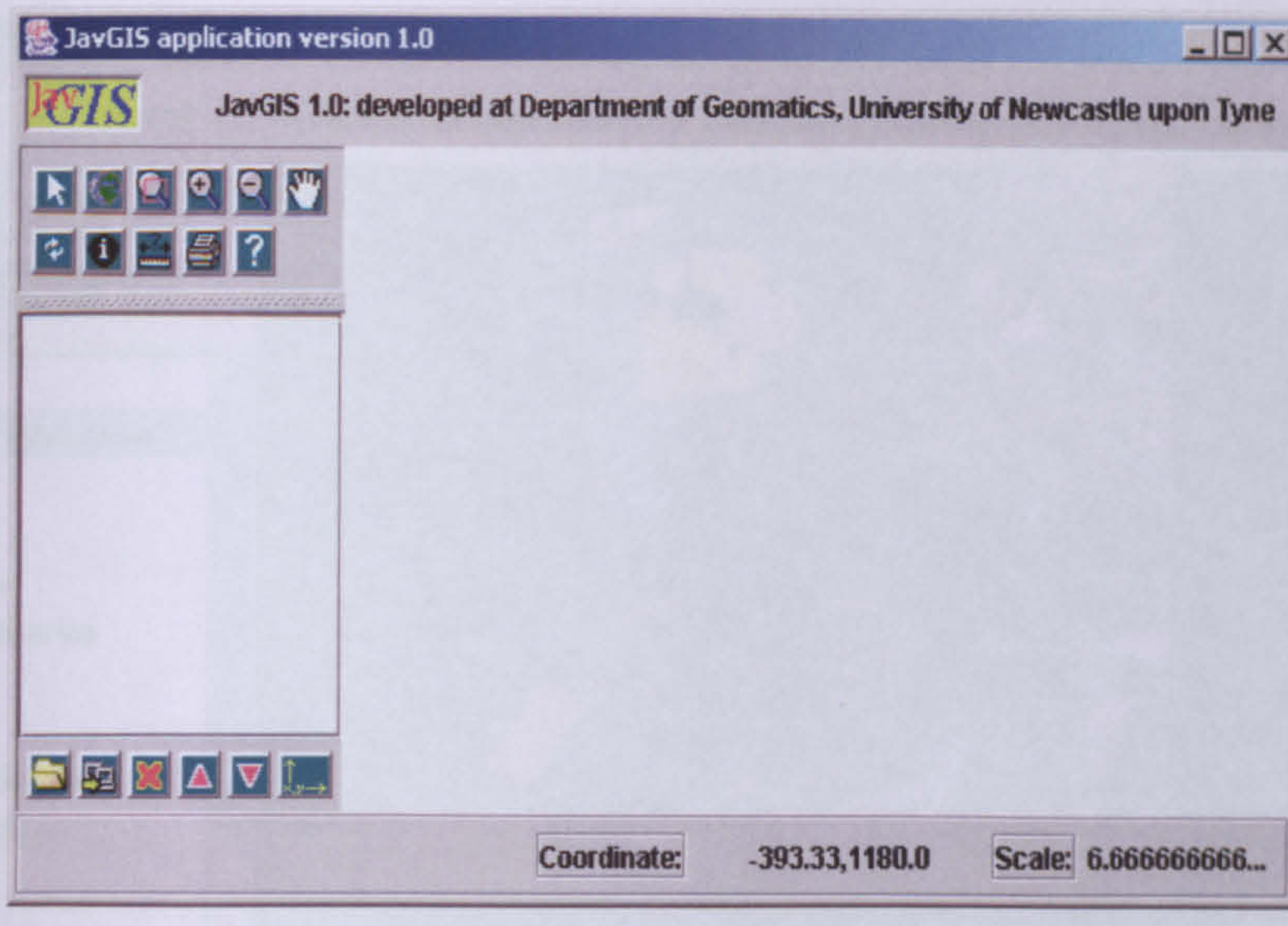


Figure 6.7: The interface for the client map viewer

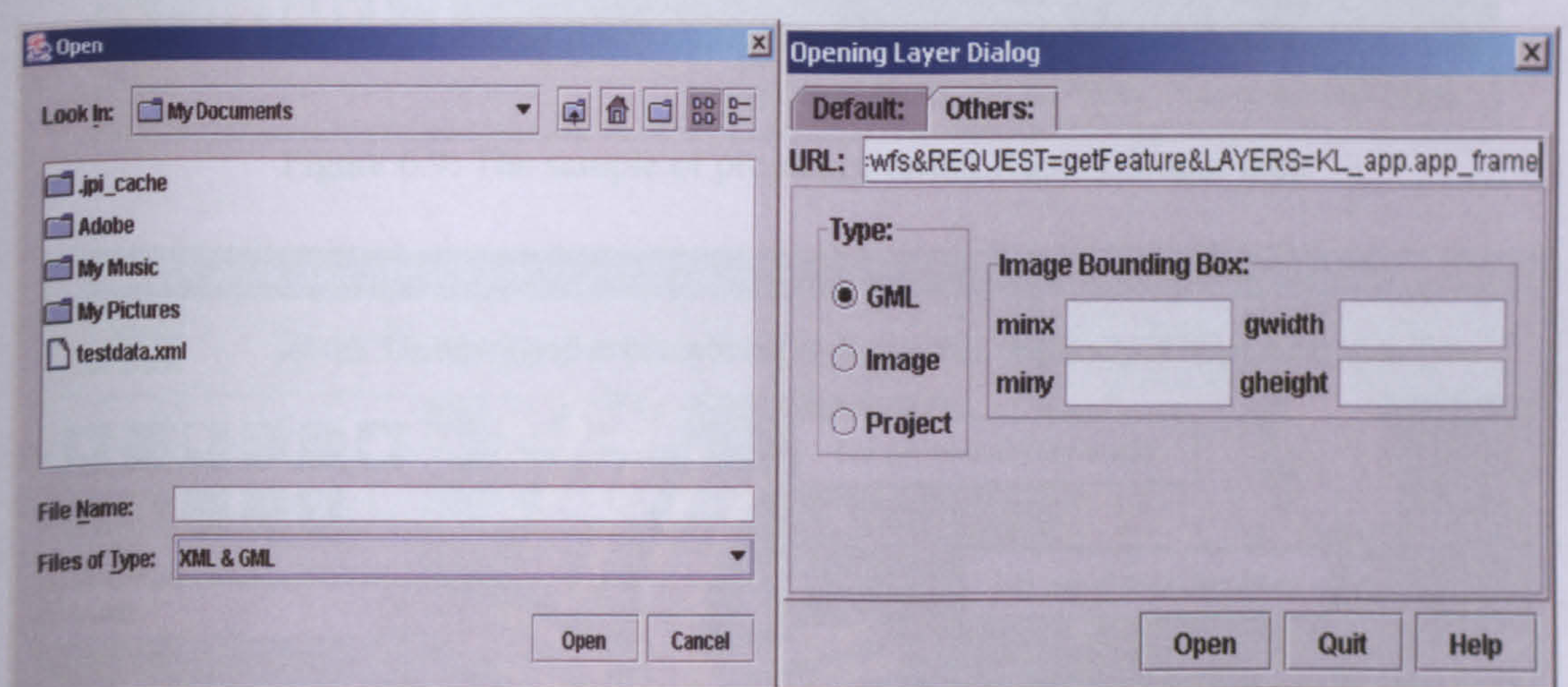


Figure 6.8: Local file and HTTP data open dialog

A sample of data that was classified as feature classes in the feature dataset *The_GIS-READY_Information* was listed (as in the code shown in 6.2.2.2) and opened using the application. Figure 6.9 demonstrates the display of the sample project datasets in the application. Figure 6.10 provides the zoom-in view of the area showing cadastral (lot parcel, lot boundary lines, lot boundary marks), topographic (buildings and roads) and image data in one single view. These datasets were stored in different servers and retrieved via the HTTP protocol. The datasets are the GIS-ready information, processed data as discussed in the earlier part of this chapter.

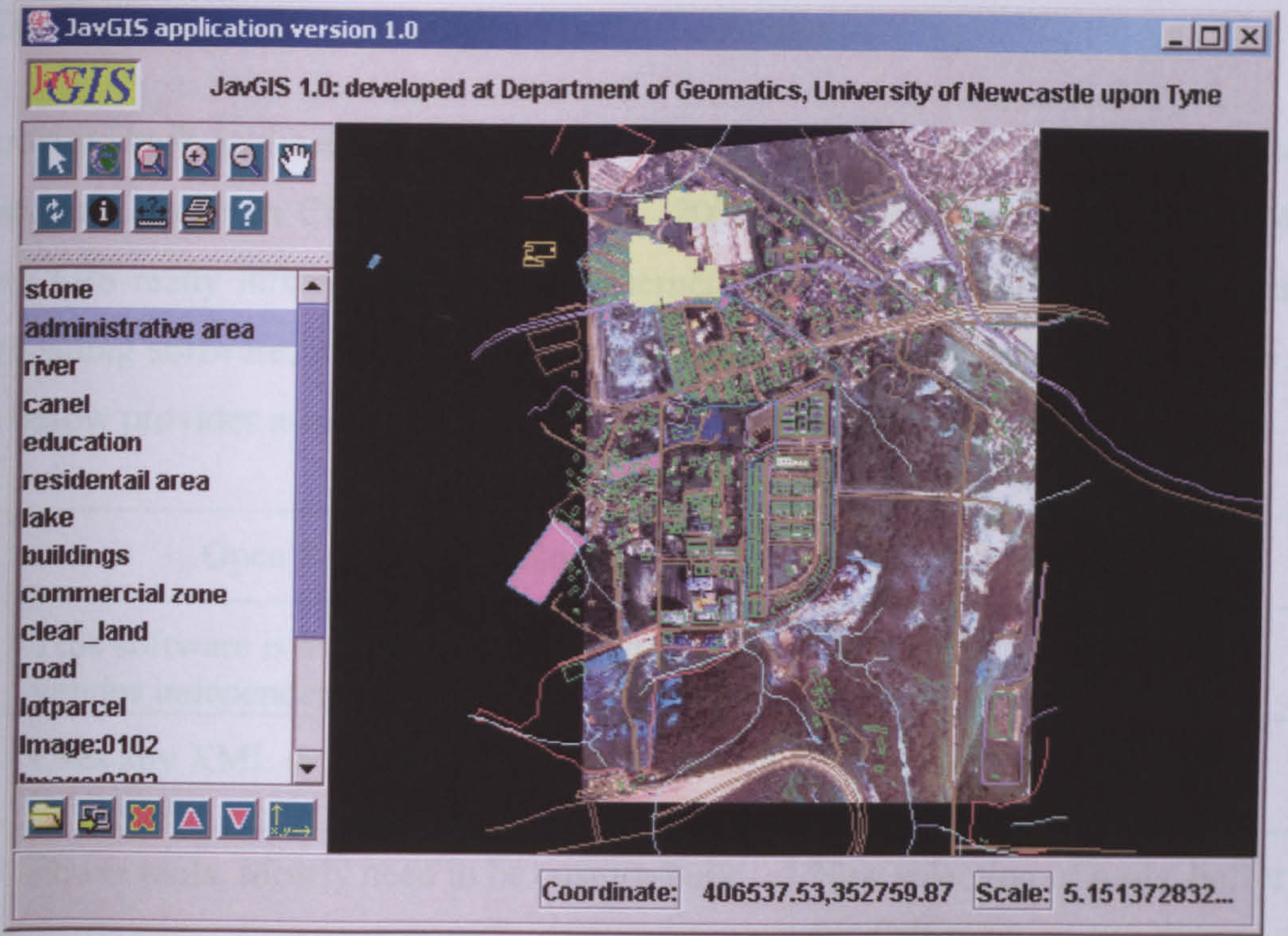


Figure 6.9: The sample of project covering Kuala Lumpur area

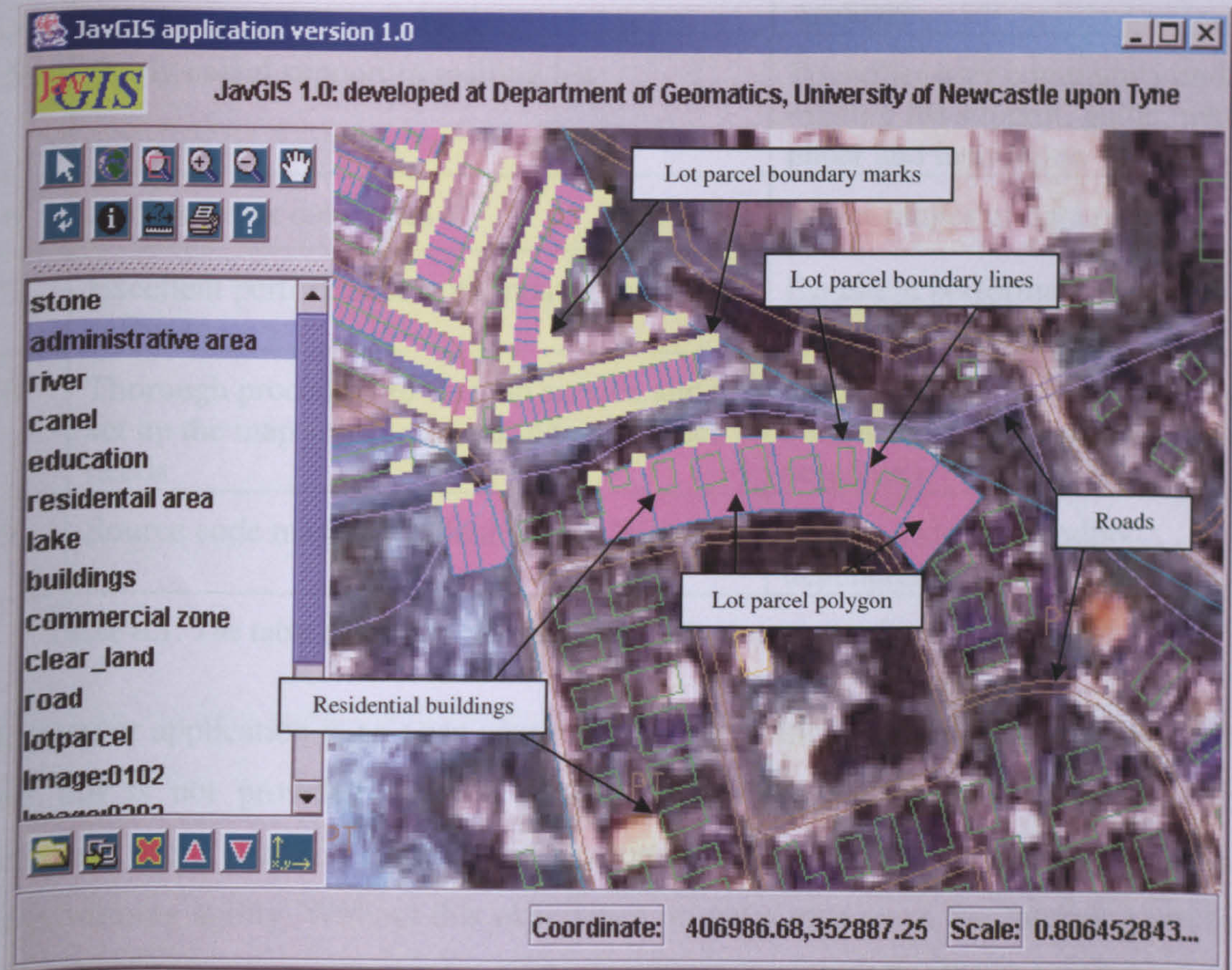


Figure 6.10: A close-up (zoom-in) of an area showing the data in single view

6.3 Contrasting ArcIMS and Open Source Technology

Contrasting both applications is led by the experiences throughout the implementation of the research paper in Chunithipaisan et al. (2003) and development using ArcIMS to deliver GIS-ready information over the Internet. Discussion is put forward concerning downloading software, installation, configuration and the testing of the application. The table below provides a summary of their capabilities and weaknesses.

	Open Source Technology	ArcIMS
1	This software is available freely, open source, vendor independent	Standard and commercial software
2	Uses any XML or other programming language	Using proprietary ArcXML
3	Fewer tools. Mostly need to be custom-built	Nice selection of tools, buffering for example
4	Database connectivity to be set	Easy with ArcGIS geodatabase or ArcSDE
5	Fairly useful support in mailing list	Friendlier user community and mailing list support, guide, white paper and instruction
6	Access fewer data types	Wider ranges of data types
7	Excellent performance with most computing software	Excellent performance with its own related software
8	Thorough procedure to do map service and to set up the map server, computing experience needs	Setting map service is easier with complete guide on screen (digital helpdesk)
9	Source code may be available in the Web	Source code are vendor-dependent and unavailable

Table 6.1: The table shows the advantages and disadvantages of using ArcIMS and OST

Mapserver application is an open source technology which can easily be obtained and is free but is not provided with an instruction manual. Installation and configuration therefore needs personal experience in computing and Internet application, as well as programming ability. Without this experience, to get a map up to the Internet costs time because you have to search the Web for documentation. ArcIMS, as 'out of box' software provides a friendlier environment by using an instruction manual and a localised digital 'helpdesk' which you can browse while installing. If the documentation is followed closely, with experience in computing, installation time can be very short.

It can be pointed out that the scopes of what ArcIMS and Mapserver can present are different. ArcIMS is superior to Mapserver in that it has several components designed to present GIS-ready information online that Mapserver is incapable to address. These include HTML and applets based viewer for the client that allows GIS-ready information delivery more impressive. Additionally, ArcIMS has broader functionality: components on the server including the spatial servers, manager components and application server. The spatial server renders the images, handles spatial queries, extracts features, etc. The manager components provide the facility to build map services, create Web sites using the map services and publish the sites. The application server allows distribution of data over multiple machines. In addition, ArcIMS is provided with inbuilt Website application to do map services and designing Web, such as ColdFusion that provides an in-hand connector between the web server and the application server. MapServer only uses a spatial server from the user machine and designing a Website needs licensed software. As in this research, we use Java applications to design the user interface to distribute information for the Mapserver application.

ArcIMS allows the distribution of applications across multiple computers, but every computer needs to be configured with a map viewer application of proprietary type. In addition, ArcIMS has native integration with ESRI's Geography Network while Mapserver has to be configured to work within the Geography Network or other map data providers. Raster data is supported in ArcIMS with better storing in RDBMS through ArcSDE. Mapserver relies on DBMS with a configured connector to the application.

However, ArcIMS has limited resources available because they are vendor specific and need purchases for extension or other services. ArcIMS has restricted use for some operating systems such as Linux. Mapserver can be shared with different kinds of operating systems. Another disadvantage of ArcIMS is that source code is not available whereas OST can be shared within the community. Cost is high for using ArcIMS while Mapserver is free with the open concept GIS community. Finally, ArcIMS may be useful out of box software but there are 'bugs' here and there that are still unanswered and so it is undependable. OST can be easy and problems can be solved with a high level of experience in computing and programming.

6.4 Chapter Summary

Sharing and access of GIS-ready information over the Internet has been significant for the data to be serviceable and practically usable. Current technology methods using ArcIMS and open source technology were demonstrated in this chapter. The concept and development of existing ESRI Internet Mapping software technology and application-based WebGIS were successfully tried and tested. The presentation of the resultant GIS-ready information over the Web has shown the success of making the geospatial data sharable and accessible. Both applications have advantages and weaknesses depending on the type of map services and resources available.

The following chapter gives the conclusion, achievement and future works.

Chapter 7

Conclusion, Achievements and Future Work

7.1 Introduction

In this chapter, a summary is drawn upon on the work presented earlier in the chapters to achieve the research aim and objectives. The main aim of the research was to investigate the use of geospatial data handling technology to facilitate and improve the management of the geospatial data from capture right through to the production of GIS-ready information. To achieve this main aim, investigation, review and development of GIS application were made regarding geospatial technologies, object representation, issues concerning database management technology, and processing steps, data sharing and delivery.

In this research the key aims and objectives discussed in the introductory chapter have been accomplished. The followings are pointed out as the accomplishment;

- Discovery of the current systems for managing survey datasets from their raw captured geospatial data to their output stage as GIS-ready information. Review of the current technologies of geospatial data capture and how they are being managed was made in OS, USGS and JUPEM.
- Comprehensive experience and in-depth understanding of current GIS technologies and techniques for managing geospatial data right through to the production of GIS-ready information.
- Potential of modelling all data and information, spatial and non-spatial, within raw captured data, processed data and GIS-ready information in a national survey and mapping department as geospatial objects.
- Implementation of a prototype DBMS according to the specified data model using UML technology and available database design tools.
- Demonstration of the population of trial survey datasets of JUPEM for the implemented DBMS using inherent geospatial data handling technology.

- Development and demonstration of improved management of the survey datasets using the developed GIS capability illustrating 'one-way drill' of all information and data within the raw capture right through to the GIS-ready data stage.
- Manifestation of the use of inherent GIS tools to assist with the data processing of processed data for the production of GIS-ready information.
- Implementation of vendor specific GIS and Open Source Technology Internet solution for the delivery of GIS-ready information to the clients.
- For the accomplishment of the ninth and tenth objectives, these are described in this chapter in the appropriate section.

In this final chapter, a digest of the chapters' summaries gives a synopsis of the core research technology and methodology utilised. Achievement reached throughout the course of this work is highlighted and further discussion relating to the research matter is given. This is followed by an abridgement of the conclusions established at the end of the research and recommendations for future direction.

7.2 Thesis Summary

The writing of the thesis began in the introductory chapter with an overview and background and motivations of the research work. The aim and objectives have been presented, followed by research methodology and an outline of the thesis.

Chapter Two has described the motivations that inspired the research, review of what are geospatial data and the technologies behind them. Knowledge and current development of GIS were discussed, followed by a review and investigation of the handling of geospatial datasets in three geospatial data organisations. A thorough discussion of the practice in JUPEM that drove this research was made. The literature review chapter has found the need for raw captured data, processed data and GIS-ready data to be managed within one system of management by modelling them as objects with spatial entities and structuring them in a relational model. Three entities of geospatial data capture, processed data and GIS-ready information, and all data and information existing in them have properties and capability to be modelled as objects. These three features are suggested to be harmonised into one single portal which is manageable for GIS users.

GIS data delivery and access was also discussed in this chapter. GIS-ready data are not resourceful and comprehensively used by users if they are to sit in the organisation without access to other government departments and the general public.

Chapter Three has critically examined the theory of objects in geospatial data handling technology. Objects allow world entities be represented incorporating operation and attributes. The basic concept of objects as GIS representation was discussed. Eventually, related database system and UML technology which served the main ingredient of GIS functionality and application were incorporated. Geographic database as a smart spatial database was briefly discussed in this chapter describing the entire concept and presumption of the object technology and relational database model. Intentionally, this technology served the main part in the implementation and prototype development.

Chapter Four has discussed the current survey practices and workflow of JUPEM which linked its nature to the aspects of geography and spatial modelling. The notion and theory of data models and the way to model geospatial data especially survey data has been described. Affiliation of the described data models with the survey datasets in a national survey and mapping department, JUPEM, was discussed. This is the crucial part of the discussion because the survey datasets have demonstrated significant values, connections to spatiality, capability and properties that enable them to be modelled as geospatial object, for management using GIS techniques and technology. Data management approaches to store and manage geospatial data were examined conceptually. DBMS with GIS functionality has inspired efficient geospatial data management. Geodatabase, a suitable DBMS-based data storage and management method was elaborated in this chapter to serve ideas and background for the implementation of the geospatial management. From raw captured to GIS-ready stage, all the data and information can be modelled using elements of the geographic database. Feature dataset, feature class, object class, relationship class, raster catalog and TIN dataset, as elements of geodatabase are powerful and intuitive concepts for the management of geospatial data. Aspects of sharing and access of GIS-ready information were investigated using vendor specific application and Open Source Technology via Internet solution.

Chapter Five has presented implementations and development of prototype regarding geodatabase DBMS, geospatial data storage, processing and management application using the current geospatial data handling technology. These implementations and developments were meant to support the thoughts brought up in the previous chapter. Through the practical process, several concepts to process, store and manage huge amount of geospatial data from data capture through the provision of GIS-ready information have been demonstrated, tested and proved. Data processing, data loading and improved management within the same system of geospatial management has been successfully achieved.

Chapter Six has demonstrated applications for delivery of GIS-ready information via the Internet. Delivery applications were carried out, tested and proved using inherent vendor specific application and Open Source Technology combined with GML. Discussion was made to differentiate these two applications.

7.3 Achievements and Discussion

Achievements of the research are listed as the followings:

1. A review of the current survey technologies and current methods for the management of survey datasets in government survey and mapping departments.
2. Achieved an in-depth understanding of the latest GIS technologies, techniques and standards as it applies to the managing of geospatial data capture to the production of GIS-ready information.
3. Geospatial object concept, within a geospatial database, has been used to model all the data and information from its raw capture stage, through its processing stages, right through to being GIS-ready information. The use of an object-relational model to design a geospatial database for storing and managing the datasets has also been presented.
4. Successfully implemented a prototype DBMS for the geospatial database using a UML to construct schema for database design.

5. Fruitfully demonstrated the implementation of DBMS using trial datasets from a national survey and mapping department, JUPEM. This was carried out using the current inbuilt geospatial data handling technology, ArcGIS tools.
6. Successfully demonstrated benefits to the survey and mapping organisation, of implementing a GIS-based application for the management of its packages of information. This includes the 'one-way drill down', single interface and integration of all data and information in the package.
7. A demonstration of inherent GIS tools that can assist with the data processing from processed data (CAD files) to GIS-ready information within the management system.
8. Successfully implemented two methods for the delivery of GIS-ready data to clients: vendor specific GIS (that is ArcIMS) versus an open source technology solution (GML, open source Mapserver software and XML).
9. The overarching motivation of understanding a very high level of all aspects of GIS techniques and technologies to cope with geospatial data infrastructure needs.
10. The achievement of high standards of methodology and experience to develop a GIS project to equalise the current development of GIS techniques and technology within developing countries to that of the developed countries.

The outcome of the review of the current survey technologies and current methods for the management of survey datasets in government survey departments revealed that there is no real management system that organises survey datasets from raw capture and processed data for the production of GIS-ready information. This issue is thus not receiving much concern for the full life cycle of the survey datasets currently. In addition, there was no research found that deal with this type of management system.

The latest GIS technologies, techniques, models and standards were examined and discussed. Chapter 3 and 4 has described these and they were implemented in Chapter 5 and 6. Using the standard of UML notation and visual modelling tool an object-relational geospatial database within DBMS was produced to the specified data model.

The research has demonstrated that it is possible to model all the data and information of the raw survey, processed survey and the GIS-ready information as geospatial objects in a geospatial database. Datasets were modelled as objects and objects classes within an object-relational geospatial database. A prototype DBMS and population of datasets were successfully implemented. Inherent GIS tools were used to efficiently create and develop of GIS-ready information across the system.

The 'one-way drill down' method to achieve efficient query, management and ease of processing steps is an intelligent system of digging dataset, finding historical value, eliminating redundant development of datasets and GIS-ready information. The visualisation of all raw data, processed data and GIS-ready information in one view management application system allows 'bird-eye view' of all the real world objects of surveys implementation and GIS-ready production in an organisation. This solution gives surveyors an expressive method to minimise imaginative or site exploration of all survey datasets collected in the earlier time such as field survey collection. Two-way traceability has shown the impact to sound survey data management which would allow the traceability of sources data for GIS-ready information enhancement and in the same way would allow the tracking of up-to date GIS-ready information to avoid multiple resurvey of original source data.

Contrasting the two Internet GISs implemented, advantages and disadvantages were tabled in Chapter 6. Dependent on the user specification, both systems deliver the GIS-ready information but the speed, support, ease of use, difficulty to receive data and cost are some of the matters that to be taken care off before embarking into an appropriate Internet GIS. Vendor specific, ArcIMS has shortfalls: the user has to download and configure viewer software in their machine. If the computer does not comply with specification set upon, viewing would be painstaking. Open Source Technology is freeware and with an open environment, many computing expert users can communicate and solve problems by interactions over the Internet.

Implementing Open Source Technology (OST) in this research is significant to anticipate for the realisation of geospatial data delivery in JUPEM. As the matter of fact, the Government of Malaysia through the Malaysian Administration Modernisation and

Management Planning Unit (MAMPU) has decided to encourage the use of Open Source Software (OSS) in the Malaysian Public Sector (MAMPU 2006). The Prime Minister Department is given the responsibility to implement this initiative. This initiative gives impact to JUPEM to use OST to deliver GIS-ready information. This research has thus benefit JUPEM to promote and foresee the possibility of using OST as the standard for the delivery of the raw, processed and GIS-ready information within the Intranet or Internet.

For survey data and all information which are spatial and with coordinates, it is possible to look into the use of the current open source technology, PostGIS. It supports geographic objects for the use of PostgreSQL object-relational database. PostgreSQL is a powerful, open source object-relational database system created by PostgreSQL Inc. The capability of PostGIS includes enabling spatial extension, the same as ESRI's ArcSDE or Oracle's Spatial Extension. PostgreSQL plus PosGIS (with spatially enabled geodatabase) has been a very reliable and powerful solution, and has already replaced commercial top-end solutions (Cavallini 2004).

PostGIS has been developed by Refractions Research as a research project in open source spatial database technology (PostGIS 2006). PostGIS can be very useful to be used in JUPEM because of its spatially-enabled capability and the technology is always enhanced, for example to date with the technology for data loading and dumping, user interface tools for direct data access and manipulation, and support for advanced topologies at the server side, such as coverages, networks, and surfaces.

Malaysia as a developing country needs to use OST for the development of systems and ICT based infrastructure in the government department as well as the private sector because:

- Freeware easily available.
- Relatively limited financial resources in the government sector thus lead to the adoption of low-cost solutions.

- Availability of IT staff who are nowadays able to exploit effectively the programming side of computing science, customise and further develop existing or new tools.
- By using the technology, resources are made to remain in JUPEM or Malaysia as a whole thus may assist building national software industry.
- By using vendor specific technology, money invested in the proprietary software goes mainly abroad.

It is very likely that the experience and understanding gained has provided fundamental ideas such that the author can debate the requirement of GIS and management of geospatial data at an equal level with other government department and GIS software contractors. The experience of contrasting vendor specific and Open Source Technology in geospatial data delivery has provided material which can be valuable when contracting GIS development to private companies. Most private IT and GIS specific companies are originated from parent companies overseas, particularly in the developed countries.

De Man and Den Toorn (2002) suggest that information sharing and operational planning and management would allow cultural desirability of GIS. They also stipulated that feasibility of the actual introduction of GIS would be governed by the conditions largely specific to the recipient organisation. Thus the adoption and sustained use of GIS depend on the combined influence of desirability and feasibility. So the resulting research project that suggested the sharing of information in the department and the feasibility (practical and intuitive) of the system within GIS environment would enable adoption and sustained use of GIS in the national survey and mapping department specifically JUPEM.

Further issues and discussions which are very much relevant and associated to the research ideas are related to metadata, accuracy, updating and versioning. These will be discussed and highlighted herewith and after.

Metadata is the data about the data. Metadata has a lot of meaning to survey data. There are huge amount of metadata when aspect of surveying is discussed. A certified plan from cadastral survey displays some metadata surrounding the graphic plan of the land parcel. Plan number, surveyor, date, coordinate system, state, district, type of survey, class of survey, northing sign, closure, scale bar are some of the metadata that are presented on the title block and notes on a certified survey plan. Surveyor's boundary can be known accuracy in terms of closure, but it is not disclosed specifically. Likewise coordinate system, an important metadata, is not also disclosed. Aspect of metadata in survey data, processed and GIS-ready information for GIS application is significant because ultimately this type of metadata can lead toward increasing usability and interoperability of survey data. GIS mapped survey features should contain explicit and easy-to-understand metadata. The public can be reasonably assumed to be protected if they are informed about the locational accuracy, currency, and method of compilation (lineage) of the data in a GIS. JUPEM, which incorporating large number of survey data, needs a standards-based system such as of this research to implement metadata, ultimately allows promotion of more seamless ways to integrate data from disparate sources, from the 2-d traditional surveying data all the way to high-end GIS data without losing the fidelity, meaning and context of the data.

Using GIS application such as ArcGIS enables metadata be stored for survey data from raw survey and processed data. GIS-ready information can be rich in metadata if incorporated all the metadata from raw to processed data in the application. Users of GIS-ready information can create metadata from basically any data managed using the module ArcCatalog in ArcGIS. The metadata vehicle in ArcCatalog is called a stylesheet. ArcCatalog can actually create the metadata for the user if it is not documented yet. There are five useful functionalities in ArcCatalog software icons: Edit metadata, Metadata properties, Create/Update metadata, Import metadata and Export metadata.

In many instances, the metadata in many survey datasets and processed datasets may not exist or does exist in poor quality. Metadata can be very difficult to find even it exist in survey datasets. They are sometime dynamic, and with continuous updates but are not reflected in the metadata (such as collection date, method of collection and collection accuracy) would undermine the data. On carrying out large scale survey or emergency response land survey (such land slide or Tsunamis), negative or false positive on the topographic terrain information can lead to loss of life and/or property or waste of laborious survey.

In the application of the management of survey from raw to GIS ready state, metadata serves surveyor and GIS staff to make informed decisions about usefulness of data (for example avoidance or second survey of the outdated dataset). Survey datasets with properly documented metadata, can be evaluated and prioritised for use. From the surveying world to the GIS world, the context of works and processes becomes more complicated. So, surveyors, apart from understanding locational criteria, need to use and document metadata in their product. By doing this, they are able to convey their metadata along with their data into GIS systems and Land Information System, which are eventually collaborative and are able to provide continuous quality improvements. However, GIS maps that have been adjusted (rubbersheeted) to create consistent, coherent display maps should retain the original mapped coordinates as feature attributes, as well as metadata describing the transformation adjustments that were made.

In the case of accuracy, GIS data are not accurate when compared to what surveyors observe and record in the field. Issue on quality and coverage of the source data makes the GIS-ready information in dilemma. An analysis of the quality and coverage of possible sources is required because a particular GIS feature is often shown on multiple sources of varying quality (De Man 2006). In actual fact, the quality of source materials should be equal to or greater than the desired quality of the converted database; otherwise the source can be rejected and replaced. For example, if maps do not show the required information with appropriate accuracy, precise calculations or aerial mapping may have to be used to provide the data. Therefore it is crucial to identify the quality of sources before defining them in database and before the cost of conversion is estimated and quoted and data conversion begun. It is also important that GIS-ready information should explicitly and accurately refer to the source documents from which they were compiled. Such linkage could be achieved by carrying a source document identifier in the database record of each GIS feature, or linking to scanned images of those source documents.

In the line of implementing the management for survey data for GIS-ready information, accuracy should therefore be sought not only in term of data precision but also accuracy of the datasets when converted and structured in the database. Accuracy may mean temporality, currency, loss of attributes and error in digital mapping process. An analysis of the accuracy of a data source should also consider additional error introduced during the data conversion process itself. For example, a digitiser operator will usually miss the precise mathematical centre of a line during the digitising process, and this error will be reflected in the database. In all, the combined effect of source error and data conversion error should be considered during the determination of source accuracy levels and the establishment of quality control error.

Updating is an issue that is significant because spatial data are dynamics. In surveying, updating of survey dataset is crucial, for example it is a need to establish point coverage of control cadastral monuments or GPS station which eventually will yield accurate GIS datasets. Mapping, updating, and distributing these points to the geospatial community, and the surveyors as well as developers is essential to the success of the GIS system. Using GIS application and management techniques, updating can be carried out in a single portal application such as ArcGIS. Documenting the information and data concerning updating of survey, processed and GIS-ready information can be made accessible by Internet GIS. Datasets need updating periodically so that data quality and currency can lead to efficient use by the JUPEM staff and public.

Using the open source technology as part of the research ideas promoted the use of OpenGIS Web Feature Server (WFS) protocol to do the updating of GIS-ready information in a heterogeneous networked environment. It is possible to design an Internet based geospatial information processing and updating environments which include disparate servers offering data layers and different client types (Brentjens et al., 2005). Brentjens et al. (2005) reported that the retrieval, updating and combination of geospatial data were relatively easy primarily because, firstly WFS Specification clearly describes the requests and responses that WFS should support and secondly, because WFS uses standard Web technology such as HTTP and GML. This was the reason for the success of the open source prototype dealt in this research except there is no addressing of updating issue specifically. As the government has drawn up the policy of using OST for the ICT development, updating the data online can be implemented along side in JUPEM and by all GIS stakeholders in Malaysia. This move thus will enable the cooperation and collaboration of private sectors and government departments for the development of geospatial data sharing and interoperability. Many private companies, because of its low cost and computing development have been using OST for their business function.

In managing flow of data from raw, processed to GIS-ready information, it is critical to provide mechanism for the maintaining of multiple states in the geodatabase. This will ensure the integrity of the GIS database. This ability to manage and view multiple states and to work with them is based on versioning. As the name implies, versioning explicitly records states (versions) of individual features and objects as they are modified, added, and deleted. A version records each state of a feature or object as a row in a table along with important transaction information (Hedstrom 2003). ArcSDE plays a critical role in versioned geodatabase applications and is used to manage long transactions in each RDBMS as well as across different systems. As the flow of survey data all the way to the GIS-ready state are about large quantity of information, versioning allows the quality, temporality and accuracy of the datasets be known. Staff in JUPEM would be given the clear picture of current data and historical data while converting to processed data and while disseminating them to the general public. As well, public are given with accurate and up-to-date state of data.

Geodatabase allows multiple access and continuous storing of spatial data through the use of versioning. Versioning enables simultaneous editing (subjects to DBMS permission). The versioning mechanism allows the stages of the flow of survey to GIS-ready information be established and the differences between versions can be reconciled, and master version of the geodatabase can be updated.

7.4 Conclusion

A way to manage surveys and other geospatial technologies, and its resources has been developed in a GIS environment. Through this research, the resultant data models, concepts and development can be acknowledged to be a new means to manage geospatial data from survey capture to the delivery of GIS-ready data to the clients. The designing of the feature dataset, feature classes inheriting from feature datasets, spatial objects entities, within geospatial data capture, processed data and GIS-ready information has been proven a concept successfully implemented.

It is believed that object orientation using relational arrangement has been an intuitive and effortless methodology to cater for the management of spatial entities such as data and information within the geospatial data capture towards the provision of GIS-ready information. The data model was implemented as an object-relational geographic database and served as an information foundation for exploration and analytical tools that supported data query, data loading, data exploration and data visualisation within the activities of data capture towards GIS-ready information production.

With the management model designed it is expected that the world of all surveys and data capture technologies including GPS survey, air survey, cadastral and topographic survey, engineering surveys, laser scanning and even hydrographic survey can more easily be aligned with the world of GIS. Moreover with the use of an object-oriented data modelling approach, representation of these complex datasets can be facilitated in a more intuitive manner. The approach has been seen to be flexible and perceptive and the data model can be modified as the need rises. In addition, by modelling the data and with improved management of 'one-way drill down' method users are able to more clearly observe and specify connections and relationships between objects in the database. This may lead to the construction of inputs for explanatory and predictive model. This thus encourages continued use of the data and promotes a more flexible environment for future extension and modification.

A web portal application of geographic data is vital for the implementation of a Spatial Data Infrastructure. Maguire and Longley (2005) explained that a web portal can add value to raw data by encapsulating data and tools in a single user-oriented application. This research has seen that this is possible as a sub-service to an SDI disseminating raw survey or survey package as metadata. A single portal application of a survey package can act as a core inventory description of raw and processed data.

The Internet or Web application provided in a national survey and mapping department would allow round the clock datasets access online throughout the government organisations, private agencies and the public. The online access to raw and processed data within the department and the resultant GIS-ready information provided by the survey and mapping department would open up the eyes of survey technicians, surveyors and external professionals to current technologies and trend, and in return give valuable response regarding their needs and demands.

By applying and implementing the research ideas, JUPEM would possess better geospatial data management and sound data dissemination means for survey datasets and GIS-ready information. On the whole, the general public will gain benefits and supports for the purpose of various developments and planning. They include infrastructure development, housing projects, territorial planning, political boundaries, and environmental assessment as well as for the building of NSDI. With the management of the raw and processed and dissemination of GIS-ready information, general public will have accurate, current, versioned and up-to-date metadata. Other government departments will feel the impacts of having data available online. Private companies get the ease of not going to the counter services in the government department to obtain geospatial data. Bureaucracy in the work flow of using and analysing geospatial data will be reduced. Daily works would be easier and users would save time, money and energy when GIS-ready data can be accessed via the Internet. Licensed surveyors will have the cadastral survey datasets for field reference available through the Internet thus moderating their need of accurate previous survey and its attributes to deploy new surveys. The time of searching and buying data from JUPEM can be saved.

With the metadata online (attribute of the GIS-ready information which may include source survey information), general public are able to obtain first-class information that will enhance their need and application on their work. The application to various professionals whom are lawyers, planner, architects, engineers, politician, state executive administrator, local authority and also to the academician can be elongated. Lawyer firm in the north part of the country who handles land disputes situated in the south may use the JUPEM single portal to access source survey data and metadata without travelling to the JUPEM state in south to obtain the source information.

Using the conceptual and logical modelling techniques such as in the research, many seamless databases in different sections in JUPEM can be stored in one database and can be retrieved by single GIS system. JUPEM has different system storage of geospatial data from topographic digital map, cadastral database, geodetics data and remote sensing data. A single portal data accessibility and availability through the Internet would necessitate the following suggested goals be achieved:

- To make available to clients locally and globally the capabilities to browse and downloading data for all JUPEM products, i.e. Cadastral, Mapping and Geodetic from a single point of Internet access.

- To establish a One-Stop-Centre at the JUPEM HQ where clients can walk-in and buy all DSMM products in hardcopies, printed copies and digital data in one place.

- To establish a Geomatic Data Centre where JUPEM becomes the reference agency for any GIS data requirement for the nation including government and private agencies.

Through the experience of the author working in JUPEM for nearly more than 20 years, the following can be drawn up in quantifying the amount of work, time and cost that can be saved using the suggested prototype.

Normally, by using total station method, a cadastral survey of 200 lots of housing land parcels takes 3 months from the processing of the survey data until they can be edited and stored in a cadastral database as GIS-ready information. This is carried out by different sections within JUPEM comprising electronic computation, certified plan plotting and cadastral base map charting. With the capability of a single portal geospatial data management system using the same viewing and editing application system, effort and time of the process and development can be projected to save about 30-40% of the actual work. Thus, period of 5 weeks can be removed from the 3 months schedule. Redundant data processing can be reduced by viewing all aspects of the survey and GIS-ready processing activities. Extra work of searching source data and processed data in separate files and locations can be saved. Costs of human resource therefore can be lower. The management of the geospatial data in the same GIS system can be more efficient due to interactive application of 'drill down' capability, 'bird-eye view' and 'single view' system.

Digital mapping of topographic base map data in JUPEM is carried out by three Sections: Data Capture, Database and Aerial Photography which implement separate systems in different locations. In a field topographic survey, by having a capability of knowing which parts of the region of survey have the aerial photo coverage or satellite images, producing GIS-ready information can be easy and straightforward. GIS-ready data can be produced by digitising the photogrammetrically-adjusted raster images. In the process of developing CAD processed data towards the production of GIS-ready information within one system, a total knowledge of the adjacent project area can lead to the reduction of extra work of locating and producing redundant data. In JUPEM, for example, a survey for producing GIS-ready information for a topographic base map of a 10 kilometer by 10 kilometer area will normally take up to 6 months to be completed passing through all the tasks in the three different locations and systems of the Sections. Using GIS within the same management system, time can be saved up to 40-50% for normal large area survey coverage for topographic mapping. Duration of survey can therefore be reduced to 3 months instead of the normal 6 months.

In conclusion, it can be said that the described way of storing and managing surveys and their datasets using GIS has promoted an intelligent, highly scalable, time-saving, many-to-one portal access administration and integration of survey works and, in so doing, can make the execution and querying of surveys less laborious, more timely, fitting and economic. Efficient management of surveys by GIS creates skill to track, allocate, analyse, manage and conserve survey activities. This has been proved by the 'one-way drill down' exploration and management. The one system management suggests that surveys can be conserved by storing and managing them in an intuitive and expressive way and by constantly updating as they were displayed during processing and analysing. With this skill, it is hoped to make the organisation's core business and efforts more effective. Data sharing and access is a development in GIS that has seen the need for the final product be usable, time saving and may provide spontaneous updating process. The global understanding and experience achieved in this research had provided essential skill and knowledge to handle geospatial data from raw to GIS-ready stage in a national survey and mapping department in same level to the commercial GIS contractors and expertise from the developed countries.

7.5 Recommendations for Future Works

This research did not address the methodology and process of updating the data by multiple users by way of Intranet or Internet. How to update and modify spatial and non-spatial data online by multiple users or various staff according to time and need is research that can be further proposed. Since spatial and non-spatial data are voluminous in a survey and mapping core business organisation, revising, updating and modifying data and information by different staff in different unit/branches across the nation is a crucial and essential development to be looked into in the future.

Further concentration of future work can be made to highlight the issue of visualisation of the temporal information of different surveys or geospatial technologies over a number of years by different surveyors in a geospatial data providing organisation. The land surveying process of amalgamation and subdivision into parcel lot has indicated that land parcel will changed over time. The visualisation of temporal changes in survey features and their spatial and non-spatial attributes will provide an enhanced method of exploration and analysis tools in managing the datasets that exist within field data capture, processed data and GIS-ready information. The multiple user ESRI's ArcSDE geodatabase allows multiple versions of the database to coexist. If the concepts of a versioning mechanism in ArcSDE geodatabase were to be implemented and tested, a GIS capable of storing behavioural and temporal datasets may be developed.

The research did not actually provide a way to store and manage 3-D spatial data such as DEM and TIN data in a geodatabase. TIN data were stored as files in the geodatabase and not as a table like the raster catalog. Visualisation of the TIN model was done in ArcGIS using ArcScene application by opening them in a single file navigating to their file names. Ways to store and manage 3-D datasets in a database application is an active research on its own (Kidner and Smith, 2003; Arens et al., 2005; Zhang et al., 2005). There is still room for improvements in perfecting the large quantity 3D data management and interactive graphic display.

REFERENCES:

Adair, A., Berry, J., Deddis, B., McGreal, S., Keogh, G. and Key, T. (1998). "Barriers to data sharing in the surveying profession: implications for the commercial property market." Journal of Property Research **15**(4): 331-346.

Adam, N. R., and Gangopadhyay, A. (1997). Database Issues in Geographic Information Systems, Kluwer Academic Publishers, Boston.

AGI (1999). The Association for Geographic Information (AGI) GIS Dictionary. [online], http://www.geo.ed.ac.uk/agidict/welcome.html, [checked 2005].

Ahern, D. (2004). Using GIS To Lure Attractive Retail Businesses Downtown. [online], http://www.directionsmag.com/article.php?article_id=567, [checked 2005], Article from Directions Magazine.

Arctur, D., and Zeiler, M (2004). Designing Geodatabase: Case Studies in GIS Data Modelling, ESRI, Redland, California.

Arens, C., Stoter, J. and van Oosterom, P. (2005). "Modeling 3D spatial objects in a geo-DBMS using a 3D primitive." Computers & Geosciences **31**(2): pp 165-177.

Aschenwald, J., Leichter, K., Tasser, E., and Tappeiner, U (2001). "Spatio-temporal landscape analysis in mountainous terrain by means of small format photography: a methodological approach." IEEE Transactions on Geoscience and Remote Sensing **39**(4): 885-893.

Awang, H. (2004). "Human Capital and Technology Development in Malaysia." International Education Journal **5**(2): pp 239-246.

Balley, S., Parent, C. and Spaccapietra, S (2004). "Modelling geographic data with multiple representations." International Journal of Geographical Information Science **18**(4): pp 327-352.

Baumann, J. (2004). Enterprise GIS in Dubai. GIS @ Development: [online], http://www.gisdevelopment.net/magazine/gisdev/2004/may/enterprise_gis.shtml, [checked 2005].

Beeri, C., Formica, A. and Missikoff, M (1999). "Inheritance hierarchy design in object-oriented databases." Data and Knowledge Engineering **30**(3): pp 191-216.

Bernhardsen, T. (2002). Geographic Information Systems: An Introduction, John Wiley and Sons, New York.

Bishr, Y. (1998). "Overcoming the semantic and other barriers to GIS interoperability." International Journal of Geographical Information Science **12**(4): 299-314.

Blaha, M., and Premerlani, W. (1998). Object-Oriented Modeling and Design for Database Applications, Prentice-Hall, Upper Saddle River, New Jersey.

Blaha, M. R., Premerlani, W.J. and Rumbaugh, J.E (1988). "Relational Database Design Using an Object-Oriented Technology." Communications of ACM **31**(4): pp 414-427.

Booch, G., Rumbaugh, J., and Jacobson, I. (1999). The Unified Modeling Language User Guide, Addison Wesley Longman, Reading, Massachusett.

Bosak, K., and Schroeder, K. (2005). "Using GIS for gender and development." Development in Practice **15**(2): pp 231-237.

Bossler, J. D., Jensen, J.R., McMaster, R.B. and Rizos, C.(eds) (2002). Manual of Geospatial Science and Technology, Taylor and Francis, London and New York.

Brentjens, T. J., Vries, M.E.D., Guak, C.W., Vijibrief, C., and Oosterom, P.J.M. (2005). "Updating geo-information in a heterogeneous networked environment – Experiences and evaluation of OpenGIS Web Feature Services." **[online],** http://www.gisdevelopment.net/technology/gis/techgis_005.htm: [checked 2006].

Burrough, P. A., and McDonnell, R.A. (2000). Principles of Geographic Information Systems, Oxford University Press, New York.

Calkins, H. W. (1992). Institutions Sharing Spatial Information, Networking Spatial Information Systems, Belhaven Press, UK.

Capretz, L. F. (2003). "A Brief History of the Object-Oriented Approach." ACM SIGSOFT Software Engineering **28**(2): pp 6-16.

Carlisle, B. H. (2005). "Modelling the spatial distribution of DEM error." Transactions in GIS **9**(4): pp 521-540.

Casper, M. L., Barnett, E., Halverson, J.A., Elmes, G.A., Braham, V.E., Majeed, Z.A., Bloom, A.S. and Stanley, S. (2000). Women and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality, Second Edition. Morgantown, West Virginia, Office for Social Environment and Health Research, West Virginia University.

Cavallini, P. (2004). "Open Source and Free GIS: A Way Out?" **[online],** http://www.gisdevelopment.net/technology/gis/techgis_004.htm: [checked 2006].

Chainey, S., and Ratcliffe, J. (2005). GIS and Crime Mapping, John Wiley and Sons, London.

Chrisman, N. (2002). Exploring Geographic Information Systems, John Wiley & Sons, New York.

Chrisman, N. R. (1999). "What Does 'GIS' Mean?" Transactions in GIS **3**(2): pp 175-187.

Chunithipaisan, S., Majeed, Z.A., James, P., Parker, D., and Abele, S. (2003). Geospatial Interoperability via the Web: Supporting Land Administration in Kuala Lumpur. Proceedings of MapAsia 2003 Conference, Kuala Lumpur, Malaysia, 13-15 October 2003.

Clementini, E., and Di Felice, P. (2001). "A spatial model for complex objects with a broad boundary supporting queries on uncertain data." Data and Knowledge Engineering **37**: pp 285-305.

Cova, T., and Goodchild, M. (2002). "Extending geographical representation to include fields of spatial objects." International Journal of Geographical Information Science **16**(6): 509-532.

Crompvoets, J., Bregt, A., Rajabifard, A. and Williamson, I (2004). "Assessing the worldwide developments of national spatial data clearinghouses." International Journal of Geographical Information Science **18**(7): pp 665-689.

Crowther, P., and Hartnett, J. (2001). Handling Spatial Objects in a GIS Database - Relational v Object-Oriented Approaches. Proceedings of the 6th International Conference on GeoComputation, University of Queensland, Brisbane, Australia, 24 - 26 September 2001,.

Daud, M. N. (1999). Public Sector Information Management and Analysis Using GIS in Support of Property Valuation in Malaysia. Department of Geomatics. PhD Thesis, University of Newcastle upon-Tyne: 306p.

Demers, M. N. (2000). Cartographic and GIS Data Structures. *Fundamentals of Geographic Information Systems*, John Wiley and Sons, 2nd Edition: pp 117-120.

De Man, W. H. E. (2006). "Understanding SDI: complexity and institutionalisation." International Journal of Geographical Information Science **20**(3): pp 329-343.

De Man, W. H.E. and Den Toorn, W.H.V. (2002). "Culture and the adoption and use of GIS within organisations." International Journal of Applied Earth Observation and Geoinformation 4(1): pp 51-63.

Devogele, T., Parent, C., and Spaccapietra, S. (1998). "On spatial database integration." International Journal of Geographical Information Science **12**(4): 335-352.

Diepenbroek, M., Grobe, H., Reinke, M., Schindler, U., Schlitzer, R., Sieger, R., and Wefer, G. (2002). "PANGEA - an information system for environmental sciences." Computers and Geosciences **28**(10): 1201-1210.

Diviacco, P. (2005). "An open source, web based, simple solution for seismic data dissemination and collaborative research." Computers & Geosciences **31**(5): pp 599-605.

Dollner, J., and Hinrichs, K. (2000). "An object-oriented approach for integrating 3D visualisation systems and GIS." Computers & Geosciences **26**(1): pp 67-76.

Dragicevic, S. (2004). "The potential of Web-based GIS." Journal of Geographical Systems **6**(2): pp 79-81.

Dugacek, D., Hronec, J, Timcak, G. M., and Hroncova, E. (1997). GEO-an application software for the presentation of zircon morphometry data. [online], http://actamont.tuke.sk/pdf/1997/n1/9dugacek.pdf, [checked 2005], A Report, Technical University of Kosice, Slovakia.

Egenhofer, M. J., and Frank, A.U. (1992). "Object-Oriented Modelling for GIS." URISA Journal **4**(2): 3-19.

Elmasri, R., and Navathe, S.B. (2004). Fundamentals of Database Systems, Fourth Edition, Addison Wesley.

ESRI (1999). Exploring ArcObjects, ESRI Press, California.

ESRI (2001). ArcGIS Desktop: Integrating Data from the Web. An ESRI White Paper. **[online], http://www.esri.com/library/whitepapers/pdfs/arcgis_dt_web_data.pdf, [checked 2005]**.

ESRI (2003a). "Spatial data standards and GIS interoperability." An ESRI White Paper **[online], http://www.esri.com/library/whitepapers/pdfs/spatial-data-standards.pdf**: [checked 2004].

ESRI (2003b). ArcIMS 4 Architecture and Functionality. An ESRI White Paper, [online], http://downloads.esri.com/support/whitepapers/ims_/arcims4_architecture.pdf, [checked 2005].

Fotheringham, A. S., and Wegener, M .(eds) (2000). Spatial Models and GIS: New Potential and New Models, Taylor and Francis, London.

Frank, A. (1992). "Qualitative spatial reasoning about distances and directions in geographic space." Journal of Visual Languages and Computing: 343-71.

Fulghum, R. (1987). All I Really Need To Know I Learned in Kindergarten. New York, Villard Books.

Fussel, M. L. (1997). Foundations of Object-Relational Mapping. Chimu Publication: [online], http://www.chimu.com/publications/objectRelational/index.html [checked 2005].

Gartner, H., Bergmann, A., and Schmidt, J. (2001). "Object-oriented modelling of data sources as a tool for the integration of heterogeneous geoscientific information." Computers and Geosciences **27**: 975-985.

Gauna, I., and Sozza, A. (1999). "The geographic information system of the Turin City Council on the Internet." Computers, Environment and Urban Systems **23**: 485-494.

Ghandeharizadeh, S. (2003). An Educational Perspective on Database Management Systems and Object-Oriented Methodology: A 12 Year Journey. 18th Annual ACM Conference on Object-Oriented Programming, Systems, Languages and Applications (OOPSLA) 2003, Anaheim, California, 26-30 October 2003.

Gielsdorf, F., Gruendig, L. and Aschoff, B. (2004). Positional Accuracy Improvement - A Necessary Tool for Updating and Integration of GIS Data. FIG Working Week, Athens, Greece, May 22-27 2004.

GIS Development Pvt Ltd. (2005). Mapping GIS Milestones : 1960-1970. [online], http://www.gisdevelopment.net/history/1960-1970.htm, [checked 2005], GIS Resource Portal : History.

Gogu, R. C., Dietrich, V.J., Jenny, B., Schwandner, F.M. and Hurnia, L. (2006). "A geo-spatial data management system for potentially active volcanoes—GEOWARN project." Computers & Geosciences **32**(1): pp 29-41.

Gong, J., and Wang, Y. (2004). Federated Spatial Databases and Interoperability. Science and Technology Magazine Online: [online], http://www.paper.edu.cn/scholar/known/gongjianya/gongjianya-9.pdf, [checked 2005].

Goodchild, M. F. (1998). The Geolibrary. Innovation in Geographical Information System 5. In Carver, S. (ed.), Taylor and Francis, London: pp 59-68.

Gordon, S., Lichti, D. and Steward, M. (2001). Application of a high-resolution, ground-based laser scanner for deformation measurements. The 10th FIG International Symposium on Deformation Measurement, Orange, California, USA, 19-22 March 2001.

Granger, N. (2005). "Raising the profile of the surveying profession." The Journal of the Institution of the Civil Engineering Surveyor: January 2005, pp 20-22.

Greenwood, R. W. (2002). Using Mapserver to Integrate Local Government Spatial Data. Proceedings of the Open Source GIS-GRASS Users Conference 2002, Trento, Italy, 11-13 September 2002.

Gregorius, T. (2005). Geospatial Fuel for the Economy. Geoinformatics. **8:** pp15.

Guting, R. H. (1994). "An Introduction to Spatial Database Systems." Very Large Data Bases (VLDB) Journal **3**: 357-399.

Haines-Young, R., and Watkins, C (1996). "The Rural Data Infrastructure." International Journal of Geographical Information Systems **10**(1): pp 21-46.

Hedstrom, K. (2003). Understanding ArcSDE Table Relationships. ESRI's ArcUser - July -Sept 2003: pp 28-30.

Holland, D. A., and Tompkinson, W. (2003). Improving the update of geospatial information databases from imagery using semi-automated user-guidance techniques. Proceedings of the 7th International Conference on GeoComputation, University of Southampton, United Kingdom, 8-10 September 2003.

Hoult, C. (1995). The Requirement, Implementation and Use of a Generic Foundation Dataset for Large-Scales Spatial Data Management. Newcastle upon-Tyne, UK: 180p, Department of Surveying, University of Newcastle upon-Tyne: 180p.

Huang, B., and Lin, H. (2002). "A Java/CGI approach to developing a geographic virtual reality toolkit on the Internet." Computers and Geosciences **28**(1): 13-19.

Hughes, J. G. (1991). Object-Oriented Databases, Prentice Hall, London.

Huxhold, W. E., and Levinsohn, A.G. (1995). Managing Geographic Information System Projects, Oxford University Press, Oxford.

Igbokwe, J. I., and Emengini, E. J. (2005). GIS in Management of Electricity Distribution Network: A case study of Onitsha-North L.G.A., Anambra state, Nigeria. GIS Development, Jan 2005: [online], http://www.gisdevelopment.net/application/utility/power/utilityp0022.htm, [checked 22/02/2005].

ISO (2005). Geographic Information: ISO/IEC 19501:2005. International Organisation for Standardisation (ISO), [online], http://www.iso.org/iso/en/CombinedQueryResult.CombinedQueryResult?queryString=UML, [checked 2005].

Jeffress, G. A. (2003). "The Value of Cadastral Surveying to Efficient Land Administration." Surveying and Land Information Science (SaLIS) of the National Society of Professional Surveyors **63**(4).

Jericke, E. (2002). The Integration of Surveying and GIS - The Basis of Sustainable Management of Spatially Related Information. Conferences on Information Technologies in Landscape Architecture: Trends in GIS and Virtualization in Environmental Planning and Design, Anhalt, Germany, 2002.

Jerome, J. A. (2005). Interoperability: a UK Cabinet Office e-Government Unit perspective. GIS Professional. **Issue 2:** pp 10-12.

Jiang, B. (2005). "A structural perspective on visibility patterns with a topographic surface." Transactions in GIS **9**(4): pp 475-488.

JUPEM (2004). "Department of Surveying and Mapping Malaysia." **[online]** **http://www.jupem.gov.my:** [checked 2004].

Kain, R. J. P. (2002). "The Role of Cadastral Surveys and Maps in Land Settlement from England." Landscape Research, Carfax Publishing, part of the Taylor & Francis Group **27**(1): 11-24.

Kang, M. A., Pinet, F., Schneider, M., Chanet, J.P., and Vigier, F. (2004). How to Design Geographic Databases? Specific UML Profile and Spatial OCL Applied to Wireless Ad Hoc Networks. 7th AGILE Conference on Geographic Information Science, Heraklion, Crete, 29 April -1 May 2004.

Karikari, I., Stillwell, J. and Carver, S. (2005). "The application of GIS in the lands sector of a developing country: challenges facing land administrations in Ghana." International Journal of Geographical Information Science **19**(3): pp 343-362.

Kemper, A., and Moerkotte, G. (1994). Object-Oriented Database Management: Application in Engineering and Computer Science, Prentice -Hall International, Inc.,New Jersey.

Kennan, P. (1998). "Integrating GIS and Management Science." Operational Research (OR) Insight **11**(1): 23-28.

Khalsa, S. J. S., Dyurgerov, M.B., Khromova, T., Raup, B.H., Barry, R.G. (2004). "Space-based mapping of glacier changes using ASTER and GIS tools." IEEE Transactions on Geoscience and Remote Sensing **42**(10): 2177 - 2183.

Kidner, D. B., and Smith, D.H. (2003). "Advances in the data compression of digital elevation models." Computers & Geosciences **29**(8): pp 985-1002.

Kim, W. (1990). Research directions in object-oriented database systems. Proceedings of the 9th ACM SIGACT-SIGMOD Symposium on Principles of Database Systems, pp 1-15, Tennessee, United States, April 1990.

Kingsbury, D. (2005). The "IT Geo-Evolution"- The Evolution of Geospatial Technology in IT. Directions Magazine - February 2005, [online], http://www.directionsmag.com/article.php [checked 2005].

Kohsaka, H. (2000). "Applications of GIS to urban planning and management: Problems facing Japanese local governments." GeoJournal **52**(3): pp 271-280.

Kraak, M. J. (2004). "The role of the map in a Web-GIS environment." Journal of Geographical Systems **6**(2): pp 83-93.

Kriegel, H., Kunath, P., Pfeifle, M and Renz, M. (2004). Object-relational management of complex geographical objects. Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems, ACM Press, pp 109-117, Washington DC, USA, Nov 2004.

Kroha, P. (1993). Objects and Databases, McGraw-Hill Book Company, London.

Lake, R. (2001). GML lays the foundation for the geospatial Web. GEOWorld - October 2001: pp 42-45.

Lake, R. (2005). "The application of geography markup language(GML) to the geological sciences." Computers & Geosciences **31**(9): pp 1081-1094.

Laurini, R. (1998). "Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability." International Journal of Geographical Information Science **12**(4): 373-402.

Laurini, R., and Thompson, D. (1995). Fundamentals of Spatial Information Systems, Academic Press Limited, London.

Lawrence, V. (2004). Putting GIS on the map. RICS Business Magazine - December 2004: [online], http://www.rics.org.uk/Environmentalandlandconsultancy/Geomatics/Geographicinformationscience/rics_bus_nov_dec2004_gis_map.html, [checked 2005].

Lemmen, C. (2005). Geo Databases. GIM International - January 2005: pp 40-41.

Leung, Y., Leung, K.S. and He, J.Z (1999). "A generic concept-based object-oriented geographical information system." International Journal of Geographical Information Science **13**(5): pp 475-498.

Longley, P. A. (2000). "The academic success of GIS in geography: Problems and prospects." Journal of Geographical Systems **2**(1): pp 37-42.

Longley, P. A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W. (2002). Geographic Information Systems and Science, John Wiley & Sons, Chichester.

Loomis, M. E. S. (1995). Object Databases: The Essentials, Addison-Wesley Publishing Company, Reading.

Lu, P. F., and Ping A. (1999). "A metric for spatial data layers in favorability mapping for geological events." IEEE Transactions on Geoscience and Remote Sensing **37**(3): 1194 - 1198.

MacDonald, A. (2001). Building a Geodatabase, ESRI Press, Redlands, California.

Maguire, D. J., and Longley, P.A (2005). "The emergence of geoportals and their role in spatial data infastructures." Computers, Environment and Urban Systems **29**: pp 3-14.

Majeed, Z. A., and Parker, D. (2004). Geographic Information System for Managing Survey Data to the Development of GIS-Ready Information. Proceedings of the 3rd FIG Regional Conference, Jakarta, Indonesia, 3-7 October 2004.

MAMPU (2006). "Malaysian Public Sector Open Source Initiative." **[online]** http://opensource.mampu.gov.my/: [checked 2006].

Mattos, N. M., Meyer-Wegener, K. and Mitschang, B. (1993). "Grand tour of concepts for object-orientation from a database point of view." Data & Knowledge Engineering, **9**(3): pp 321-352.

McBride, J. (2004). Using Historic Survey Data in Conservation GIS. ESRI 2004 User Conference, San Diego, California, 9-13 August 2004.

McConnell, T., and White, C. (1999). Using ArcView GIS as a project management tool. ESRI 1999 User Conference, San Diego, California, July 1999.

McKee, F. (2002). The Importance of GIS in Modern Society. GIM International - June 2002**:** pp 68-71.

Medeiros, C. B., and Pires, F. (1994). "Databases for GIS." ACM Special Interest Group for Management of Data(SIGMOD) Record **23**(1): 107-115.

Microsoft Corporation. (2005). Microsoft Developer's Network (MSDN) Library, [online] http://msdn.microsoft.com/library/default.asp, [checked 2005].

Mitasova, I., Visoovcova, J. and Hajek, M. (2002). Approaches to Data Modelling in GIS. [online], http://e-collection.ethbib.ethz.ch/ecol-pool/inkonf/inkonf_65.pdf, [checked 2005], Laboratory of Geoinformatics, Slovak Technical University, Slovakia.

Mitchell, A. (1998). Zeroing In: Geographic Information Systems at works in the community, ESRI, Redland, California.

Molenaar, M. (1998). An Introduction to the Theory of Spatial Object Modelling for GIS, Taylor and Francis Ltd, London.

Nash, E., James, P. and Parker, D. (2005). "A model for spatio-temporal network planning." Computers & Geosciences **31**(2): pp 135-143.

Neild, J. (2004). Sharing Multiband Imagery Analysis Using ArcSDE. ESRI's ArcUser - April-June 2004**:** pp 50-51.

Neldner, V. J., Crossley, D.C., and Cofinas, M. (1995). "Using GIS to determine the adequacy of sampling in vegetation surveys." Biological Conservation **73**(1): pp 1-17.

Neteler, M., and Mitasova, H (2002). Open Source GIS : a GRASS GIS Approach, Kluwer Academic, London.

Newell, R. G., and Theriault, G. (1995). Ten Difficult Problems in Building a GIS. [online], http://emea.smallworld.co.uk/support/techpaper/tp1.html, [checked 2005], Technical Paper, SmallWorld Inc.

Nix, M., and Hill, C. (2001). Rapid Accurate GIS Updates with Improved Link to Survey System. 5th Global Spatial Data Infrastructure Conference, Cartegena, Columbia, 23 May 2001.

Nyerges, T., Janskowski, P., and Drew, C. (2002). "Data-gathering strategies for social behavioural research about participatory geographical information systems use." International Journal of Geographical Information Science **16**(1): pp 1-22.

OGC (2003). "Web Feature Service Implementation Specification." Open GIS Consortium **[online], http://www.opengeospatial.org/specs/?page=specs, [checked 2005]**.

OGIS. (1998). "The OpenGIS Guide: Introduction to interoperable geoprocessing." Open GIS Consortium **[online]** **http://www.opengeospatial.org:** [checked 2004].

OMG (2004). CORBA, XML and XMI: Object Management Group Resource Page. **[online], http://www.omg.org/technology/xml/index.htm, [checked 2005]**.

Oosterom, P. v., and Vandebos, J. (1990). An Object-Oriented Approach to the Design of Geographic Information Systems. Proceedings of the First Symposium on Design and Implementation of Large Spatial Databases, Springer-Verlag London, UK.

Parker, J. H., Miller, J.E. and Schoenherr, J.R (1996). "Collection and processing of multibeam and side-scan data to hydrographic standards." IEEE Conference Proceedings on OCEANS -'Prospects for the 21st Century', 23-26 Sept. 1996 **3**: pp 1108-1113.

Parsley, S. (2001). Methodologies for Distributed and Higher Dimensional Geographic Information. PhD Thesis, Department of Geomatics, University of Newcastle upon-Tyne: 232p.

Peng, Z. R., and Tsou, M.H. (2003). Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks, John Wiley and Sons, Hoboken, New Jersey.

Peng, Z. R., and Zhang, C (2004). "The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS)." Journal of Geographical Systems **6**(2): pp 95-116.

Pieraccini, M., Fratini, M., Parrini, F., Pinelli, G., and Atzeni, C. (2005). "Dynamic Survey of Architectural Heritage by High-Speed Microwave Interferometry." IEEE Geoscience and Remote Sensing Letters **2**(1): 28-30.

Pooley, R., and Stevens, P. (1999). Using UML: Software Engineering with Objects and Components, Addison Weslley Longman Limited, Harlow, Essex.

PostGIS (2006). "PostGIS by Refractions Research." **[online],** http://postgis.refractions.net/index.php: checked 2006].

Premerlani, W. J., Blaha, M.R. and Rumbaugh, J.E (1990). "An Object-Oriented Relational Database." Communications of ACM **33**(11): pp 99-109.

PSM (2004). History of Surveyor. Professional Surveyor Magazine. [online], http://www.profsurv.com/newpsm/, [checked 2004].

Pundt, H., and Bishr, Y. (2002). "Domain ontologies for data sharing - an example from environment monitoring using field GIS." Computers and Geosciences **28**(1): 95-102.

Raafat, H. M., Xiao, Q., Gauthier, D.A. (1991). "An extended relational database for remotely sensed image data management within GIS." IEEE Transactions on Geoscience and Remote Sensing **29**(4): 651-655.

Rainis, R., and Sulaiman, S. (2001). The Development of Internet-Based GIS Metadata for Penang Integrated Coastal Zone Management Pilot Project. Malaysian Surveyors Magazine**:** pp 24- 29.

Rajabifard, A., and Williamson, I.P (2004). SDI Development and Capacity Building. Proceedings of 7th Global Spatial Data Infrastructure Conference, Bangalore, India, February 2-4, 2004.

Rajabifard, A., Feeney, M.F. and Williamson, I.P (2002). "Future directions for SDI development." International Journal of Applied Earth Observation and Geoinformation **4**(1): pp 11-22.

Ramroop, S., and Pascoe, R. (2001). "Use of LDAP to partially implement the OGIS discovery service." International Journal of Geographical Information Science **15**(5): pp 391-413.

Raper, J., Dykes, J., Wood, J., Mountain, D., Krause, A. and Rhind, D. (2002). "A framework for evaluating geographical information." Journal of Information Science **28**(2): pp 39-50.

Raza, A., and Kainz, W. (1999). Cell tuple based spatio-temporal data model: an object oriented approach. Proceedings of the 7th ACM international symposium on Advances in geographic information systems, Kansas City, Missouri, United States, ACM Press, New York, pp 20-25.

Reed, M. D., Landry, C.E. and Werther, K.C (1996). The application of air and ground based laser mapping systems to transmission line corridor surveys. IEEE Symposium on Position Location and Navigation, 22-26 April 1996, pp 444-451, [online] http://ieeexplore.ieee.org/iel3/3679/10819/00509113.pdf?tp=&arnumber=509113&isnumber=10819, [checked 2006].

Rigaux, P., Scholl, M., and Voisard, A. (2002). Spatial Databases: With Application to GIS, Morgan Kaufmann Publishers, London.

Rodriguez, A. (2004). Brief Introduction to Database Concepts. [online], http://edugi.uji.es/Andrea/DBreading.pdf, [checked 2005], Article from University of Concepcion, Chile.

Rollins, R. (2005). GIS in estate management, Institution of Civil Engineering Surveyors's Year Book 2005: Geographic Information System, pp 43-48.

Shariff, M. R. M., and Krishna, B. (2003). Keranamu Malaysia : We are moving ahead.... GIS Development, Oct 2003: [online], http://www.gisdevelopment.net/magazine/gisdev/2003/october/kmwm.shtml, [checked 21/02/2005].

Smith, R. A., and West, G.R (1999). "Airborne lidar: a surveying tool for the new millennium." IEEE on OCEANS-Riding the Crest into the 21st Century, 13-16 Sept. 1999 **3**: pp 1290-1299.

STARPLUS (2005). User Manual Supplement for Star*net Pro V6 for WIndow. Starplus Software, Inc.

Steimann, F. (2000). "On the representation of roles in object-oriented and conceptual modelling." Data and Knowledge Engineering **35**(1): pp 83-106.

Stoimenov, L., and Djordjevic-Kajan, S. (2005). "An architecture for interoperable GIS use in a local community environment." Computers & Geosciences **31**(2): 211-220.

Stonebraker, M., Brown, P., and Moore, D. (1999). Object-Relational DBMSs: Tracking the Next Great Wave, Morgan Kaufmann Publishers, San Francisco.

Strobl, J. (2002). Georeferenced Internet Communications in a Geoinformation Society. Conference on Information Technologies in Landscape Architecture: Trends in GIS and Virtualization in Environmental Planning and Design, Anhalt, Germany, 2002.

Su, Y., Slottow, J., and Mozes, A. (2000). "Distributing proprietary geographic data on the World Wide Web - UCLA GIS Database and Map Server." Computers & Geosciences **26**(7): pp 741-749.

Tang, W., and Selwood, J. (2003). Connecting Our World : GIS Web Services, ESRI Press, Redlands, California.

Taylor, G. (1991). The Design, Development and Implementation of a Surveyor's Land Information Management Package,. Department of Geomatics. PhD Thesis, University of Newcastle upon Tyne: 299p.

Teng, C. H. (2000). Object recognition in aerial imagery assisted by GIS for topographic dataset revision. Department of Geomatics. PhD Thesis, University of Newcastle upon-Tyne: 221p.

Thurston, J., Poiker, T.K. and Moore, J.P. (2003). Integrated Geospatial Technologies: A Guide to GPS, GIS and Data Logging, John Wiley, New Jersey.

Tilley, S., and Huang, S. (2003). A qualitative assessment of the efficacy of UML diagrams as a form of graphical documentation in aiding program understanding. ACM's Special Interest Group for Design of Communication (SIGDOC) Conference, San Francisco, CA,12-15 October 2003.

Tong, C. W. (2002). Department of Survey and Mapping Malaysia: The Major Building Block for NaLIS. Open Seminar on Spatial Data Infrastructure in Asia and Pacific, 7th Permanent Committee on GIS Infrastructure for Asia and Pacific (PCGIAP), Tsukuba, Japan, April 2001, [online], http://www.gsi.go.jp/PCGIAP/tsukuba/seminar/seminar.htm, [checked 2005].

Turker, M., and Kocaman, S. (2003). The Design and Implementation of a Cadastral Database with a Spatiotemporal Modeling Approach in Turkey. Proceedings of MapAsia 2003 Conference, Kuala Lumpur, Malaysia, 13-15 October 2003.

UNESCO (1999). Spatial Analysis. [online], http://ioc.unesco.org/oceanteacher/resourcekit/, [checked 2005], Report from The International Oceanographic Data and Information Exchange (IODE) Resource Kit, UNESCO.

Utler, E. (2002). ArcSurvey: A Data Model for Land Survey. [online], http://www.leica-geosystems.com/corporate/en/ndef/lgs_1844.htm, [checked 2004], Technical Paper, Leica Geosystem AG.

Van der Vegt, M. (2001). Evaluating GIS in local government in developing countries. [online], http://www.gisdevelopment.net/thesis/thesis1/index.htm, [checked 2005], Internship Report, ITC, The Netherland.

Vanderpoorten, A., Sotiaux, A., and Engels, P. (2004). "A GIS-based survey for the conservation of bryophytes at the landscape scale." Biological Conservation **121**: 189-194.

Verbyla, D. L., and Chang, K.T. (1997). Processing Digital Images in Geographic Information Systems: A Tutorial Featuring ArcView and ARC/INFO, OnWorld Press, Santa Fe, New Mexico.

Voisard, A., and David, B. (2002). "A Database Perspective on Geospatial Data Modelling." IEEE Transaction on Knowledge and Data Engineering **14**(2): 226-243.

Wang, J., Hassett, J.M. and Endreny, T.A. (2005). "An object oriented approach to the description and simulation of watershed scale hydrologic processes." Computers & Geosciences **31**(4): pp 425-435.

Weers, W. (2004). NATO-compatible GIS for German Federal Armed Forces: Bundeswehr GeoInfo Data Base. GIM International - April 2004**:** pp 61-63.

Wei-Ning, X. (1996). "A GIS-based method for trail alignment planning." Landscape And Urban Planning **35**(1): pp 11-23.

Weir, M. F. (2004). The Next Generation GIS/LIS - A Survey Information System Integrated within a GIS. FIG Working Group Week, Paris, France, 13-17 April 2004.

Wheatley, J. H. (1998). A study of selected development in the use of GIS for infrastructure management. Department of Geomatics. Newcastle upon-Tyne, UK: 236p, University of Newcastle upon-Tyne**:** 236p.

Wiesel, J. (2002). Technical Trends in GeoInformation. Conferences on Information Technologies in Landscape Architecture: Trends in GIS and Virtualization in Environmental Planning and Design, Anhalt, Germany, 2002.

Wikle, T. A. (1998). "Continuing education and competency programmes in GIS." International Journal of Geographic Information Science **12**(5): 491-507.

Wilmersdorf, E. (2003). "Geocoded information incorporated into urban online services-the approach of the city of Vienna." Computers, Environment and Urban Systems **27**: pp 609-621.

Worboys, M. F. (1994). "Object-oriented approaches to geo-referenced information." International Journal of Geographical Information Systems **3**: pp 385-399.

Worboys, M. F., and Duckham, M. (2004). GIS: A Computing Perspective, CRC Press, London.

Worboys, M. F., Hearnshaw, H.H. and Maguire, D.J. (1990). "Object-oriented data and query modelling for spatial databases." International Journal of Geographical Information Systems **4**: pp 369-383.

Xu, Z. (2003). Geodata Integration and Semantic Interoperability: Network-enabling GIS. GIM International - November 2002**:** pp 35-37.

Yang, C., Wong, D.W., Yang, R., Kafatos, M. and Li, Q (2005). "Performance-improving techniques in web-based GIS." International Journal of Geographical Information Science **19**(3): pp 319-342.

Young, R. L. (2004). Using ArcView to Manage a Surveying Company in South Texas. American Congress on Surveying and Mapping, USA, April 2004, [online], http://www.acsm.net/YoungGISSmallBus42004.pdf [checked 2005].

Yunus, M. Z. M. (1997). Geospatial Data Management Throughout Large Area Civil Engineering Projects. Department of Geomatics. PhD Thesis, University of Newcastle upon Tyne: 309p.

Zeiler, M. (1999). Modeling Our World, ESRI Press, Redlands, California.

Zhang, J., and Goodchild, M.F. (2002). Uncertainty in Geographical Information, Taylor & Francis, London.

Zhang, L., Yang, C., Liu, D., Ren, Y. and Rui, X. (2005). "A web-mapping system for real-time visualisation of the global terrain." Computers & Geosciences **31**(3): pp 343-352.