Detection of Spatial Inconsistencies in Land Use Data Updating

Xiao Shengjun

M.Phil
The Hong Kong Polytechnic University
2013

The Hong Kong Polytechnic University

Department of Land Surveying and
Geo-informatics

Detection of Spatial Inconsistencies in Land Use Data Updating

Xiao Shengjun

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Philosophy

Jan. 2012

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____________________________ (Signed)

__Xiao Shengjun_______________ (Name of student)

Abstract

The purpose of this thesis is to develop a method for detecting spatial inconsistency in land use data updating based on the information of geometry, attribute and topological relationship of land use data. GIS data need to be updated regularly to maintain its timeliness, especially for land use data, which is very important when the strategy of land resource and local city planning is determined. However, during updating, errors or inconsistency may occur and influence the quality of land use data. Updating land use data needs a set of steps: surveying/mapping data, manual and automatic data interpretation and so on. Current updating methods cannot maintain the consistency and validity of data perfectly, which inevitably generate errors and inconsistency. Errors could be wrong value, such as wrong perimeter or area value, repeated key value, such as ID. In normal situation, errors could be detected easily, while data inconsistency is a special kind of error occurred in a potential way. Inconsistency is the error or illogical meaning implied by data. To detect data inconsistency, meaning and logic of data should be checked. Typical land use data inconsistency could be a conflict in attributes of two neighboring lands, such as limited height of buildings near airport land. Land use data was created under land classification standard formulated by the Government. Application of variant standards when checking data inconsistency increases the difficulties of detecting data inconsistency. If this is ignored, data quality will have problems. Moreover, inconsistency may accumulate during updating which cause further problems. Based on the study of land use data updating procedure, a new method is proposed to identify potential semantic inconsistency of data, which is ignored or hard to be detected in data quality control. It is useful for maintaining the quality of land use data after updating. Geometric, attribute and topological relationship information of land use data form the basis for analysis. Specific combination of these three aspects could be set as a rule to judge whether inconsistency occurs. Three sets of rules are extracted as references to identify inconsistency according to the types of inconsistency. Rules are designed and tested in a prototype system which is established via .NET Frame Work 2005 and C# language by using Arc Engine (AE9.2). The experiment land use data are obtained from Land and Resources Bureau of Beijing. The results show that detection rules proposed by the study detect 296 inconsistencies within the total of 1517 records, “the same class land merging problem” occurs most frequently, which is about 16.74% records in the experiment data. Careful verification suggests that all of these detected inconsistencies are true. The results of experiment indicate that detection methods proposed by this study are feasible and helpful for controlling the accuracy of the meaning of land use data, as well as the accuracy of data format and structure.

Key words: Land use data, Inconsistency, Rules, Data semantic

Acknowledgements

First and foremost, I would like to show my deepest gratitude to my supervisor, Dr. Lilian Pun, a respectable, responsible and resourceful scholar, who has provided me with valuable guidance in every stage of the writing of this thesis. Without her enlightening instruction, impressive kindness and patience, I could not have completed my thesis. Her keen and vigorous academic observation enlightens me not only in this thesis but also in my future study.

I shall also extend my thanks to Prof. Li for all his kindness and help. I would also like to thank all my teachers who have helped me to develop the fundamental and essential academic competence. My sincere appreciation also goes to the teachers and students from room HJ716, department of LSGI, who participated this study with great cooperation.

Last but not least, I would like to thank my family, especially my parents, for their encouragement and support.

TABLE OF CONTENTS

1. Introduction

1.1 Background

Land is one of the most fundamental and essential resources for human life, because most of human activities occur on land. Two concepts or phrases are used when people manage land-- “land cover” and “land use”. Land cover describes the natural environment or physical material of the surface of land, while land use is distinct from land cover even these two phrases are often used interchangeably. Land use data usually are shown as the maps that depict the distribution of a set of land use classes over a territory represented as the map legend. The legend contains a fixed number of mutually exclusive and collectively exhaustive classes (each one represented by a particular label), so that any given terrain unit can be assigned a label (Castilla et al., 2005).

Land use describes socio-economic activities happening on the land. Land use represents the perspective of how people use the land and what kind of socio-economic activities happen on the land, rather than the natural physical material of the earth surface.

Generally, land use information is recorded and presented in the land use map. Digital land use map has two formats: raster and vector. Raster land use map can be the raw remote sensing images. After classification and vectorization, raster map can be converted to the vector map. Normally, standard land use map is in the vector format. An important advantage of vector map is that it can contain lots of attribute information. For land use map, as an instance, vector map can not only present the land map, it can also record other attributes information such as land owner. Casting land use information to other thematic maps can be easily achieved by overlapping the two vector maps and updating the attribute of thematic map according to that of land use map. Therefore, the vector land use map is widely used in further mapping and

GIS application. Vector maps will be discussed as the representative example of land use data in this study.

The content of land use map, as mentioned earlier, is the information of socio-economic activities which occur on the land surface. In detail, land use maps present socio-economic activity within its boundary. Various socio-economic activities happened on the land are classified according to the specific standard or application purpose. Land use map published by different institutes may have totally different classification standards and contents which present quite different socio-economic activities and distributions. The boundary of land determines the area of a certain land use. The definition of "boundary" is also diverse. For a vector map developed from remote sensing images, the "boundary" between two polygons should be the natural boundary of two lands which have different socio-economic activities in real world. When a street block map was updated according to the land use attribute, the "boundary" of polygons might be the street line and road network in the area, or the natural boundary inside the block. For example, figure 1.1-1 shows land use divided by the natural boundary and by the street block, respectively. In summary, vector land use map consists of areas with attributes which describe the activities happen on lands and the boundaries of those activities.

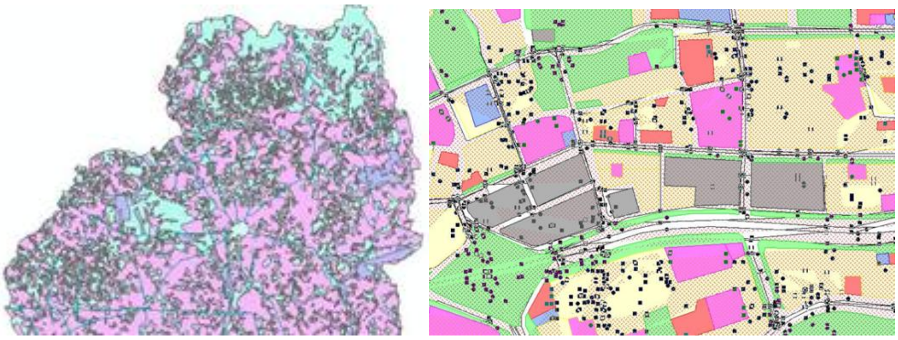


Figure 1.1-1 Boundary styles in vector land use map

To maintain the validity of land use map, regular information updating is necessary. All GIS data face the challenge of updating mission as geodata often contains diversified information and requires mass data processing techniques.

Without timely updating, land use map will soon become outdate and useless because socio-economic activities are developing and changing rapidly. There are a number of solutions or techniques can be used to complete the updating mission which will be introduced in detail in the following chapter.

However, updating land use map may cause inconsistency problems. Data updating is a very complex job. It needs site surveying, office mapping, image interpretation and automatic or man-made classification. During all the steps of updating, it is hard to maintain the consistency of data because update data means changing the old one to a new one. Updating method will not consider if the changed new data is conflict or consistent with the other data. As maps were updated constantly, data quality control should be paid attention to preserve the integrity and consistency of data because the up-to-date data may have negatively effect on land use information. Data consistency requires common logical and standard rules of map semantic. Consistency demands not only correction of data format and structure, but also the meaning and logic of data. Current quality monitoring approaches have done well in data format checking but have limitations in checking the semantic of data. This study focuses on analyzing the reasons of causing inconsistency problems and proposes a set of approaches to solve several specific inconsistency problems.

Reasons causing land use data inconsistency and formats of inconsistency are complex. The solution suggests that starting analysis from the land use data updating technology, because inconsistency normally occurs after the data updating.

This research analyzed three main types of procedures of land use data updating, as well as reasons causing inconsistencies. The purpose of analyzing updating procedures is to find the principle of why and how the inconsistencies were created. Finding and sorting inconsistency cases to extract general rules is one of feasible methods to detect and prevent inconsistency. Three main categories of land use data inconsistencies are defined as the following: geometric, attribute and logical inconsistencies. Based on analysis of land use data and three main land use data updating methods, a set of rules were established to detect 9 types of land use data inconsistencies. A prototype system has been developed to implement and test the detection rules to see whether the rules can find out inconsistencies in land use data. The prototype system is expandable as it is developed by DLL coding technique. The advantage of this method is that new rules and functions can be added into the prototype system conveniently for other research purposes.

1.2 State-of-the-art research in land use data

Socio-economic activities require building a land use database to make better use of land resources in current or future development. As the world is developing rapidly, new maps are needed to present new situation of land use. Surveying and building a new land use database frequently costs too much, as socio-economic activities developed rapidly. Therefore, updating the current data is another option. Updated data will be the base data in the next updating procedure, and this procedure can be repeated to make the database up-to-date. However, errors or inconsistencies will be accumulated during this procedure. Typical example for error accumulation is that a wrong area number of land will be used or calculated in the next round of updating and this wrong number will keep influence the quality of the data. A number of studies have pointed out the land use conflict problems in the United States and Europe, which are caused by the land use plan or the conflicts of land use alternatives.

The reasons, consequences, and management of land use change have become one of the major topics of current society. Urban/nonurban land use setting and urban/rural interface are major interests. Urban land use and land use change are major concerns as it almost affects every socio-economic activity. For instance, it builds an environment where cities develop and employments happen, also affects the natural environment, in terms of air quality, water quality, water supply and natural hazards. Fundamentally, land use parcel is so important that it affects socio-economic activities from the individual to the community level. Land use change reflects the evolution of natural land surface and expansion of urban activities. One decision made by single person of land department could affect many other people’s life and the natural environment.

Rational land use planning and changing are very important. Contradictory land use policy inconsistent with land use data is one of major factors which leads to the poor planning or land use change. Inconsistency in land use data may due to the long procedure of data updating or the planning policy (Goetz et al., 2005). Therefore, inconsistency in land use data must be taken account of to avoid the negative influence of updating.

Land use research on urban expansion in the United States has been started since 1960s. However at present, U.S. focuses on farmland change management much more than on urban land use change, because the pressure of shifting farmland is more than that of urban in U.S. (Libby et al., 2005). Changing of land use patterns is recorded in National Resources Inventory (NRI) of USDA, which is a consistent and reliable database. The database describes how land is being used at the specific time point. Land use data are collected from the 48 contiguous states and Hawaii. As the ground truth being recorded on satellites images, NRI focuses on the categories of land use which can be easily collected from remote sensing images.

Land use planning should be rational and consistent. The consistency can be a logical result of interaction between land use and human activities. Studying on history data of U.S. shows that socio-economic activities, farm programs and policies will affect the changing of land. The higher crop price leads to farmland increasing, population growing, and urban area expanding. Analyzing rural housing land use data can reveal cultural trends and general economic living conditions. For an example, farmers pursue smaller house, and meanwhile land use data shows many of the houses located at the urban fringe, which illustrates a trend of rural living: as economy improved, people are looking for social independence and development space. Therefore, if economy is growing, yet urban land is decreasing, there is an inconsistency of land use data in data updating. Similarly, when the food price is increasing but farmland is decreasing, there is an inconsistency too.

Land use data also affect human activities and policy planning. Analysis on farmland in land use data of each state at U.S. shows that developed farmland area is not in proportion to the total area of state. Therefore there is less possibility that farmland distribution change will threat the food production. However, its allocation change does affect the available composite services for rural living, and then affects the decision making of policy. Therefore, related services are consistent with the changing of the farmland location.

Land use problems of Europe are more complex and have their unique characteristics because of larger area and diverse land use. In Europe, there are two major issues about land use. The first one is land developing. For instance, in England, governments try to open hill-land and moorland for agriculture. The second one is land reform, which is consideration of transaction on land ownership to a new landlord in open market (Hanley et al., 2005).

Land use data of Europe more focus on land owners and users. For farmland, the first priority of land owner should be the farmer who works on the land. This rule obeys the idea of protection of farmland uses and food supply. Therefore, during the land use data updating, if there is a conflict between farmland owner and user, an inconsistency may occur.

In 1970s, concerns were most expressed in land conversion from cropland to urban uses in U.S.. Currently, concerns are expressed in rural/urban conflicts, i.e. the interface of urban and rural lands (Barnard et al., 2005). Reasons could be population increasing and local food supply pressure. The tradeoff between urban and rural land uses, i.e. structure and size of growing cities, is influenced by the population growth. Population growth is the force that leads to land use changes at urban fringe. Recently, most cities grow outward in U.S.. Therefore, land use changing may be more frequently at the edge of urban and rural land. An inconsistency for this changing is easily to modeling. For the land located at the urban/rural interface, if food supply became the major problem, these lands may be switched to services land which can be used for food production. On the contrary, if population is growing fast or housing supply became the main pressure, lands may be changed for housing hold or a new residential land as the trend of city growing is outward.

The problem is that market cannot manage land use perfectly by itself. Consistent and rational land use needs planning and analysis of accurate data (Seidl et al., 2005).

Land provides a series of goods and services to human society, which is dependent on land use. Under the marker rules, as its goal is always about profit, it may conflict the rational land use in countries without controlled market.

In the developed areas, markets affect the allocation of land resources. For instance, market determines food price and therefore affects how much land will be allocated to agriculture purpose. Similarly, demand for housing supply will affect how many land will be allocated to housing. As mentioned before, market will affect land use changing and some of those may be conflicted with current land policy. If accurate updating of land use data can reflect the rise and fall of market, then with the reference of food market and household market, updating result of land use data should be consistent with current market situation (Bergstrom et al., 2005).

Urban development affects ecosystem, both wildlife and human habitats. Human activities also affect land use change. Interaction between land and human is complex (Marshall et al., 2005). In land use data updating, changing of land use attributes should be consistent with relative developing activities.

Land is immobile; hence land use data should describe its location. Location of land uses has some common rules, which reflect the logic of human life (Bell et al., 2005). For instance, residential land parcel may depend on its proximity to transportation network, schools, employment center, shopping opportunities, parks, and recreation areas. A land parcel’s proximity to these items may change over time as some new facilities are built or old facilities were shut down. Lands provide not only single service but also the services flow. Services may solely rely on the attribute/use of a specific land parcel, while most likely they will depend on a suite of land parcels. One single land parcel produces services to itself as well as the surrounding parcels.

Land use data are spatial data, which have spatial configuration and location. Land use map shows ownership’s boundary and describes features of land parcel such as area, and sale price. Basic land use map is a parcel-level land use data, because parcel directly shows the situation of land use units. Historical data allow users to recreate the land conversion; however, those data are not consistent, which makes land use analysis complicated. Inconsistency problem is an important issue of land use data.

1.3 Objectives

The main objective of this study is to develop the methods for detecting spatial inconsistency in land use data updating based on geometric, attribute and relationship information of land use data.

In traditional land use surveys, the findings are commonly presented in the form of large-scale to medium-scale maps. According to the information of land use classification, area, spatial distribution and ownership, organizations will perform their management and planning in a local, regional or even national scale. With the advance of spatial technology, updating land use maps is no longer merely rely on on-site surveying/observation, but can also be supplemented with GPS (Global Positioning System) for location determination, aerial photos or satellite imagery for fast tracking of changes, especially over a wide and remotely-accessed territory, and GIS (Geographic Information System) for speedy data manipulation, visualization and production of various thematic maps. However, accuracy of the map derived from primary or secondary data sources is a major issue. Current solution of data quality control can maintain accuracy of data format and structure after updating, but not handle the accuracy of spatial meaning of data very well.

In this study, a new method is proposed that information of data such as geometric, attribute and spatial relationship can be used as reference to extract rules for detecting spatial inconsistency of data, which are useful to inspect the meaning of data and the consistency. Inconsistency detection can be implemented through comparing or calculating geometric, attribute and relationship of land use data.

1.4 Methodology

Basically, there are two main steps we used to implement this study. The first step is to extract rules to establish the standard of consistency of land use data. Then, the second step is to identify whether the land use data obey the rules. The identification elements are geometric, attribute and spatial relationship of land use data.

As land use data obey their own classification standards, in this study the China Mainland land use data were chose as an instance. Rules and analysis described in this study are all based on this land use classification standard. The study analyzes information of spatial geometric relationship, attribute and logical relationship of polygons in land use map to extract consistency rules that the map should obey. The study also analyzes updating method of land use map to detect inconsistency. Based on the two points above, inconsistency detection rules might be extracted as a reference to check land use data. Experiment method is to use ArcEngine 9.2 and C# to program inconsistency detection algorithms and apply them in a prototype system.

1.5 Organization of whole thesis.

The first chapter of this thesis is the introduction of the study. In this section, it will state why inconsistency detection is necessary in land use data updating. Besides, current studies on inconsistency are reviewed, and a new method to solve the limitation of current methods is proposed. The second chapter states two main points: The first one is about three main methods currently used in land use data updating. This is helpful to find out the reasons for inconsistency and rules of detecting inconsistency. The second point is to introduce the current methods for quality control on land use data. The third chapter states the rules and method that will be applied on inconsistency detection in land use data updating. The forth chapter states how to design and implement an experiment system to test the proposed method. The fifth chapter states the experiment result and analysis on inconsistency detection in land use data updating. Finally, the sixth chapter is the conclusion and summary of the study.

2. Existing methods for detection of spatial data inconsistency

2.1 Main methods for updating land use data

The procedure of updating land use data in China is used as a demonstration, as its three main updating methods have been used widely: site surveying, change detection via remote sensing images and the combine of the above. The procedure of updating land use data at national level is complex in China. Since China is a large country, it is difficult for the country to arrange a national level land investigation and update land use data by a single department or institution. Generally speaking, as shown in figure 2.1-1, the solution is to divide the national task to the smaller ones for land resource departments or local surveying institutions at province or city level (Zhang et al., 2005). Local department will arrange and implement the land use updating task by using their own methods and techniques. Instead of managing updating job by single leading department, implementing task by local may cause inconsistency problems in China.

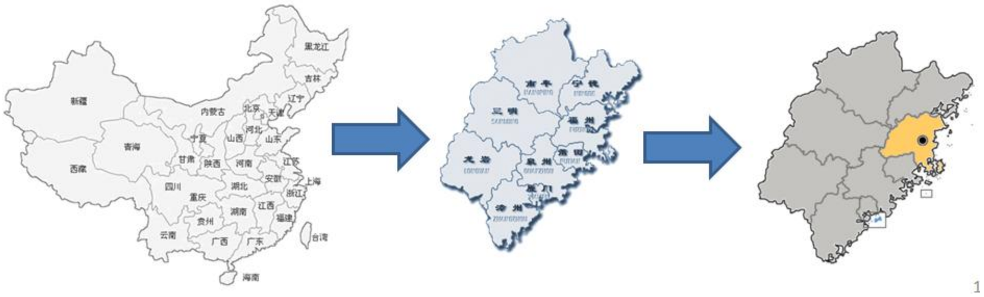


Figure 2.1-1 Dividing national task into smaller ones, then gets support from local government.

Local land updating methods used by different surveying institutions are not always the same. Three main methods are applied: totally site surveying, change detection by remote sensing image and the combination of them.

Traditional site surveying is the most reliable method for updating land use data. However, site surveying is very expensive and time consuming. For some regions, such as forest, no man's zone and so on, delivering surveying crew there is unreasonable and dangerous. Otherwise, site surveying might be the best option for the local institution to update land use data. For instance, Fujian province applied site surveying combined by a part of interior mapping work to implement land use updating task. The work flow has been summarized as follows (Li et al., 2005):

1. Using geometric rectification and image fusion techniques to get orthophoto map (DOM) of interested area from high resolution and up-to-date remote sensing image.

2. Overlapping administrative boundary and placing name mark information onto the DOM to get the site surveying base map.

3. Using the base map to implement site surveying to get land cover type and distribution location information of each land. Considering the convenience of applying site surveying, land use data were recorded in the point format, instead of the polygon format.

4. Using interior mapping to update old land use map data, based on updating information collected from site surveying.

The main working flow is like the figure below:

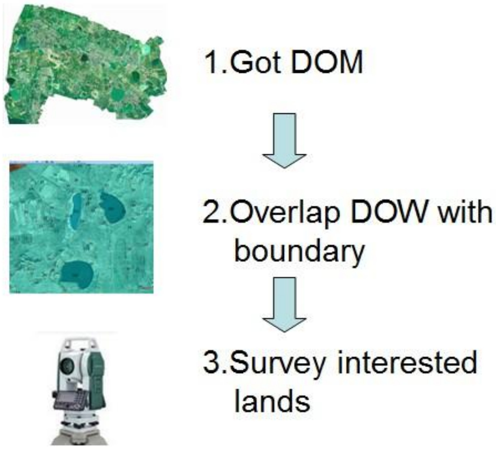


Figure: 2.1-1.2: Working flow of site surveying updating method

The second main method of updating land use data is using change detection techniques based on up-to-date remote sensing images, as depicted in figure 2.1-2. As shown in figure 2.1-3, images can be overlapped by vector data for fast updating. Using change information to update land use data is cheap and fast. Change information is greatly significant and useful for updating land use data automatically (Xiao et al., 1999; Rong et al., 1999). The process of extracting change information is so called change detection. Change detection is to identify differences in one area by interpreting the image collected in different time (Singh et al., 1989). Applying change information to old data can get the new version of data. Change detection techniques can be divided into two main methods: pre-classification and post-classification change detection. Pre-classification methods compare the image directly before classification to find out the change information. Post-classification methods are applied to the classification results of satellite images. Post-classification methods are useful for data sources which are difficult to combine before classification. Matching or overlapping the change detection result with existing data will show where has been changed or need to be updated. Since this method is highly automatic, the result may have various potential inconsistencies.

Figure 2.1-2 Detected changes based on up-to-date remote sensing images detection techniques

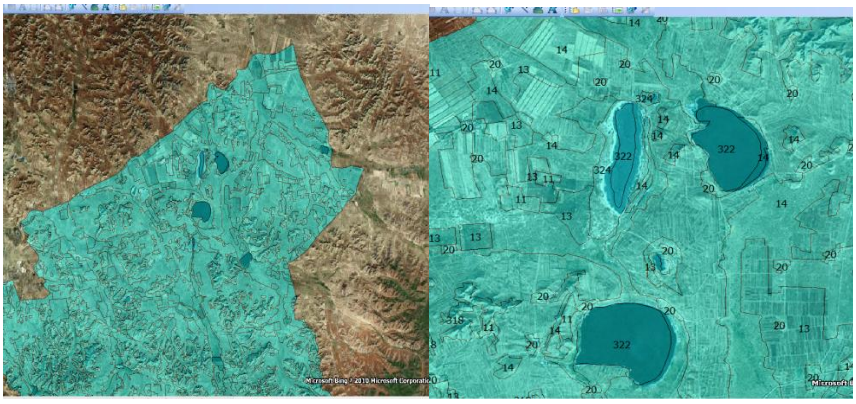


Figure 2.1-3 Images overlapped by vector data

The third main method is the combination of site surveying and changing detection, which is the most popular and acceptable one. This method is less difficult and faster than totally site surveying but much more reliable than changing detection. Therefore, it is used by majority of local surveying departments, such as surveying departments in Hebei province and Beijing, to update their local land use data. Experiment land use data of this study is produced by this updating method. Work flow of this updating method can be summarized as follows (Tang et al., 2006):

1. Get orthophoto map (DOM) of interesting area via DEM and up-to-date remote sensing image. Then the DOM is overlapped by administrative boundary, place name mark information and topographic map (scale in 1:10000, for instance). This DOM will be the base map for interior interpretation and outside site surveying.

2. Overlap the DOM indicated above to old land use data map for interior interpretation. As the current automatic remote sensing image classification is not reliable enough, changing area is extracted and digitized manually via man-machine interaction. Boundary information is digitized and attribute of change data are recorded via human interpretation.

3. Overlap the change area data to DOM, and then use this new DOM as the base map to guide site surveying. Site surveying can confirm the result of human interpretation and get extra land updating information such as new line objects, point objects, and other details which are valuable for updating.

4. Based on the site surveying confirmation result, establish changing data map database and update old land use map.

The combination method is different from complete site surveying. Complete site surveying conducts site surveying first and then applies interior mapping. The combination method carries out interior interpretation first and then uses site surveying as a confirmation. This method is inexpensive, more efficient and reliable. The key point is that no matter which method used by the local surveying institution, the final land use updating result and land use data format must obey the national standard. For instance, for the latest “Second National Land Use Investigation”, China give a set of standards which include land object classification code, layer name format, attribute table structure and so on (Lin et al., 2001).

These three main updating methods are used in the large area. For land use updating in city level, there are various local updating method which is much easier than that at nation level.

2.2 Existing methods for spatial data quality control

From a general perspective, the main reason for inconsistency problems occur in China land use data updating is to split it to small tasks. Chinese government divides the national or region updating task to local surveying institutions. Although the final data standard is the same, technology abilities of various local surveying institutions are different. For instance, using totally site surveying will get more accurate and reliable updating information than using totally remote sensing change detection. Besides, local institutions do not implement the updating task simultaneously. The land use updating data at the final national level are compiled from many smaller local updating databases which have different reliabilities and there may be long time gaps between each other (Zhang et al., 2008; Zhu et al., 2008). As shown in figure 2.2-1, when land use maps are collected from different local surveying department, it is easy to cause inconsistency on boundary. As the updating task is divided by administrative boundary, instead of natural boundary of land, it may cut one land polygon into two different surveying tasks. The direct consequence is that a single land will be divided into two lands by administrative boundary. In national database, these two lands should be merged to a single land.

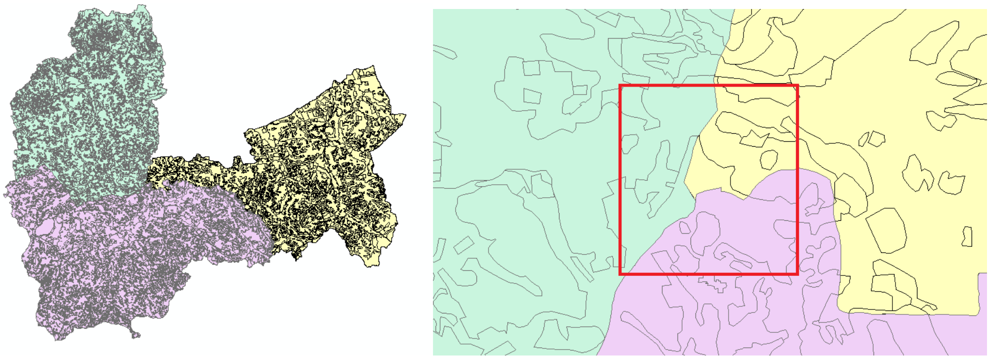


Figure 2.2-1 Local maps were merged into national map

Another quality control problem in national land use data updating is that the updating purpose of local updating institutions is different, as local institution will collect and update land use data based on their own benefits and data application. However when building the database, all city level maps will be combined as a national map. Maps from different cities may have conflict as the mapping purpose and mapping standard are different.

Besides of above reasons, there are three main reasons for causing updating inconsistency:

1. Source data contains error. Source data includes base map used in site surveying, DOM and remote sensing image used in interpretation and so on.

2. Human errors. Such as false interpretation on remote sensing image and mistake judgments in site surveying. Besides that, even for the same site viewing or image, different experts or surveying crew may have different opinion on the same land object.

3. As site surveying and combination method involved both outside task and interior mapping task, data transformation between surveyed data and map data, especially map projection and format transformation, it will cause inconsistency problems.

As various updating methods are used by local surveying institutions while final data should follow a uniform standard, this study focuses on inconsistency when standardizing final land use data. Previous studies have developed a set of inconsistency handing methods to maintain the quality of land use data updating results. The main reasons causing inconsistency in site surveying method is the transformation works between interior mapping job and outside surveying job (Li et al., 2008). For totally site surveying updating method, three gaps occur during transformation from surveyed data to map data.

1. The outside surveying collects point data which contain information of land use types and location, but interior mapping needs maps in polygon format. Point maps need to be translated into polygon maps.

2. The outside point data were recorded in one data layer, while interior mapping data have different layers for different land objects.

3. Site surveying base map was separated based on standard shape which is the natural boundary of land, while interior mapping data should be separated according to the administrative boundary.

Transformation of data from surveyed data to map data needs to match point location to polygon objects in old land use map, and then divide data to different layers, merge or separate polygons to match the administrative boundary. And inconsistencies may occur during these transformations. This checking method assumes that the point data collected via site surveying is correct, and then overlaps the original point data onto final polygon data. The attribute fields of both data are compared to find out whether records match each other. The advantage of this self-checking method is that site surveying do have higher accuracy than other methods. Therefore using original surveying data to check final data is reasonable. However the drawback of this method is, when point data from site surveying itself contains conflict or error, the inconsistency cannot be checked out by self-checking. Therefore, in order to upgrade data, the checking process needs some logical inconsistency checks to detect errors caused by human judgments. In this study, a set of logical attribute inconsistency detection rules is developed, which is useful for site surveying updating method. Another drawback of site surveying method is that it can only update the land cover types, but cannot follow the updating of the boundary of land.

A “4 Points” checking method is useful for general inconsistency check. Previous study indicated that a set of simple checking methods can be applied to final land use data after updating, no matter which updating method was used (Li et al., 2007):

1. Integrity checking: This will check whether all data layers and the attribute table are completed, and whether all land classification codes are correct.

2. Attribute table checking: This will make sure that the key field has unique value and all fields do not contain “null” value.

3. Topological checking: This will ensure that the boundary data do not contain gaps, intersections and overlapping.

4. Area logical checking: This will calculate the total area to see whether it matches the sum of every smaller polygon inside the bigger polygon.

Idea of “4 points” method is to provide the general checking functions which can be applied to all land use data. However, for some potential error, such as logical errors caused by false human interpretation, these functions are relatively weak.

There are some similar checking points, such as (Zhang et al., 2008; Zhu et al., 2008):

To check whether land polygons have been recorded twice.

To check the sequence of layers.

To check every number of attribute field to make sure their values are all not less than zero.

Based on these general checking methods, some advanced and special checking points were developed, such as to check whether area of the polygon is smaller than a threshold, and delete these records that are less than threshold as they were treated as redundancy (Zhang et al., 2009).

General checking methods for land use data updating is useful as it can find out some basic and common data inconsistency. It ensures that updating data are at least formally correct. However, the disadvantage of the general checking methods is that they are not specific for land use updating data. These methods can be applied for any theme data in polygon content. It is hard to detect superior or potential logical inconsistency occurred in land use data updating. A set of logical inconsistency detection rules is developed in this study, which can overcome this disadvantage.

General checking is so important that it has been paid close attention by government and surveying institutions. Some business companies also developed software to do the checking task. Such as TD-Checker developed by GeoWay, which is especially used for quality control of land use data. Checking software implements and upgrades all the general checking functions, and it also presents users friendly interface and automatic task processing method (www.geoway.com.cn). This kind of software is the popular quality control solution for official land use data and is used in major land use data collection and updating projects in China. However, less attention is paid to detect logical and potential inconsistency.

Rule-based checking method is a reasonable and well-performed framework to handle inconsistency in land use data collection and updating task (Hong et al, 2009). Although a framework cannot describe specific checking and algorithm in details, it can present a comprehensive idea of work flow.

This idea can be applied over any final land use data after updating to implement a comprehensive and systematic inconsistency detection task. Basic formula for inconsistency detection is:

land use data + standard + element = detection result

Standard of land use data to be checked is a set of requirements of data content and quality, such as “Specifications of China Second Land Use Investigation Database”. The standard specifies the land object classification system, land type code reference table, theme layer classification system, attribute table structure and so on. It is the criterion to judge whether the land use data contains inconsistency. Element is a set of checking rules. Rule is the abstract concept which describes possible error or inconsistency. Rule is also the unit of inconsistency checking and the real entity who implement the checking task. Rules are normally divided into spatial and attribute types, while a new type of rules is added to this study: logical rules. For a reasonable check solution, it should give a standard interface for adding new rules. Unit rules can be combined together to compose advanced rules via “and” and “or” relationship. An instance is that “check area number” can be composed by “checking area field” and “not zero checking”. Purpose of this design is to meet the requirement of different land use data checking tasks which have different features, such as land use data updated via different method would contain different inconsistencies. Another benefit of this design is that it can provide a flexible inconsistency checking method.

Rule-based checking idea can be used as a basic reference to establish land use data inconsistency detection functions, especially useful for designing advanced inconsistency detection algorithm, as they are more difficult to be implemented than general checking functions.

Besides above applications on land use data inconsistency in China, some theoretical studies in oversea have also been done.

Achieving and preserving integrity of data is very important. However, within the scope of geographic applications, special problems come up due to the location of data (Plumber et al, 1997). In the spatial domain, one idea is to integrate constraints into database systems by means of queries (Gaede et al, 1996). Queries can be interpreted as a constraint; each query specifically represents a constraint to make sure the objects satisfy the requirements.

The major types of integrity constraints frequently used in database modeling are: domain constraints, key and relationship structural constraints, and general semantic integrity constraints. Extended constraints for spatial data are classified in order to encompass the peculiarities of spatial data. This classification is based on the distinction between topological, semantic, and user rules, as follows (Cockcroft et al, 1997):

Topological consistency constraints: Topology is the study of geometrical properties and spatial relations. The OGC (Open GIS Consortium), an organization dedicated for developing standards for spatial operations and spatial data, defines a set of standard topological relations: disjoint, overlaps, touches, contains, within, crosses and equals. There have been some related studies about the principles of formally defining spatial relationships. These principles can be applied to specific entities and relationships to provide a basis for inconsistency control. One example of this constraint is area subdivision. Administrative region of a city must be contained within the city limits, and there must not have any spot in the municipal territory that belongs to more than one administrative region or to none. Topological consistency describes the relationship of two or more lands. This consistency may relate to local land use policy. Such as a school land cannot touch a hospital land, a road land must touch a vegetation buffer land. An instance of this situation is like below:

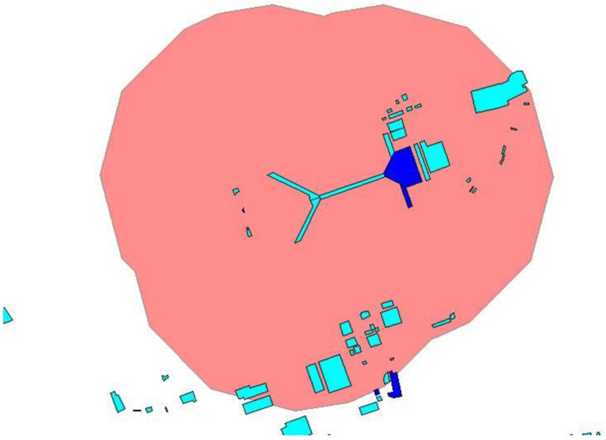


Figure 2.2-2: Logical height of buildings near airport must be lower than a limited value

Lands showed in Figure 2.2-2 are the building lands near airport area. Local government policy may give a limited height value of buildings near airport. Red area in figure 2.2-2 is the buffer area of 1.5 kilometers away from airport. With the help of calculating “touches” relationship, buildings in dangerous height may be found, which are displayed as dark blue in figure 2.2-2. Here one of the two dark blue buildings is the terminal itself.

Semantic consistency constraints**:** These constraints focus on the meaning of geographic data and entity. Semantic inconsistency constraints can be applied to database. An example of this constraint is the rule that does not allow a building to be intercepted by a street segment.

User defined consistency constraints**:** User defined integrity constraints allow database consistency to be maintained, as defined by the equivalent of business rules in non-spatial database management system. This type of constraint acts, for instance, on the location of a gas station, which for legal reasons, must lie farther than 200 meters from any existing school. The municipal permitting process must consider this limitation in its analysis. User defined rules may be stored and enforced by an active repository.

Iran scientists proposed a methodology of rule-based spatial query to detect inconsistency in land use data (Fathi and Alimohammadiet, 2010). Within a geographic context, topological relations and other spatial relationships are fundamentally important in the definition of spatial consistency rules. A number of consistency constraints must be observed when or after updating a database, to preserve the semantics and the quality of data. Constraint rules can be stored in rule repository. During the application of rules, they can be applied on data by making spatial queries, which is a rule-based spatial query. A spatial constraint rule for hospital and factory, for example, can be defined as: any hospital should be disjoint with any factories. A spatial constraint for island and water is that, for example, every Island has a within relationship with only one water body.

Rule-based spatial query should be in the form of "**IF** condition-pattern **THEN** actions". Instances are stated as follows:

IF *(land use is educational)*
THEN *(must not be neighbor by industrial Land use)*
IF *(land use is hospital)* ***AND*** *(has bedridden section)*
THEN *(must not be neighbor by educational land use)*

One typical example is an inconsistency rule for airport land use:

In the range of 500m around LOCAL CENTRAL TRANSIT [LCT] (Centers of Communication Company), the height of buildings should be less than 50m. In the range of 6000m up to 15000m from the end of airport, the height of any building must not be more than 150m. School site should not at airport nearby (school site within two miles of airport runway).

Similar land use consistency rules can be found in land use planning reports, such as one of Snohomish County of Washington (Jones and Stokes, 2005). In this report some land use consistencies were mentioned. For an example, agriculture land should be close to river network, and rural resource should be adjacent to rural residential land.

In this local land use case, some situations were considered as inconsistency. For instance, a new industrial land is next to a low density land, but local government policy requires that new industrial, commercial and public facilities land uses should include vegetative buffer. In some cases, inconsistency land uses locate on either side of roads, depending on the traffic volume and its width.

Another inconsistency case is that: low-density residential area that is adjacent to a medium-density residential uses is considered as an inconsistency.

Government report of land use planning normally makes some local land use consistency constraints when the plan is applied. For instance, a set of local land use policy consistencies is made in the city of Oceanside, San Diego: noise source should be separated from residential area to avoid affecting health; existing mature trees should be retained wherever possible, which means land use change from forest to other uses is forbidden. For the land use configuration: sharp angles of intersection with natural terrain should be avoided. In potential flooding area, drainage is used to reduce its damage. If the area of such type of land parcels increases, an inconsistency occurs.

3. Development of Rules for inconsistency detection

Before we start to develop the rules, one thing must be paid attention to: The rules we designed in this study are only available under the land use classification system which the experiment data obey.

Detail content of this land use classification system can be read from Table 3.1-1: Land Use Classification System.

This classification system is the temporary version of national land use classification standard issued by China Mainland government. It was widely used in the second national land investigation, which was a project to survey all the lands in China mainland. The experiment data used in this study came from this project and used this land use classification system as standard. Analysis on this classification can help us to extract and form the concept of a piece of rule. Such as:

In China land use updating in military area is confidential. Briefly speaking, this kind of land use changing is hard to be detected and updated by civil land use data. Therefore, it is reasonable to doubt one land with code 281 in land use data, because code 281 represents land for military installations. Furthermore, this concept can be developed into a piece of rule.

Similar rules can be developed based on analysis of data, experience of people and interpretation of land use classification system.

However, based on the situation of experiment data we used in this study, we found that most area in experiment district included several kinds of farmlands, residential lands and fewer unused lands. There were not special lands in this area, i.e. the land of Class 28, nor any commercial lands and services sites, such as land of Class 21, because this area is just a normal town in north of China. Therefore, in this study we chose some rules focus on the town’s situation to develop and implement. Analysis on classification of inconsistency and understanding of data is the fundamental basis of rule developing.

Table 3.1-1: Land Use Classification System

Land Use Classification of China Mainland (Temporary Version)					
Third Class		Second Class		First Class	
Class Name	Class Code	Class Name	Class Code	Class Name	Class Code
Irrigated Paddy	111	Cultivated Land	11	Farmland	1
Rained Paddy	112				
Irrigated Cropland	113				
Dry Land	114				
Vegetable Plot	115				
Orchard	121	Orchard and Perennial Plantation	12		
Mulberry Field	122				
Tea Garden	123				
Rubber Plantation	124				
Other Garden	125				
Woodland	131	Forest Land	13		
Shrubland	132				
Open Forest land	133				
Young Afforested Land	134				
Slash	135				
Nursery	136				
Natural Grassland	141	Grassland	14		
Improved Grassland	142				
Artificial Grassland	143				

Land Use Classification of China Mainland (Temporary Version)					
Third Class		Second Class		First Class	
Class Name	Class Code	Class Name	Class Code	Class Name	Class Code
Confined Feeding Operations	151	Other Farmland	15	Farmland	1
Green House and Aquatic Operations	152				
Country Road	153				
Pond	154				
Aquiculture Area	155				
Irrigation Land	156				
Riser of Terrace	157				
Grain-Sunning Ground	158				

Land Use Classification of China Mainland (Temporary Version)					
Third Class		Second Class		First Class	
Class Name	Class Code	Class Name	Class Code	Class Name	Class Code
City	201	Residential Land, Industrial Land and Mining Land	20	Land of Construction	2
Organic Town	202				
Rural Residential Area	203				
Industrial Land and Mining Land	204				
Salt Field	205				
Land for Special Uses	206				
Commercial Land	211	Commercial Land & Service Sites	21		
Land of Business and Finance	212				
Land of Accommodation and Catering	213				
Other Commercial and Service Land	214				
Land of Industrial	221	Land of Industrial, Mining and Warehouse Space	22		
Land of Mining	222				
Land of Warehouse Space	223				
Land of Public Infrastructure	231	Land for Public Facilities	23		
Public Garden & Greenbelt	232				

Land Use Classification of China Mainland (Temporary Version)					
Third Class		Second Class		First Class	
Class Name	Class Code	Class Name	Class Code	Class Name	Class Code
Land of Government Agencies and Organizations	241	Land of Public Construction	24	Land of Construction	2
Land for Education	242				
Land for Scientific Research	243				
Land for Recreational and Sports Activities	244				
Land of Medical Treatment	245				
Land for Charity	246				
Land for Residence Only	251	Residential Land	25		
Land for Residence,industry and commerce	252				
Countryside Homestead	253				
Free homestead	254				
Land for Railway	261	Land of Transportation	26		
Land Used for Highways	262				
Civil Airport	263				
Land for Harbour and Wharf	264				
Land for Pipage	265				

Land Use Classification of China Mainland (Temporary Version)					
Third Class		Second Class		First Class	
Class Name	Class Code	Class Name	Class Code	Class Name	Class Code
Reservoir	271	Land of Water Conservancy Facilities	27	Land of Construction	2
Land for Hydraulic Architecture	272				
Land for Military Installations	281	Land for Special Uses	28		
Land of Embassies and Consulates	282				
Land of Religion	283				
Land of Prison	284				
Land of Funeral	285				
Unused Grassland	311	Unused land	31	Unused land	3
Saline-Alkali Soil	312				
Marsh Land	313				
Sandy Land	314				
Bare Land	315				
Bare Rock & Gravel	316				
Other Unused land	317				
River	321	Other Land	32		
Lake	322				
Reed Marsh	323				
Mud Flat	324				
Glacier & Permanent Snow Cover	325				

3.1 Classification of inconsistency

Basically, inconsistencies are divided into three types: spatial/geometric inconsistency, attribute inconsistency and logical inconsistency.

Attribute data inconsistency is the error or conflict recorded in data attribute table. “Null” value recorded in key column of the table, illogical or unreasonable value of area or perimeter numbers are some common instances of attribute data inconsistencies.

Spatial data inconsistency is spatial related or location related inconsistency. For instance, one single polygon in land use classification data represents a land with one type of land use code. Therefore, a reasonable deduction is that two adjacent polygons should have different land use types. Yet, after updating land use change data, a typical spatial inconsistency is that two adjacent polygons have the same land cover type. Updating data without merging adjacent polygons with same type will cause data redundancy and will have low data quality. The reason of this problem is that adjacent polygons do have different land use types in old database but one of polygons or both of them have been changed to a same land use type. Unfortunately, current updating method does not take account of the topological relationship of updated objects, especially the neighboring relationship of objects. This is just one case of drawbacks of updating process that could cause spatial data inconsistency.

Logical data inconsistency is an imperceptible phenomenon. However, it has significant influence when data are used. Logical data inconsistency is special data inconsistency which has correct form with wrong meaning. Generally speaking, data with logical inconsistency is correct when only attribute and spatial inconsistency detection rules are applied. However, logical data inconsistency is an error because it overrides common sense. For instance, desert contains oasis or artificial irrigated land is possible, but, if a large desert contains an orchard or hydraulic structure in the land use data, it obviously overrides common sense. Another example would be that a land changed from building land to farm land. In these cases key values in the attribute table and spatial information are all correct. Therefore, when applying attribute and spatial inconsistency rules to detect data inconsistency, the results may suggest the data is without inconsistency because they cannot detect the logical inconsistency.

To detect the possible data inconsistency, the information about type, feature, reason, and detect concepts are needed. Such information could be obtained from analysis on land use classification data and land use change data. Then detection rules can be established based on those analysis results. And then the rules can be implemented by programming method or data processing steps to detect data inconsistencies. These procedures are the basic methodology of this study. According to three main types of data inconsistencies, three sets of rules are developed, which focus on attribute, spatial and logical meanings, respectively. Data will be considered containing inconsistency when it breaks any of these three sets of rules. The results will be useful for data user and decision maker to avoid the wrong information in the data.

3.2 Rules for geometric inconsistency

Detection inconsistency in land use data relies on rule sets. To detect inconsistencies, the concepts of rules need to be transformed to algorithms or data processing steps. If the transformation is implemented via GIS software such as ArcGIS or MapInfo, this method is so called data processing steps. In this study, transformation of concepts to algorithms is implemented by programming a software system. The prototype system offers two main sets of functions which focus on attribute data detection and spatial data detection, respectively. Logical inconsistency detection functions are included in these two sets of inconsistency detection functions as: attribute related logic or spatial related logic.

In spatial related detection function set, details of the detection rules are as follows:

- Adjacent lands share same land use type lands without merging: As mentioned in the previous chapter: adjacent polygons in land use classification data have same land cover type but are not merged together after updating processing. This will cause data redundancy and confusing data expression. Here, rules are designed as: if the polygons have the adjacency topological relationship and have same land use type code in the attribute tables, it is considered as data inconsistency. By comparing the land use code and checking the topological relationship of the pair of polygon, this function can be implemented.

The land use data contain the land use code, which represents land cover information of the land. For example, as shown in figure 3.2-1.1, each land is labeled by class code to describe the land use information. Particularly, some adjacent lands (highlighted by red and yellow) have the same class code “14”.

In real world, polygons in red and yellow should be merged as single land to avoid data redundancy. The merged result should be like the colored land in figure 3.2-1.1., which only has one class code and one boundary.

This case is easily happening when updating data. In the old land use data, red land and yellow land may have different class code, and that’s why they had their own boundary. Surveying indicated that both lands are class 14 now. Therefore, after updating, both of lands in red and yellow are changed to class 14, without updating their boundary.

Red land: change from land in class x to land in class 14

Yellow land: change from land in class y to land in class 14

Rule of this case can be formalized as follows:

If (land A connect land B is true and land A.Class = land B.Class)

Then an inconsistency detected

However, such inconsistency detection needs more detail information. For instance, if the land use data contain ownership information, and red land and yellow land may belong to different owners, then they should not be treated as inconsistency and merged as a single land. Otherwise lands should be merged if all value is the same in the attribute table. Therefore, the rule can be advanced as follows:

If (land A connecting land B is true and land A.Class = land B.Class and carrying different attribute information is wrong)

Then an inconsistency detected

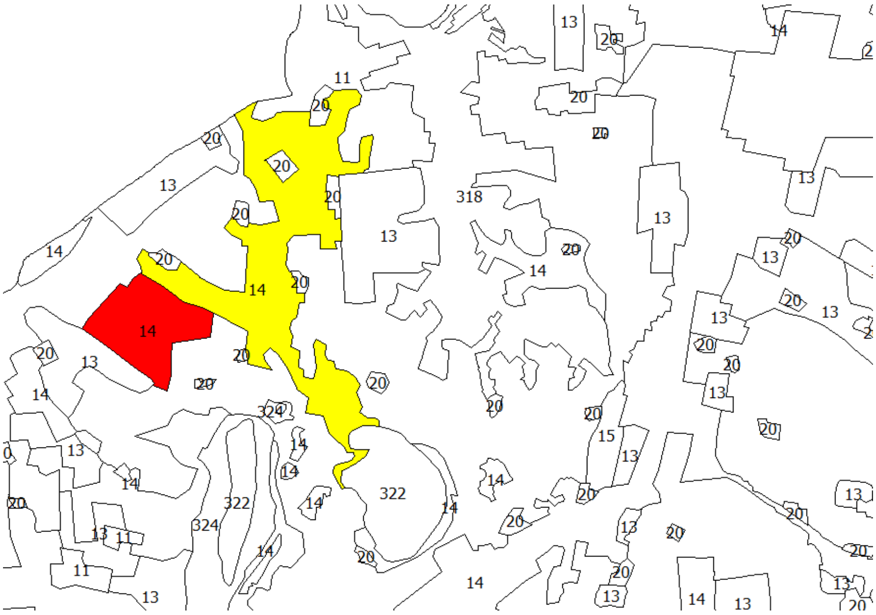
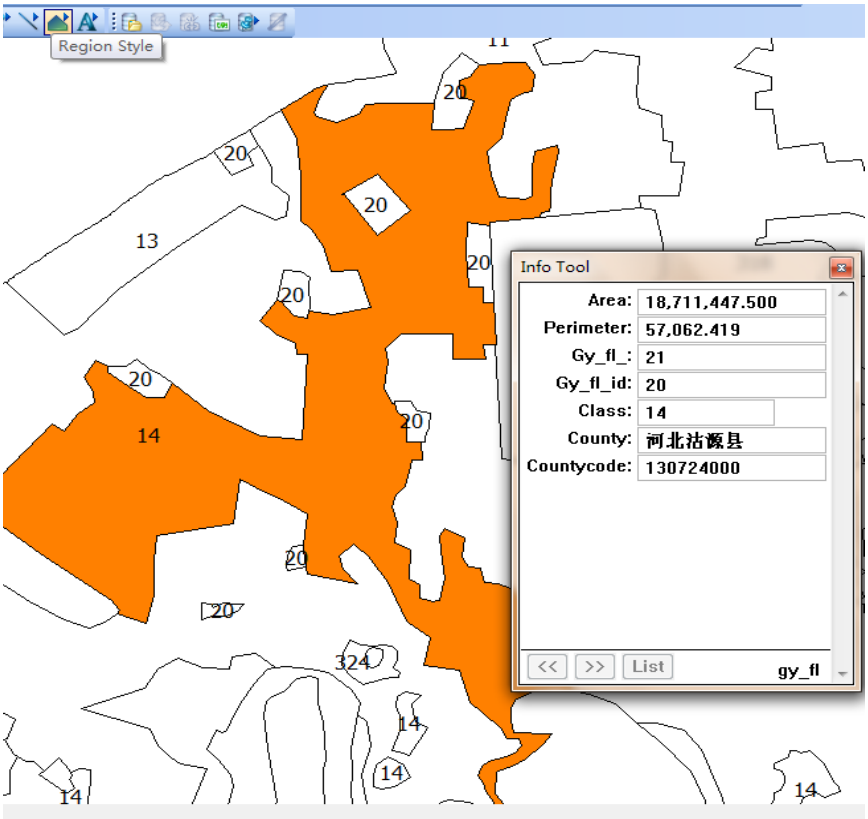


Figure 3.2-1.1 Polygons labeled with class code, and adjacent polygons have same land cover type but are not merged



Area: 18,711,447.500
Perimeter: 57,062.419
Gy_fl_: 21
Gy_fl_id: 20
Class: 14
County: 河北沽源县
Countycode: 130724000

Figure 3.2-1.2 Polygons labeled with class code, and adjacent polygons which have same land cover type can be merged

- Matching problems when overlapping land use change data and land use classification data: If a pair of polygons in these two types of land use data are overlapped each other but have different land type code value, this pair of polygons will be considered as possible inconsistency. Usually it is the last three digital numbers in change data is different with that of in classification data. As discussed earlier, change data indicated the change direction between two periods. Therefore, the code value of classification polygon should be consistent with the value of change polygons. Unsuccessful updating process or mistaken manual operation will result in this type of data inconsistency.

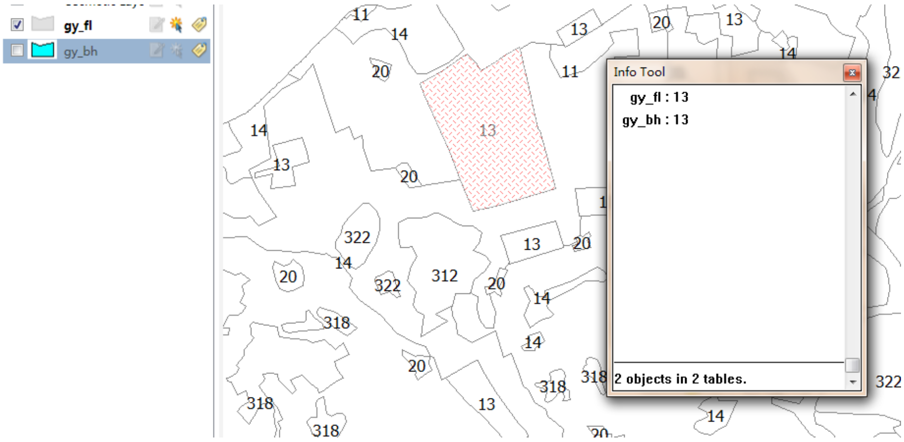
In land use data updating, change data is widely used as the base to produce new land use data. Land use change data have information about both of old land class and new land class. Old land class tells the land cover used to be and new class tells what the land cover is now. If the new land use data was correctly updated, its class code should be corresponding to that in change data. As shown in figure 3.2-2, one land was updated from old class 14 to new class 13 in change data, and its corresponding land in updated data should have class 13. Otherwise this would be an unsuccessful updating operation.

Before check this situation, the overlapping relationship of one land and its corresponding land is needed. For example, check one land in updated data, and then overlap updated data with change data, calculate the space relationship to find the corresponding land in change data and finally compare their class code to see whether it is correct. It is also useful to check merged land in updated data. Although lands have same class in updated data may merged as a single one, this merged data may overlap more than one corresponding lands in change data, and class of updated merged land should also be the same as new class of lands in change data.

Therefore, rule of this case is:

If (land A in updated data overlapping land B in change data is true and land A.Class <> land B.New Class)

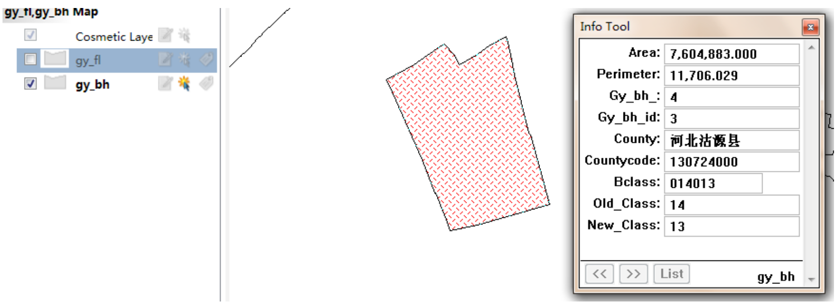
Then an inconsistency detected



gy_fl : 13
gy_bh : 13

2 objects in 2 tables.

Figure 3.2-2.1 Land class in updated date should obey the class of corresponding land in change data



Area: 7,604,883.000
Perimeter: 11,706.029
Gy_bh_: 4
Gy_bh_id: 3
County: 河北沽源县
Countycode: 130724000
Bclass: 014013
Old_Class: 14
New_Class: 13

Figure 3.2-2.2 Land class in updated date should obey the class of corresponding land in change data

- Area number of relative polygons breaks common senses: This rule is logical and spatial related inconsistency detection rule. Polygons represent area of land in various administration levels. The land use data surveying starts from smaller level, then the data are aggregated into larger units. For example, as shown in Figure 3.2-3, Jiangsu province consists of 13 polygons, each representing an administrative region. Each polygon consists of a number of sub-polygons. If the area of a town polygon is larger than the city which it belongs to, this is contrary to normal administration. For example, Polygon 11 consists of 8 sub-polygons. The total area of Polygon 11 should be equal to the sum of these 8 sub-polygons and the total area of Jiangsu province should be equal to the sum of these 13 polygons. If the size of any of the 8 sub-polygons in polygon 11 is larger than the size of polygon 11, then there must be a problem with the size of the sub-polygon or the size of the polygon 11. A rule can then be formalized as follows:

If (sub-polygon>polygon) then a inconsistency detected

Also, if the size of polygon 11 is not equal to the sum of the 8 sub-polygons, then there must be a problem with the size of the sub-polygon or the size of the polygon 11. A rule can then be formalized as follows:

If (sum-of-sub-polygons <> polygon) then an inconsistency detected

It is easy to cause such an inconsistency because survey task of sub-polygon and polygon may not be done by the same team or by the same method.

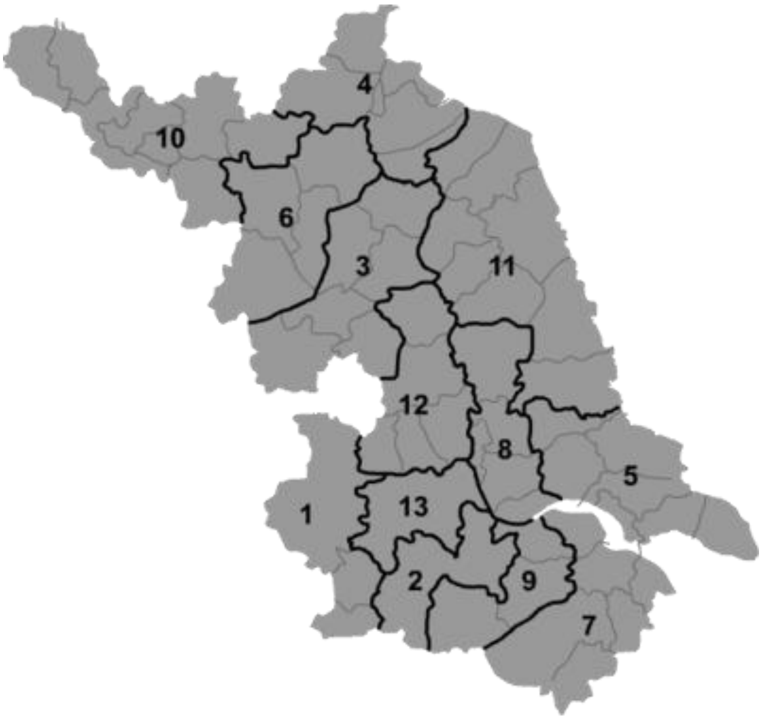


Figure 3.2-3 Jiangsu Province consists of 13 polygons; each polygon consists of a number of sub-polygons

3.3 Rules for attribute inconsistency

In attribute related detection function set, details of the detection rules are as follows:

- Sequence detection of vegetation evolution stage: This is logic and attribute related consistency detection rule. The evolution of vegetation follows specific stage and sequence. Normally, this procedure follows the steps as bare land, lichen, grassland, woodland and finally forest. If the evolution is not disrupted, then each stage should following this sequence. It is rarely to see a bare land change to a forest in 10 years. Therefore, the target of this detection rule is the attribute table of land use change data. As change data is produced based on personal judgment and experience, false interpretation occurs normally. In land use change data, if change direction violated the common sequence of vegetation evolution, this rule will consider that the record of data contains logical inconsistency.

Change data is the very important base of final product—new land use data. As depicted in figure 3.3-1, change data have both old and new classification information of the land. If both of them are vegetation land, then the evolution of the vegetation should be considered. Without human disturbance, natural evolution has its sequence and need certain time. For instance, it may take 20 years for grassland to become a forest, but only need 1-2 years for a cultivated land to become a grass land. It is not possible for evolution to happen in a very short time, unless under the influence of natural disaster or human disturbance. On the other hand, human disturbance can make vegetation type change quickly. For instance, after lumbering, a forest can become grassland in 1-2 years. But normally, it won't make it happen in a very short time. For example, it took the famous project--"Returning farmland to forests" 5-10 years to see vegetation recovering. As the highlight land shown in figure 3.3-1, vegetation in class 11 changing to class 13 may need 10 years in natural, but if the data updating cycle is 1 year, then it could be a problem.

Therefore, to check whether vegetation type has been changed in a very short time, some thresholds were set to test the rules. The summary of this rule is as follows:

If (land A.Old Class is vegetation type and land A.New Class is vegetation type and vegetation of old class need shorter time than new class to grow and updating cycle time> threshold)

Then an inconsistency detected

Area: 637,821.500
Perimeter: 4,292.081
Gy_bh_: 62
Gy_bh_id: 61
County: 河北沽源县
Countycode: 130724000
Bclass: 011013
Old_Class: 11
New_Class: 13

Figure 3.3-1 Vegetation growing stage sequence in change data

- Irreversible land type changing: Similar as irreversible vegetation evolution sequence, some land use change is logically irreversible. A typical instance is that a farmland can be changed to a building land. On the contrary, normally a building land is impossible to be changed to arable land, orchard or other agricultural lands. When the land use change breaks common sense, it will be considered as inconsistency.

Different from the vegetation evolution rule, this rule considers all impossible change of land. Normally, when a land changed to urban land, it can be shopping area, industrial area or residential area etc. But urban land would barely be changed to forests, grass or farmland etc, unless there are some disasters such as earthquake, flood, tsunami or typhoon etc. Without human disturbance, some land use changes are irreversible, such as returning farmland to forests, and desert become forests. This changing cannot happen by natural.

Irreversible changing has another feature that changing in reverse order may be correct. Such as, industrial cannot change to grass land, but the change of grass land to industrial is possible. For example, highlight land in figure 3.3-2, changing from class 13 to 14 may be possible, however, the old class is 14 and new class is 13, which is unreasonable

In this case, rules is like a mathematics set, a list of irreversible change is needed. The more factors are collected in this list, the more inconsistencies can be detected. The rules can be shown as follows:

Irreversible changing set:

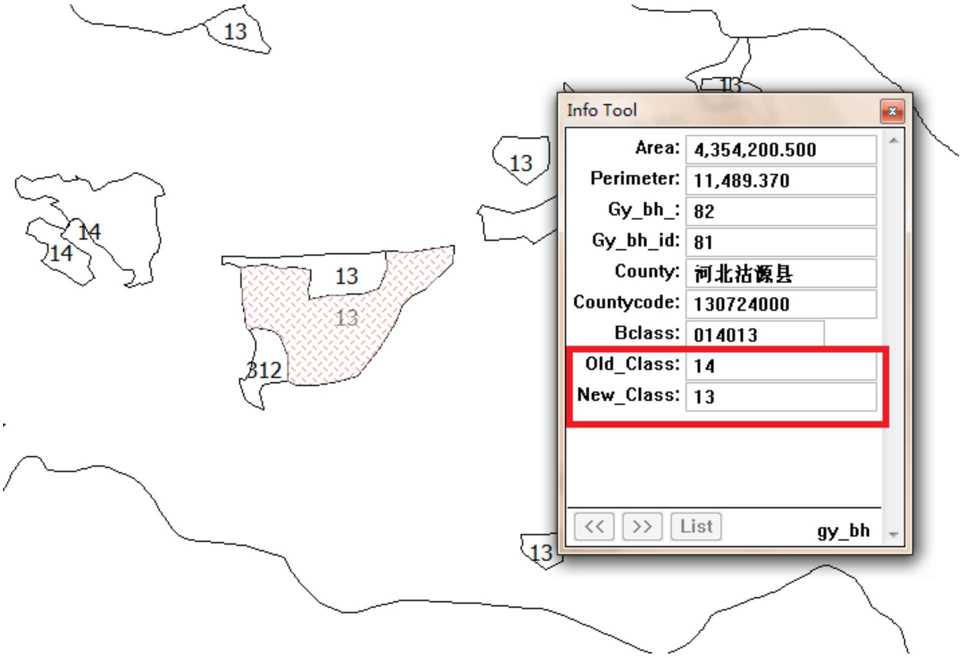
Class a1 cannot change to class array a1[i], a1[1]=class b1, a1[2]=class c1…

Class a2 cannot change to class array a2[j],a2[1]=class b2, a2[2]=class c2…

…

Rules:

```
Switch (Land A.Old class )
  Case a1:
    If Land A.New class is in Class array a1[i]
    Then an inconsistency detected
  Case a2:
    If Land A.New class is in Class array a2[i]
    Then an inconsistency detected
…
```



Area: 4,354,200.500
Perimeter: 11,489.370
Gy_bh_: 82
Gy_bh_id: 81
County: 河北沽源县
Countycode: 130724000
Bclass: 014013
Old_Class: 14
New_Class: 13

Figure 3.3-2 Irreversible land type changing

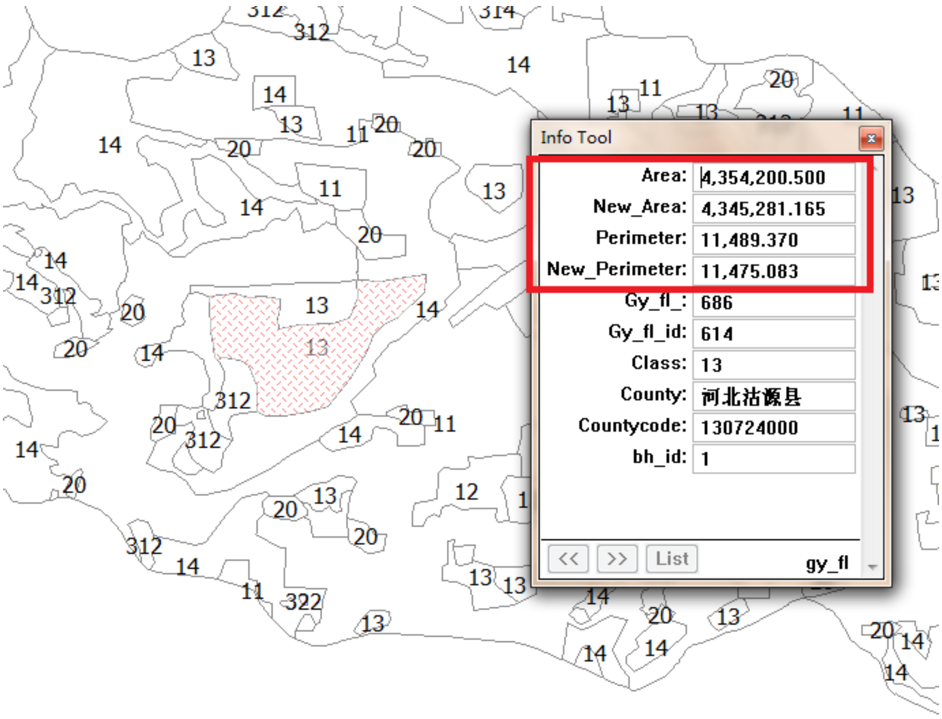
- Inaccurate area and perimeter key value: Both of these two land use data are matched under correct projection and coordinate system. However, after recalculating area and perimeter value of every polygon and comparing recorded in the attribute table and recalculated value, there are some inaccurate key values. If the discrepancy between two values is larger than the threshold, this rule will consider a record error.

Normally, inaccurate value or false value barely occurs in attribute table since value checking has been almost done in every data quality control. However, there are still some reasons which may cause it happen. When data were updated, they may be transformed into different formats or projection systems for further using. These data processing steps may cause inaccurate area or perimeter. As shown in figure 3.3-3, calculated value is not equal to the surveying value. Data user should pay attention on it to decide which value to be used. Difference between calculated vale and surveying value may range from very small to very large. Therefore, a threshold is needed as an acceptable margin of error, which depends on data user’s choice.

The rule is as follows:

If (Abs(land A. area - land A.Calculate area)> threshold of area Or Abs(land A. perimeter - land A.Calculate perimeter)> threshold of perimeter)

Then an inconsistency detected



Area: 4,354,200.500
New_Area: 4,345,281.165
Perimeter: 11,489.370
New_Perimeter: 11,475.083
Gy_fl_: 686
Gy_fl_id: 614
Class: 13
County: 河北沽源县
Countycode: 130724000
bh_id: 1

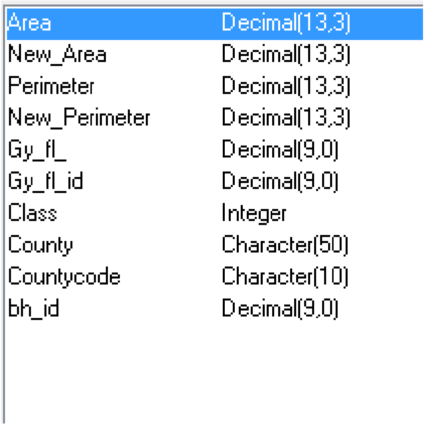
Figure 3.3-3 Inaccurate value of area and perimeter

- Attribute field inspection: Unacceptable attribute field values are detected by using this rule. The unacceptable values in the data attribute table include “null” value in the main field and repetitive ID value which should be unique.

Attribute data have their structure and rules, just like database do. As shown in figure 3.3-4, data structure defines the frame of attribute, it contains field name, data type, whether it is a key value or null value etc. Details of rules to check these inconsistencies depend on how many rules are applied on data attribute table. The more rules are applied; the more rules will be included in checking.

An example of checking rule to detect attribute field is as follows:

```
Do while record is not None
{
If(record[i].Key value is null)
Then an inconsistency detected
If(record[i]. non-repetitive value is in array non-repetitive value[j] )
Then
an inconsistency detected, non-repetitive value[j+1]= record[i]. non-repetitive value
…
i++
}
```



Area	Decimal(13,3)
New_Area	Decimal(13,3)
Perimeter	Decimal(13,3)
New_Perimeter	Decimal(13,3)
Gy_fl_	Decimal(9,0)
Gy_fl_id	Decimal(9,0)
Class	Integer
County	Character(50)
Countycode	Character(10)
bh_id	Decimal(9,0)

Figure 3.3-4 Attribute data structure of in land use data

3.4 Rules for logical inconsistency

In logic related detection function set, details of the detection rules are stated as follows:

- A city contains large area of agricultural land: This rule is used to detect a typical logic related land use data inconsistency. In land use classification data, if the topological relationship of a pair of polygons is “within”, while the container polygon has a land type of building land such as city, village or residential area, and the inside polygon’s land cover type is agricultural land such as farm or orchard, then this rule will consider this pair of data contain logical inconsistency because these phenomenon breaks common senses.

In land use data, contain-within relationship is very common. When one land contains another type of land, then the potential question is whether this contain relationship is reasonable. For an example, as the highlight lands shown in figure 3.4-1.1, the larger yellow land of class 318 contains the smaller red land of class 11. The rule can be used here to check whether an urban land contains a large area of farmland and the result will suggest that it is an inconsistency. If the yellow land in figure 3.4-1.1 is urban land and contains a large area of vegetation land, such as the small red land, it is possible that the red land is a public park or a big garden. However, if the red land is a farmland, then it may be an inconsistency. Farmland normally located beyond the area of urban land, connected by road net work or divided by some transitional lands. Land class changing can happen both in the container land and the island lands within a container land. If the inconsistency happened, both of the island lands and container land need to be checked to make sure the change is real.

Another factor may lead to inconsistency is the area of the island land. User can set a threshold to judge whether the island land is an inconsistency. If the suspicious island has a very small area, it can be merged with container land as one land use class, as shown in figure 3.4-1.2.The rule of this checking is as follows:

If(land A contain land B is true and land A.Class is urban and land B.Class is farmland and land B.area > threshold)

Then an inconsistency detected

When checking an island land such as the red one shown in figure 3.4-1.1, the rule should be like follows:

If(land A within land B is true and land B.Class is urban and land A.Class is farmland and land A.area > threshold)

Then an inconsistency is detected

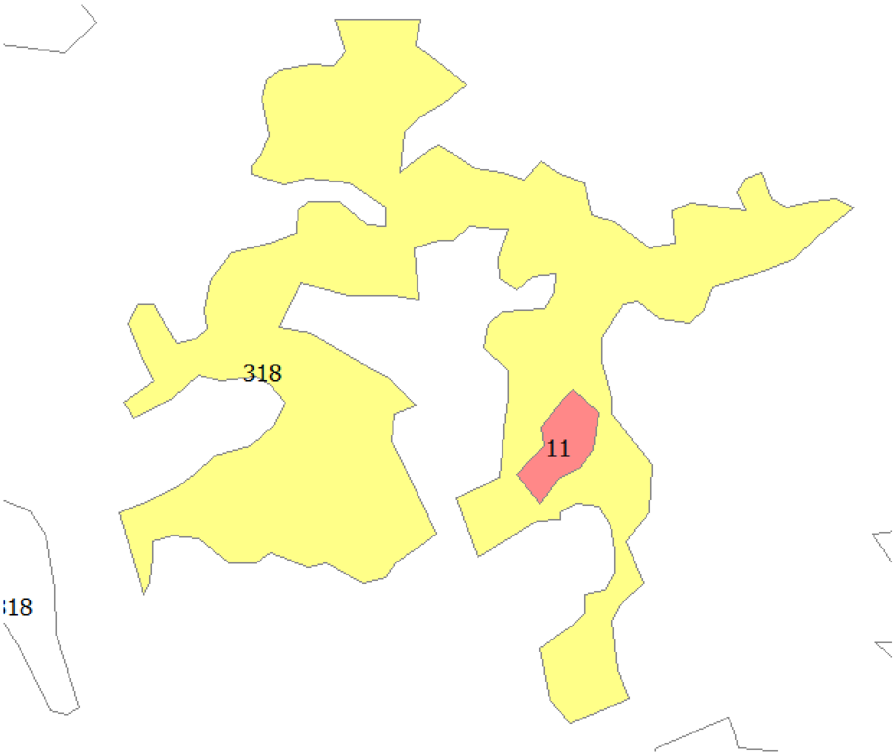
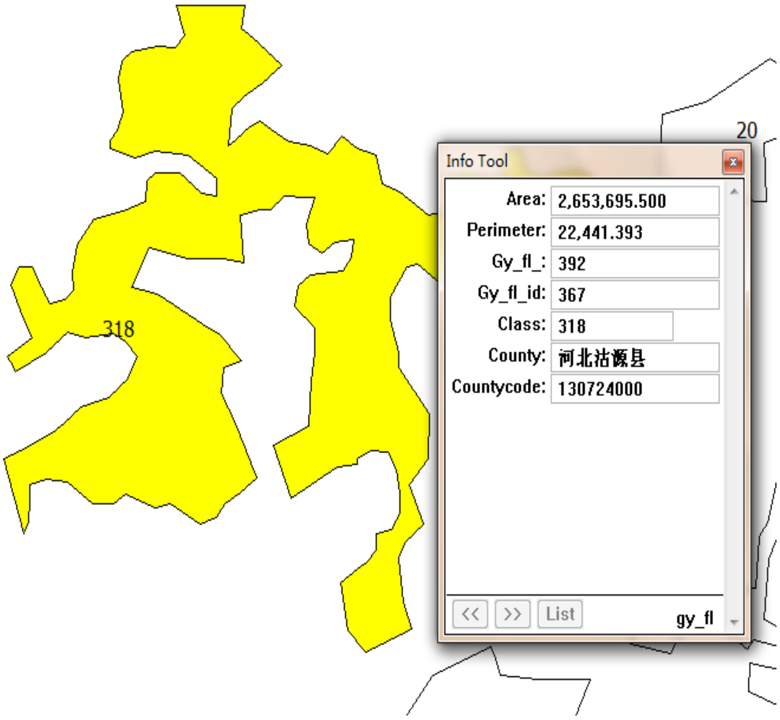


Figure 3.4-1.1 City region contains large area of agricultural land



Area: 2,653,695.500
Perimeter: 22,441.393
Gy_fl_: 392
Gy_fl_id: 367
Class: 318
County: 河北沽源县
Countycode: 130724000

Figure 3.4-1.2 City region contains large area of agricultural land

- Desert contains large area of agricultural land or vegetation: Similar with the above rule, this is also a typical logic related inconsistency detection rule. In land use classification data, if the topological relationship of a pair of polygons is “within”, but the land type of container polygon is desert or saline/alkaline land, while the inside polygon’s land cover type is agricultural land or vegetation such as forest or garden, then this rule will consider this pair of data contain logical inconsistency which breaks common sense. In this rule, only when the area of the inside polygon is larger than a threshold, the pair of polygons could be treated as being consistent.

Land use surveying often focuses on a single land but ignore the relationship between the target land and its neighborhood. As shown in figure 3.4-2, one land contains 2 island lands. When updating data, logical relationship between container land and island lands is as important as the accurate of one single land. This rule defines an inconsistency when a container land is desert land and island lands become unreasonable land type such as large area of forest, farmland, water system such as lake, etc. Desert land is barren land, logically within desert lands is still desert lands. Therefore, if an island is urban land or farmland, it could be an inconsistency and need careful verification to make sure it is correct. If it is yes, then overlap the water system with land use data to see whether the island land is close to the water system. Otherwise, an inconsistency is detected. The reason is that if an urban land or vegetation land is surrounded by desert land, it must be supported by some water system. The rule can be formalized as follows:

If(land A contains land B is true and land A.Class is desert and land B.Class is farmland or landB.Class is forest land or land B.Class is urban land)

Then an inconsistency is detected

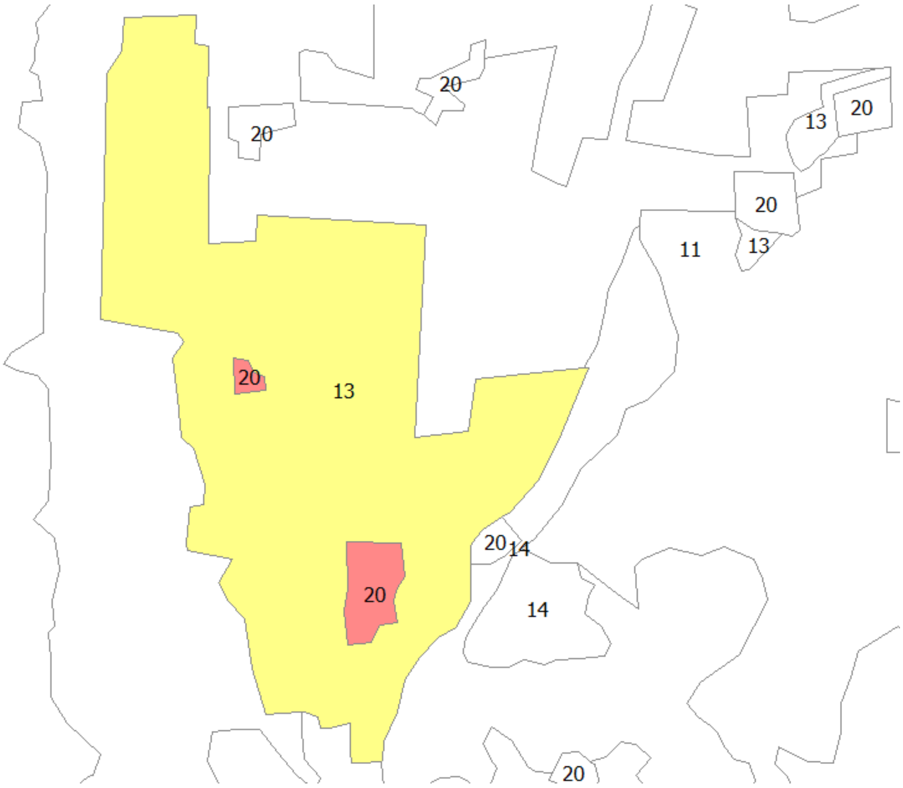


Figure 3.4-2 Desert contains large area of agricultural land or vegetation

- “Relative area of polygons breaks common sense”, as previously mentioned in spatial inconsistency, it is also a logical inconsistency. It uses the geometric information-- “area” to identify whether the logical meaning of land data breaks common sense.

- “Sequence detection of vegetation evolution growing stage” is also mentioned in attribute inconsistency before, which is logic and attribute related consistency detection rule.

- “Irreversible land use type changing”, the same as above, which is logic and attribute related consistency detection rule.

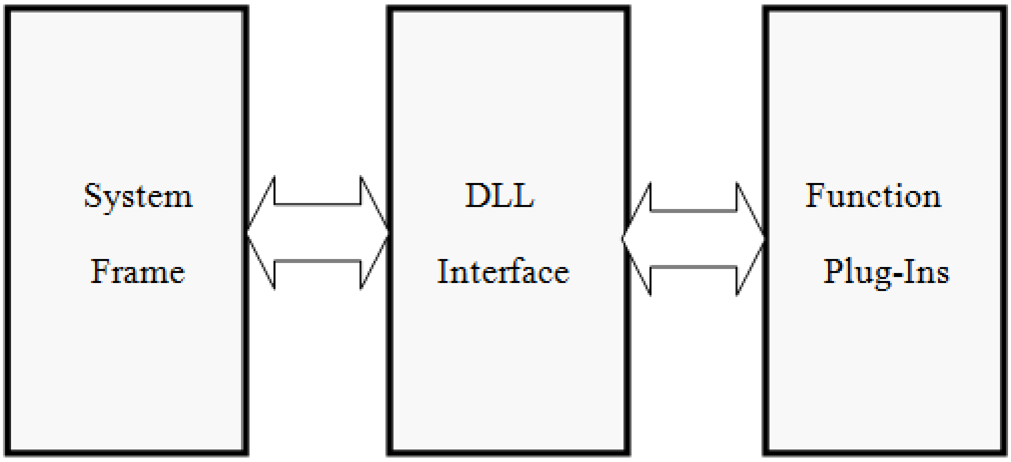
4. Design and implementation of system

Based on previous work, a land use updating inconsistency detection system is designed, which focuses on final land use data after updating. According to a rule-based framework, detection functions are implemented via rules and all rules are developed under a uniform interface to leave space for adding new rules. As combination method is the most widely used one in land use updating, and updating is not only updating the land use data but also changing them, this study will mainly focus on inconsistency detection via analyzing both of these two aspects. The detection functions include the main aspects of general inconsistency checking and developed some new logical inconsistency detection solutions. Developing of general checking functions consulted some current solutions such as TD-Checker. Because land use updating in China involves a lot of human interpretation, the logical inconsistency detection functions also need to be developed for the real situation. Moreover, the proposed detection functions for land use data can be used for data updating by both site surveying and change detection method. For both site surveying and image related updating, proposed logical checking rules are useful to detect errors caused by surveying crew.

4.1 Design of experiment system

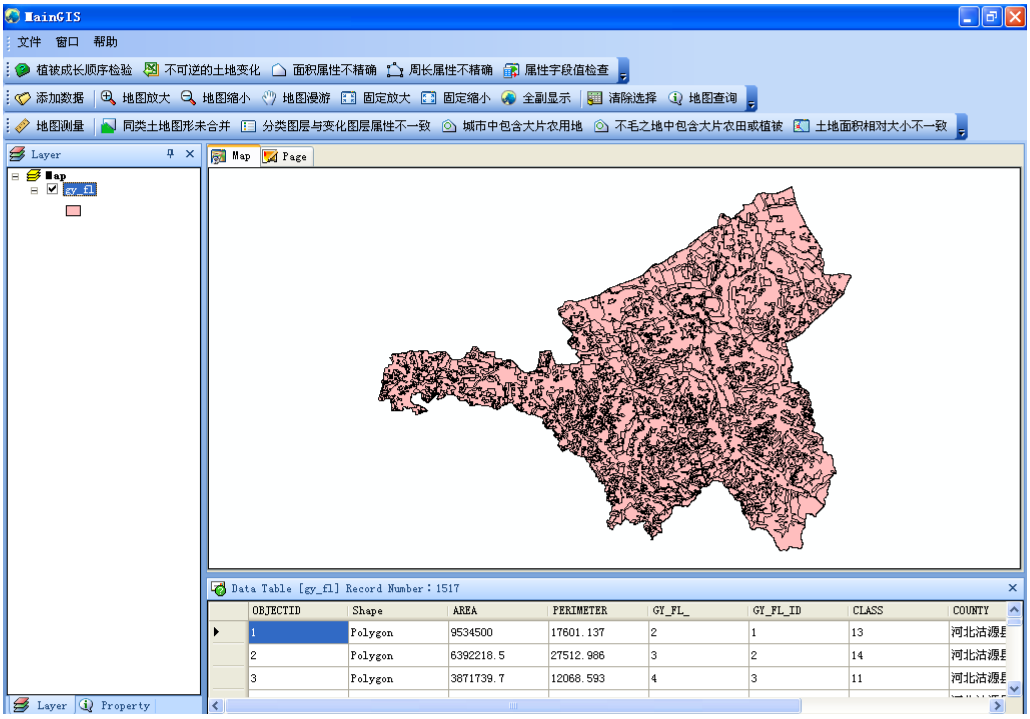
As indicated before, this study not only designed detection rules, but also established a prototype system as an environment to implement and test the detection algorisms. The system focuses on inconsistency detection. For inconsistency correction, the system has a series of interfaces for programmer to expand the functions. It is convenient to expand via interface because programming DLL Plug-In is much easier than coding a whole system.

The Prototype system was established via .NET Frame Work 2005 and C# language. It uses Arc Engine (AE9.2), which is a typical module library for GIS development to build basic user interface, display windows and advance inconsistency detection functions. An overview of the structure of the system is depicted as follows:



Functions that the System Frame offered include: Adding GIS data, Zooming in or out of maps, Panning maps, Zooming as scales, Displaying whole map space, Map layers control tools, Map symbol edit tools, Cleaning selection tools. Plug-Ins offers two sets of inconsistency detection functions as mentioned before. DLL interface links frame and plug-ins as a whole completed system.

Whether the inconsistency detection functions can detect inconsistency of data is the key point of this study. Therefore, a set of land use data of KuYuan county of Heibei province, offered by Beijing Land and Resources Bureau, was used to test the prototype system. As shown in figure 4.1-1, 1517 records of land use data were applied by all 10 available inconsistency detection rules.



Data Table [gy_fl] Record Number：1517

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS	COUNTY
1	Polygon	9534500	17601.137	2	1	13	河北沽源
2	Polygon	6392218.5	27512.986	3	2	14	河北沽源
3	Polygon	3871739.7	12068.593	4	3	11	河北沽源

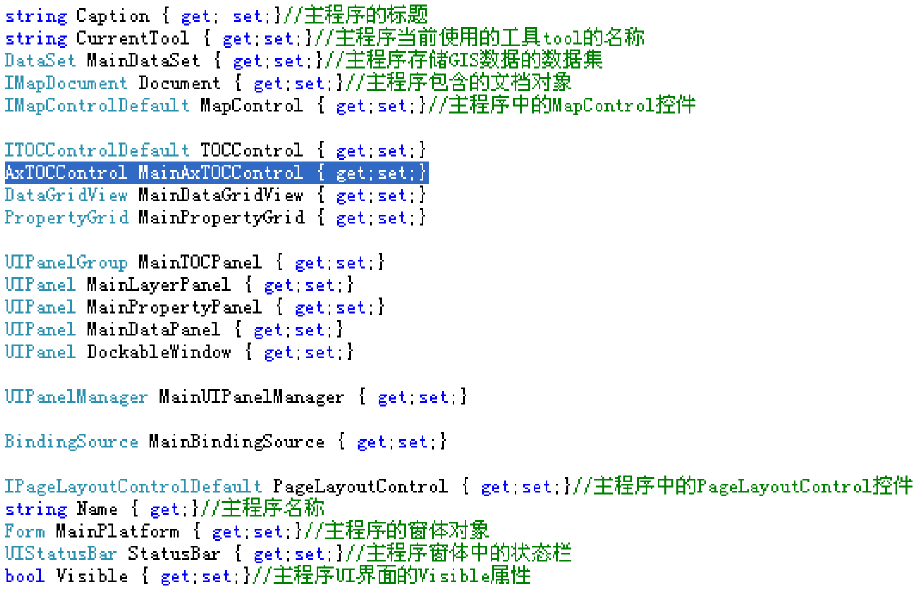
Figure 4.1-1 Test land use map

4.2 Implementation of experiment system

4.2.1 Implementation of system interface

Interface is a special class in .Net programming. The main use of interface class is to establish a standard access channel to different systems and functions. Therefore, a frame system is build to use interface class to expand functions of the whole system. The advantage is that it is not necessary to build a new system. Instead, functions can be added or deleted. By easily programming a DLL which contains the function needs, this DLL can be added to the frame system via interface. Vice versa, if a function is to be deleted from the frame system, it just needs to be deleted in the file system of Windows. This convenience allows us to expand new inconsistency detection rules easily to the experiment system.

Here “IApplication” is used to create the interface class, which contains many interfaces as the channels between DLL and frame system. For instance, the interface “AxTOCControl” shown in figure 4.2.1-1 is the interface used to demonstrate how to link controls on frame system.



```
string Caption { get; set;}//主程序的标题
string CurrentTool { get;set;}//主程序当前使用的工具tool的名称
DataSet MainDataSet { get;set;}//主程序存储GIS数据的数据集
IMapDocument Document { get;set;}//主程序包含的文档对象
IMapControlDefault MapControl { get;set;}//主程序中的MapControl控件

ITOCControlDefault TOCControl { get;set;}
AxTOCControl MainAxTOCControl { get;set;}
DataGridView MainDataGridView { get;set;}
PropertyGrid MainPropertyGrid { get;set;}

UIPanelGroup MainTOCPanel { get;set;}
UIPanel MainLayerPanel { get;set;}
UIPanel MainPropertyPanel { get;set;}
UIPanel MainDataPanel { get;set;}
UIPanel DockableWindow { get;set;}

UIPanelManager MainUIPanelManager { get;set;}

BindingSource MainBindingSource { get;set;}

IPageLayoutControlDefault PageLayoutControl { get;set;}//主程序中的PageLayoutControl控件
string Name { get;}//主程序名称
Form MainPlatform { get;set;}//主程序的窗体对象
UIStatusBar StatusBar { get;set;}//主程序窗体中的状态栏
bool Visible { get;set;}//主程序UI界面的Visible属性
```

Figure 4.2.1-1 Interface code instance

The definition of the content of this interface is:

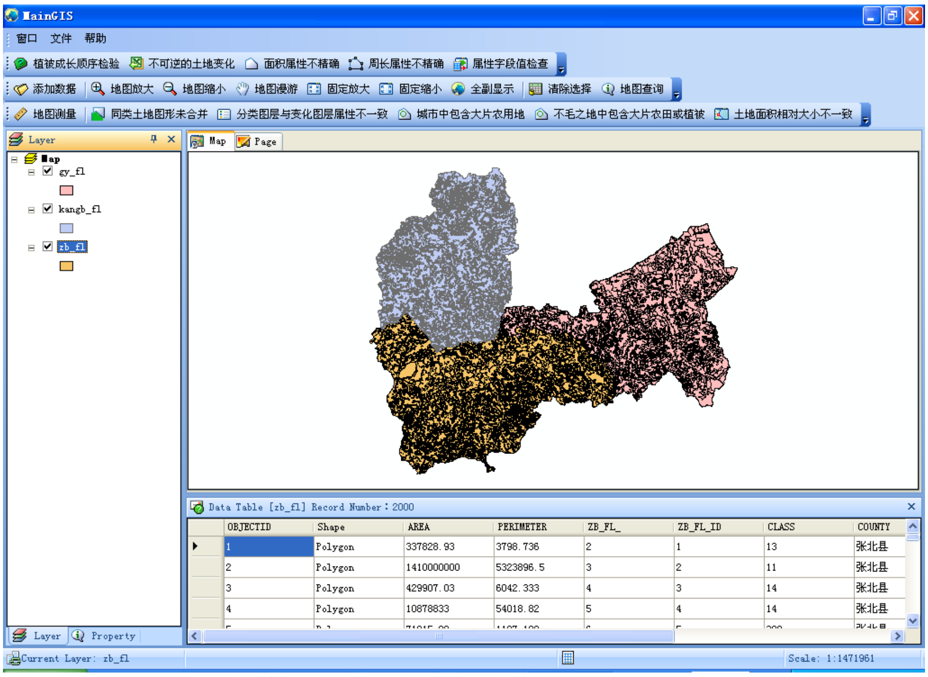
```
public ESRI.ArcGIS.Controls.IMapControlDefault MapControl

{

    get

    {

        return this._MapControl;

    }

    set

    {

        this._MapControl = value;

    }

 }
```

The reason why using interface is that interface build the connection channel between the main programs, i.e. the UI(User Interface) programs and the function programs, i.e. the DLL part. The code above is to use interface as a channel to share the key control “MapControl” between DLL and UI. Therefore both part can use the same one “MapControl”

By using this method, all of the interfaces can be achieved. Then, the interface class “IApplication” itself can be considered as the frame system. This class must be transferred to DLL.

4.2.2 Implementation of frame system

Basically, the frame system is a user interface (UI) of the whole system, which cannot provide any functions. Functions must be added via DLL. After loading function, the interface is like figure 4.2.2-1.



Data Table [zb_fl] Record Number：2000

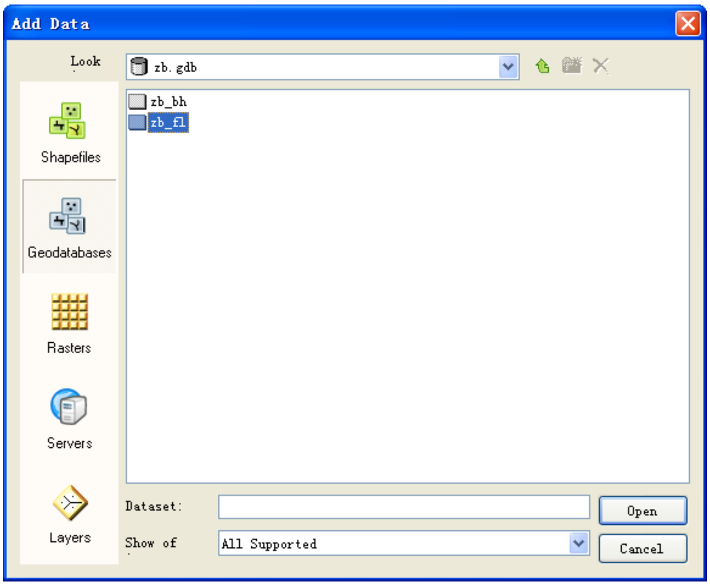
OBJECTID	Shape	AREA	PERIMETER	ZB_FL_	ZB_FL_ID	CLASS	COUNTY
1	Polygon	337828. 93	3798. 736	2	1	13	张北县
2	Polygon	1410000000	5323896. 5	3	2	11	张北县
3	Polygon	429907. 03	6042. 333	4	3	14	张北县
4	Polygon	10878833	54018. 82	5	4	14	张北县

Current Layer: zb_fl

Scale: 1:1471961

Figure 4.2.2-1 Whole interface of the system

When the UI button is pressed, action window will pop out. For instance, adding a map window as it is shown below:



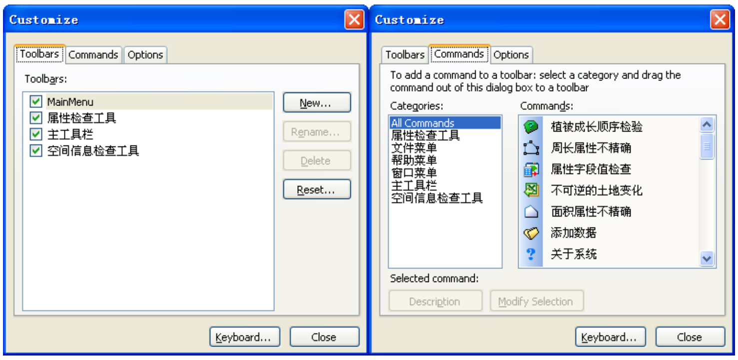
Look: zb. gdb

Dataset:

Show of: All Supported

Figure 4.2.2-2 Add map window

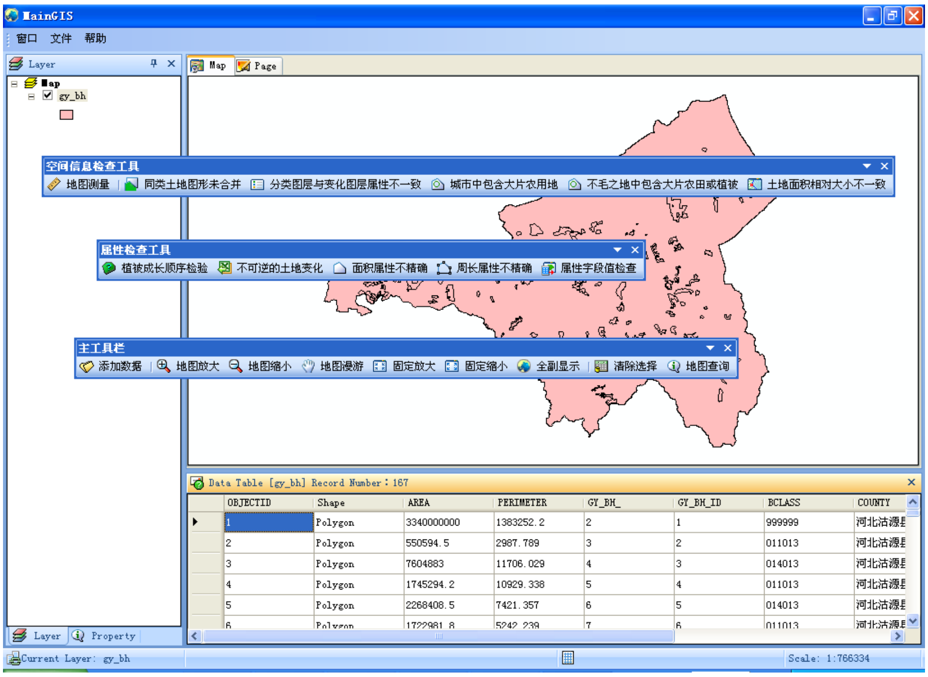
UI can also provide function management tools to manage loaded functions, as shown in figure4.2.2-3:



To add a command to a toolbar: select a category and drag the command out of this dialog box to a toolbar

Figure 4.2.2-3 Tool management window

Menu toolbar, pad-window of UI are both packable, as shown in figure 4.2.2-4 and 4.2.2-5:



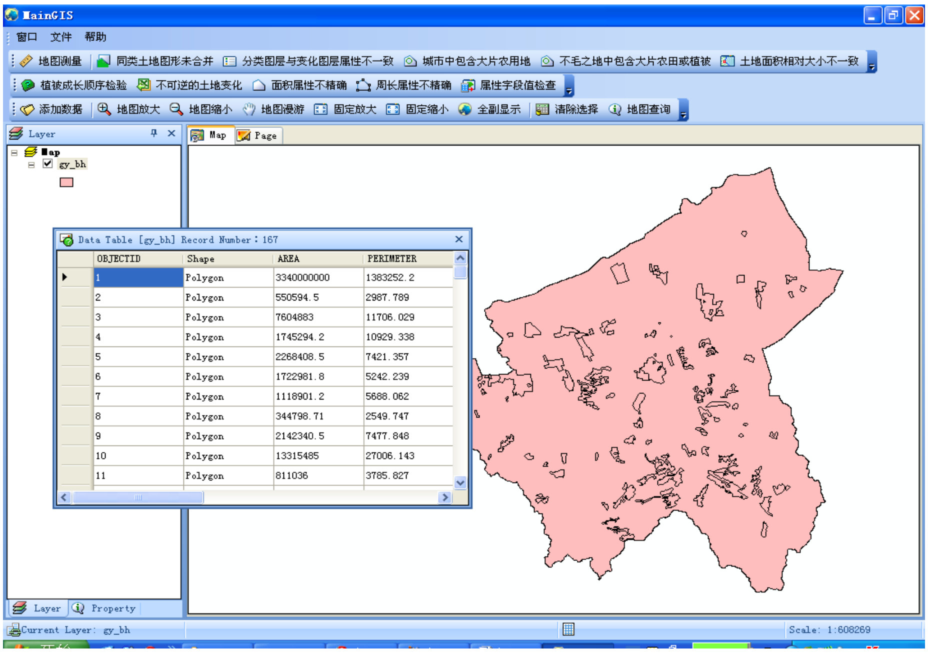
Data Table [gy_bh] Record Number：167

OBJECTID	Shape	AREA	PERIMETER	GY_BH_	GY_BH_ID	BCLASS	COUNTY
1	Polygon	3340000000	1383252.2	2	1	999999	河北沽源
2	Polygon	550594.5	2987.789	3	2	011013	河北沽源
3	Polygon	7604883	11706.029	4	3	014013	河北沽源
4	Polygon	1745294.2	10929.338	5	4	011013	河北沽源
5	Polygon	2268408.5	7421.357	6	5	014013	河北沽源
6	Polygon	1722981.8	5242.239	7	6	011013	河北沽源

Current Layer: gy_bh

Scale: 1:766334

Figure 4.2.2-4 Packable toolbar



Data Table [gy_bh] Record Number：167

OBJECTID	Shape	AREA	PERIMETER
1	Polygon	3340000000	1383252.2
2	Polygon	550594.5	2987.789
3	Polygon	7604883	11706.029
4	Polygon	1745294.2	10929.338
5	Polygon	2268408.5	7421.357
6	Polygon	1722981.8	5242.239
7	Polygon	1118901.2	5688.062
8	Polygon	344798.71	2549.747
9	Polygon	2142340.5	7477.848
10	Polygon	13315485	27006.143
11	Polygon	811036	3785.827

Figure 4.2.2-5 packable sub window

Right-Click menu of UI, as shown in figure 4.2.2-6:

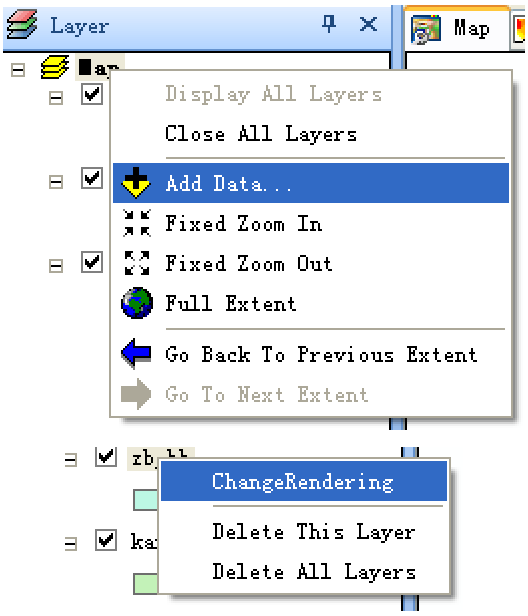


Figure 4.2.2-6 Right-click submenu

Layer color customization window, as shown in figure 4.2.2-7:

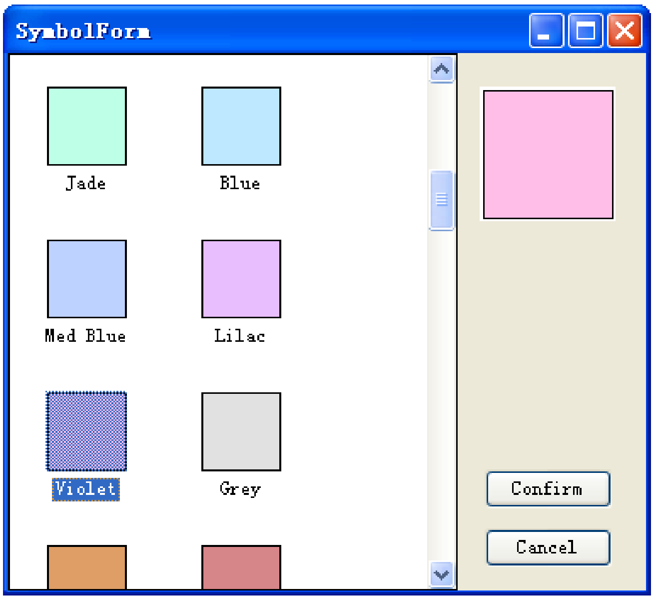
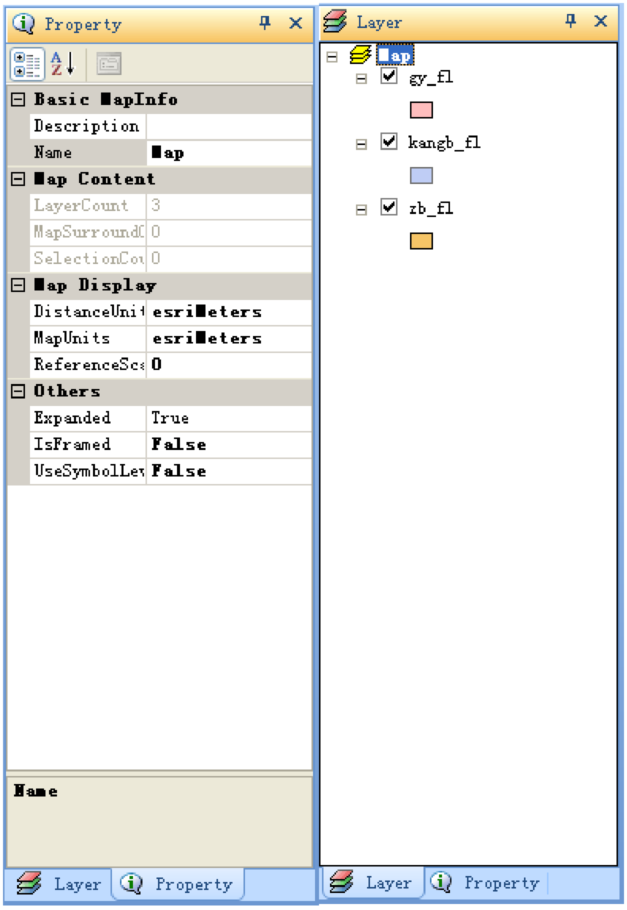


Figure 4.2.2-7 Layer symbol window

Table content / map layer control, as shown in figure 4.2.2-8:



Basic MapInfo	
Description	
Name	**Map**
Map Content	
LayerCount	3
MapSurround(	0
SelectionCou	0
Map Display	
DistanceUnit	**esriMeters**
MapUnits	**esriMeters**
ReferenceSca	**0**
Others	
Expanded	True
IsFramed	**False**
UseSymbolLev	**False**

Name

Figure 4.2.2-8 Table content and layer control window

4.2.3 Implementation of functions

After implementing the interface class of whole system, implementation of function is more like filling a form than programming. For an example, if a function called “ITool” need to be added, the following figure shows all necessary information that needs to be filled in to implement the function.

```
public interface ITool : IPlugin
{
    Bitmap Bitmap { get;}//工具按钮图标
    string Caption { get;}//工具按钮的文字
    string Category { get;}//工具按钮所属的类别
    bool Checked { get;}//工具按钮是否被选择
    bool Enabled { get;}//工具按钮是否可用
    int HelpContextID { get;}//快捷帮助ID
    string HelpFile { get;}//快捷帮助路径
    string Message { get;}//鼠标移动到按钮上时状态栏出现的文字
    string Name { get;}//按钮的名称
    void OnClick();//按钮点击时触发的方法
    void OnCreate(NBGIS.PluginEngine.IApplication hook);//按钮产生时触发,把宿主程序传给插件
    string Tooltip { get;}//按钮的提示信息
    int Cursor { get;}//鼠标在地图上的样式
    bool Deactivate();//tool的激活状态设置
    void OnDblClick();//双击地图时触发的事件
    bool OnContextMenu(int x, int y);//鼠标点击右键弹出快捷菜单时的事件
    void OnMouseMove(int button, int shift, int x, int y);//鼠标在地图上移动时触发的事件
    void OnMouseDown(int button, int shift, int x, int y);//鼠标点击地图时触发的事件
    void OnMouseUp(int button, int shift, int x, int y);//鼠标在地图上弹起时触发的事件
    void Refresh(int hDC);//地图刷新时触发的事件
    void OnKeyDown(int keyCode, int shift);//键盘上某键点击时的事件
    void OnKeyUp(int keyCode, int shift);//键盘上某键弹起时的事件
}
```

Figure 4.2.3-1 Adding function interface

The above code give a set of “attributes” a DLL function should have. When programming DLL function, only need to follow these “attributes”, add correct material. Such as attribute “Refresh”, it means the DLL should have a paragraph of code to response the acting of map refresh.

Two key points need to be concerned:

① When creating the interface class function such as “IApplication”, it must be transferred into current object. Then the function system can use the resource such as UI, of frame system.

② The content under “OnCLick()”is the detail task of the function.

Full version of this implementation is shown as follows:

```
class cZoomIn:NBGIS.PluginEngine.ITool

{

    private NBGIS.PluginEngine.IApplication hk;

    private System.Drawing.Bitmap m_hBitmap;
```

```
        private ESRI.ArcGIS.SystemUI.ITool tool = null;

        private ESRI.ArcGIS.SystemUI.ICommand cmd = null;


        public cZoomIn()

        {

            m_hBitmap = new

Bitmap(this.GetType().Assembly.GetManifestResourceStream("NBGIS.MainTools.zoomIn.ico"));

        }


        public Bitmap Bitmap

        {

            get { return m_hBitmap; }

        }


        public string Caption

        {

            get { return "Zoom In"; }

        }


        public string Category

        {

            get { return "Main Tools"; }

        }


        public bool Checked

        {

            get { return false; }

        }


        public bool Enabled
```

```
{
    get { return true; }
}

public int HelpContextID
{
    get { return 0; }
}

public string HelpFile
{
    get { return ""; }
}

public string Message
{
    get { return "Zoom In"; }
}

public string Name
{
    get { return "Zoom In"; }
}

public void OnClick()
{
    cmd.OnClick();
    this.hk.MapControl.CurrentTool = tool;
}
```

```
public void OnCreate(IApplication hook)

{

    if (hook != null)

    {

        this.hk = hook;

        tool = new ControlsMapZoomInToolClass();

        cmd = tool as ESRI.ArcGIS.SystemUI.ICommand;

        cmd.OnCreate(this.hk.MapControl);

    }

}


public string Tooltip

{

    get { return "Zoom In"; }

}


public int Cursor

{

    get { return (int)ESRI.ArcGIS.Controls.esriControlsMousePointer.esriPointerZoomIn; }

}


public bool Deactivate()

{

    return false;

}


public void OnDblClick()

{ }


public bool OnContextMenu(int x, int y)
```

```
        {
            return false;
        }

        public void OnMouseMove(int button, int shift, int x, int y)
        { }

        public void OnMouseDown(int button, int shift, int x, int y)
        { }

        public void OnMouseUp(int button, int shift, int x, int y)
        { }

        public void Refresh(int hDC)
        { }

        public void OnKeyDown(int keyCode, int shift)
        { }

        public void OnKeyUp(int keyCode, int shift)
        { }
    }
```

This full version of DLL is the code for “ZoomIn” function. Programming DLL here is like to fill a blank form. Every “item” in the “form” is listed. Filling the necessary blank with code can complete the programming of the DLL.

Function system will be created as DLL file and then copied to the path under the frame system to make the whole system work.

5. Experiment Evaluation

5.1 Experiment data

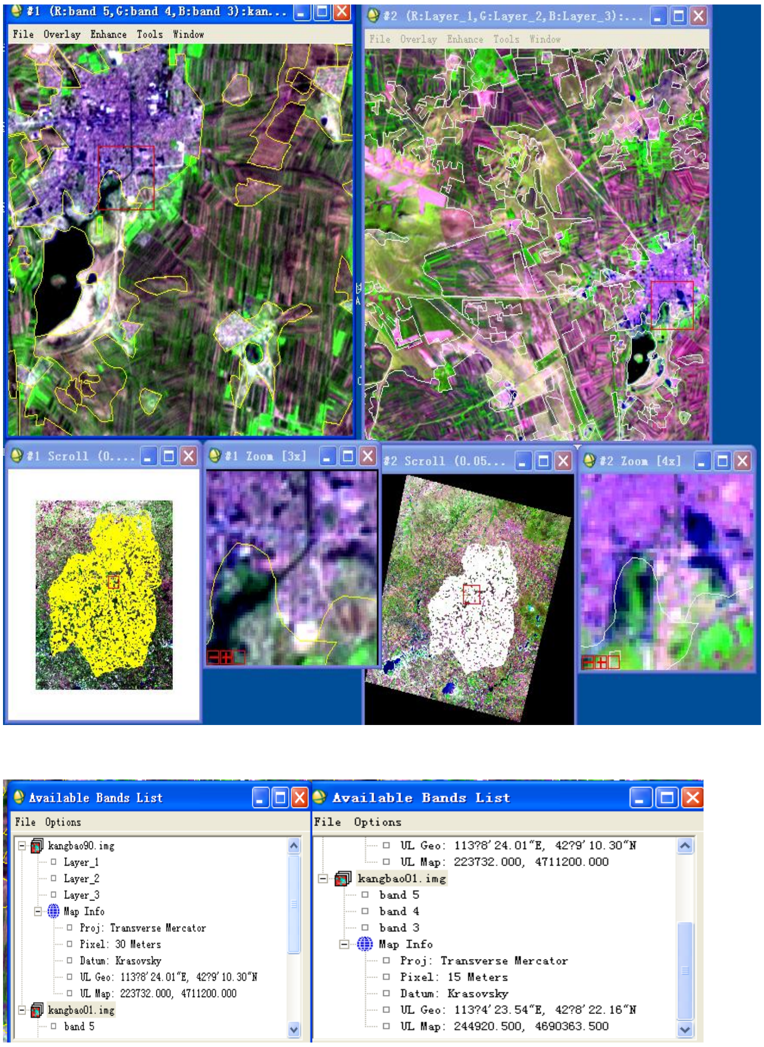
As a modern developing technique, remote sensing shows its advantages on land use investigation. These advantages include larger cover area, inexpensive to purchase and image is easy to be interpreted. Besides, interpretation methods of remote sensing image have been highly developed. Therefore, surveyor and data production department are inclined to interpret and identify the land use change information by using remote sensing image and use auto- classifying algorithm. The land use change information obtained from remote sensing images can be applied to an old land classification data by using overlapping algorithm. Using change data to update land use database is a widely used method in land use data production. This method was also used by Beijing Municipal Land and Resources Bureau to update land use data. That is the reason why the experiment data are divided into two types: land use data and land use change data.

The prototype system is considered as a collection of data processing steps, as data is the raw material for the system to manage. Data processing should be treated as a black box. In this black box, output is various types of results, while data are the unique input. Therefore, the quality, content, structure and form of data are highly related to the prototype system and output products. Content of metadata is the key point which decides whether the prototype system could attain the requirements. Pre-analysis of land use data is an important section for the entire work.

The experiment land use data is offered by Land and Resources Bureau of Beijing. The format of these data is shape file, which is a standard data format issued by ESRI and widely used in GIS research area and ArcGIS software. Metadata of the land use data are shown as follows:

- Feature Type: Polygon and image

- Content: Each file covers a single county or city. Each polygon represents a single zone with a typical land use classification.
- Projection: TRANSVER MERCATOR
- Ellipsoid system: krasovsky, proportional divisor is 1, center meridian is 117 °E
- Time: for land use classification data: 2001; for land use changing data: 1990(old class) and 2001(new class)
- Graphic data:



```
kangbao90.img
  Layer_1
  Layer_2
  Layer_3
  Map Info
    Proj: Transverse Mercator
    Pixel: 30 Meters
    Datum: Krasovsky
    UL Geo: 113?8'24.01"E, 42?9'10.30"N
    UL Map: 223732.000, 4711200.000
kangbao01.img
  band 5
```

```
    UL Geo: 113?8'24.01"E, 42?9'10.30"N
    UL Map: 223732.000, 4711200.000
kangbao01.img
  band 5
  band 4
  band 3
  Map Info
    Proj: Transverse Mercator
    Pixel: 15 Meters
    Datum: Krasovsky
    UL Geo: 113?4'23.54"E, 42?8'22.16"N
    UL Map: 244920.500, 4690363.500
```

Figure 5.1-1 image data in 1990 and 2001

Graphic data contain images of target area of 1990 and 2001. The images were merged by band 5, 4 and 3. Resolution of image in 1990 is 30 meters and that of 2001 is 15 meters. All images are produced by China surveying and mapping academy of sciences.

Graphic data also include two sets of vector maps: land use classification map and land use change map. Land use classification map is the final product of land use data, which gives the boundary and land cover class of every single land in the target area, as shown in figure 5.1-3. Land use change data only describe the land where change happened, as shown in figure 5.1-2. Land use change data are interpreted from images or surveying reports and will be used as the basic data to produce land use classification data.

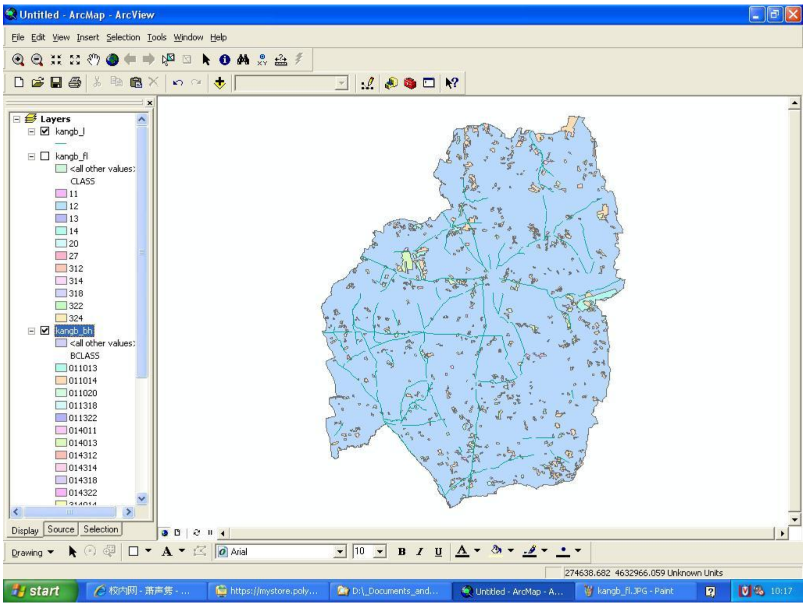


Figure 5.1-2 Vector data: change data of land use map

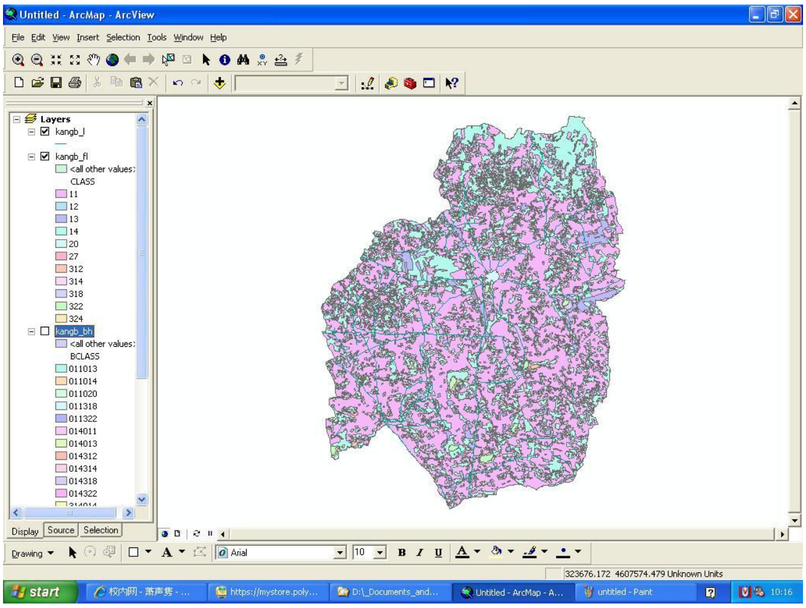
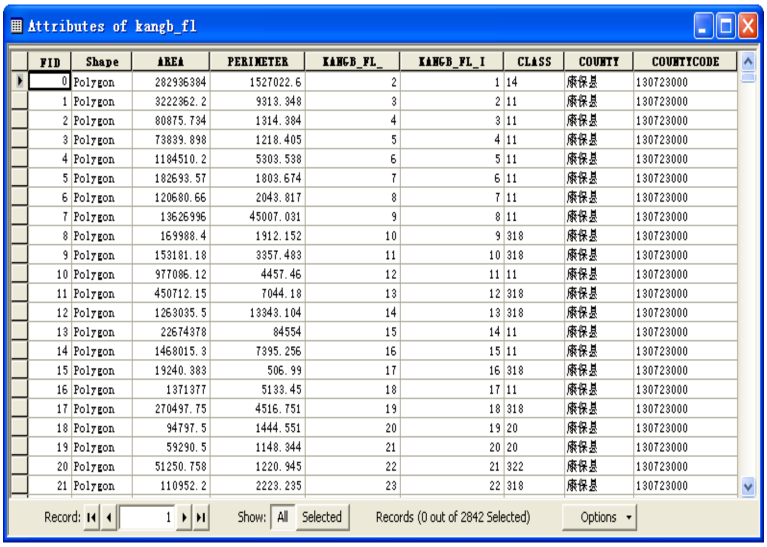


Figure 5.1-3 Vector data: classification data of land use map

- Attribute data:



Attributes of kangb_fl

FID	Shape	AREA	PERIMETER	KANGB_FL_	KANGB_FL_I	CLASS	COUNTY	COUNTYCODE
0	Polygon	282936384	1527022.6	2	1	14	康保县	130723000
1	Polygon	3222362.2	9313.348	3	2	11	康保县	130723000
2	Polygon	80875.734	1314.384	4	3	11	康保县	130723000
3	Polygon	73839.898	1218.405	5	4	11	康保县	130723000
4	Polygon	1184510.2	5303.538	6	5	11	康保县	130723000
5	Polygon	182693.57	1803.674	7	6	11	康保县	130723000
6	Polygon	120680.66	2043.817	8	7	11	康保县	130723000
7	Polygon	13626996	45007.031	9	8	11	康保县	130723000
8	Polygon	169988.4	1912.152	10	9	318	康保县	130723000
9	Polygon	153181.18	3357.483	11	10	318	康保县	130723000
10	Polygon	977086.12	4457.46	12	11	11	康保县	130723000
11	Polygon	450712.15	7044.18	13	12	318	康保县	130723000
12	Polygon	1263035.5	13343.104	14	13	318	康保县	130723000
13	Polygon	22674378	84554	15	14	11	康保县	130723000
14	Polygon	1468015.3	7395.256	16	15	11	康保县	130723000
15	Polygon	19240.383	506.99	17	16	318	康保县	130723000
16	Polygon	1371377	5133.45	18	17	11	康保县	130723000
17	Polygon	270497.75	4516.751	19	18	318	康保县	130723000
18	Polygon	94797.5	1444.551	20	19	20	康保县	130723000
19	Polygon	59290.5	1148.344	21	20	20	康保县	130723000
20	Polygon	51250.758	1220.945	22	21	322	康保县	130723000
21	Polygon	110952.2	2223.235	23	22	318	康保县	130723000

Record: 1 Show: All Selected Records (0 out of 2842 Selected) Options

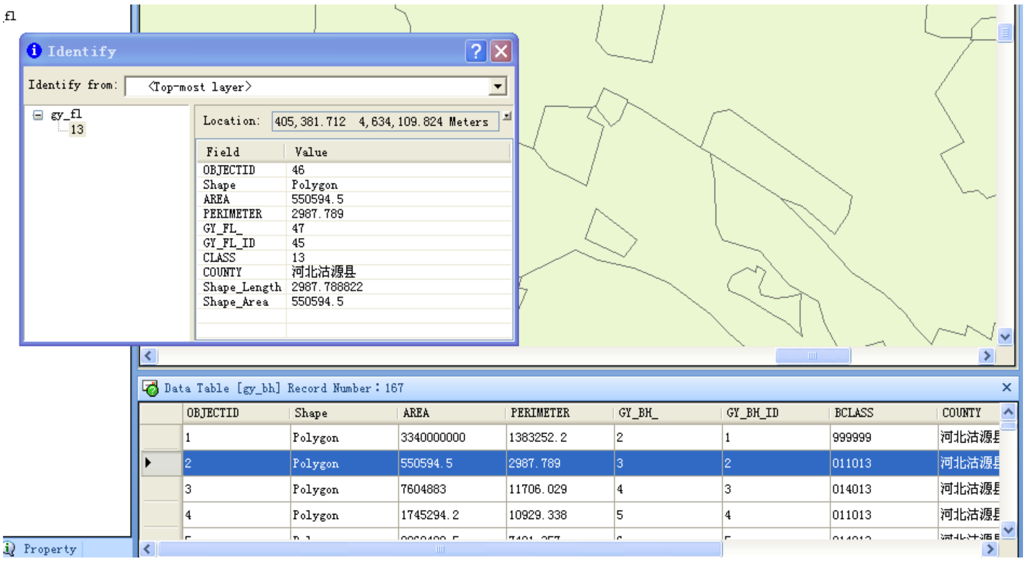
Figure 5.1-4 Attribute table of land use map

Attribute data include information such as ID, area, perimeter, county code and the most important land use class code.

The content of experiment data is one of analysis objects of this study. Therefore, pre-analysis over land use data content is very important. The content of data includes two types: land use classification data and land use change data. Classification data can directly present information of the distribution and type of land resource, while land use change data present the development trend of land resource between different time points. These data are convenient for land use data updating. Collecting land use change data and using it to update fresh land use classification data is a basic, efficient and widely used method for land use data updating.

For land classification data, each polygon has a code recorded in corresponding column of the attribute table to represents land type, such as building land, land for hydraulic structure, farm land, based on the "national land use classification coding regulation". Land classification data are normally produced by three main steps: remote sensing image interpretation, classification and vector data digitization.

For land use change data, each polygon represents the land type changing information between two periods. For instance, if surveyor found that a single land area had been changed from farm land to building land during land investigation or land monitoring, the corresponding column in the attribute table should record the information that indicated how this land has been changed. For an example, a land changed from farm land to building land can be recorded by a code link, such as 136203 which means nursery land has been changed to village settlement land. As depicted in figure 5.1-5, column "Class" recorded the land type code "13", means this land is woodland according to the "national land use classification coding regulation".



Identify from: <Top-most layer>

gy_fl
13

Location: 405,381.712 4,634,109.824 Meters

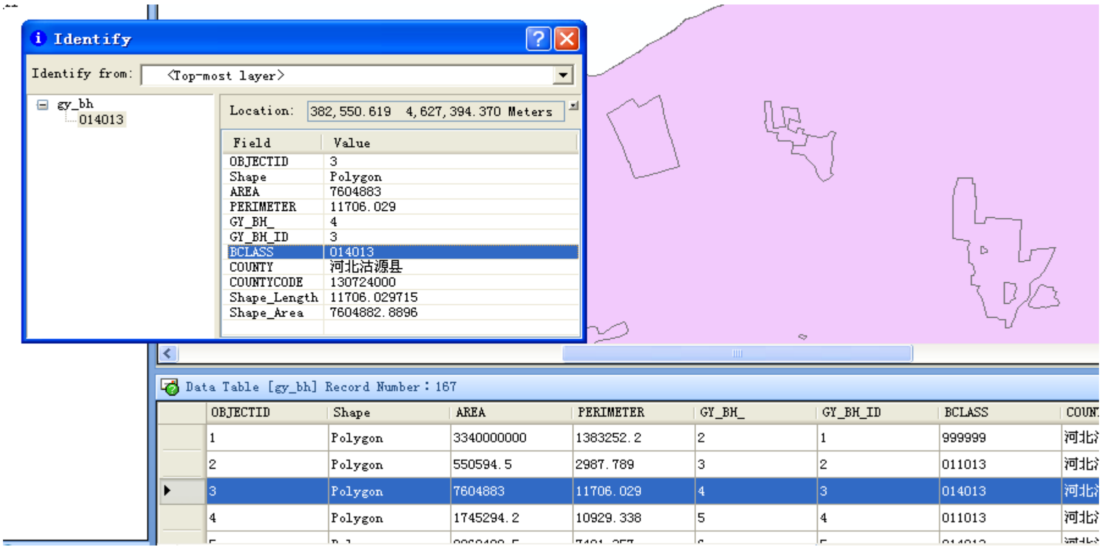
Field	Value
OBJECTID	46
Shape	Polygon
AREA	550594.5
PERIMETER	2987.789
GY_FL_	47
GY_FL_ID	45
CLASS	13
COUNTY	河北沽源县
Shape_Length	2987.788822
Shape_Area	550594.5

Data Table [gy_bh] Record Number：167

OBJECTID	Shape	AREA	PERIMETER	GY_BH_	GY_BH_ID	BCLASS	COUNTY
1	Polygon	3340000000	1383252.2	2	1	999999	河北沽源
2	Polygon	550594.5	2987.789	3	2	011013	河北沽源
3	Polygon	7604883	11706.029	4	3	014013	河北沽源
4	Polygon	1745294.2	10929.338	5	4	011013	河北沽源

Figure 5.1-5 Instance of land type code

As depicted in figure 5.1-6, column "Bclass" recorded the land use changing information code “014013”, means this land changed from meadowland to woodland.



Location: 382, 550. 619   4, 627, 394. 370 Meters

Field	Value
OBJECTID	3
Shape	Polygon
AREA	7604883
PERIMETER	11706. 029
GY_BH_	4
GY_BH_ID	3
BCLASS	014013
COUNTY	河北沽源县
COUNTYCODE	130724000
Shape_Length	11706. 029715
Shape_Area	7604882. 8896

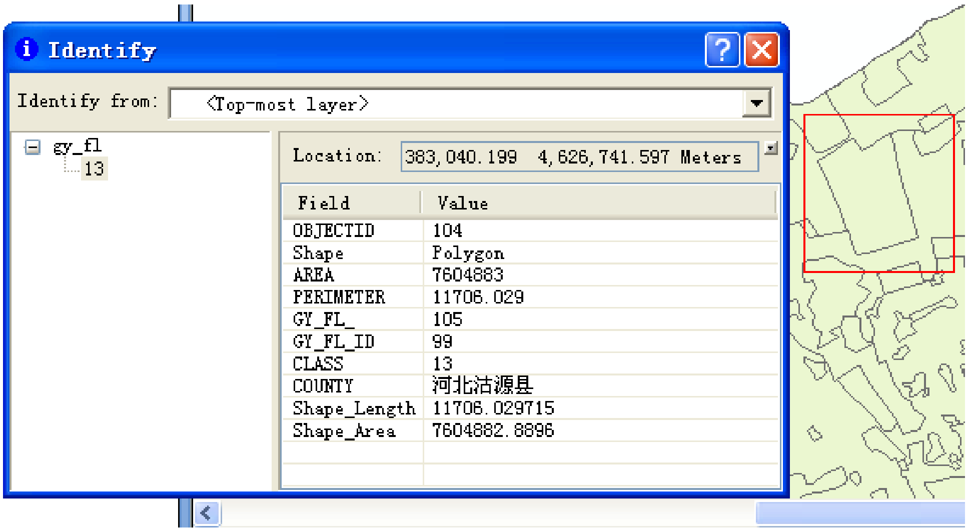
Data Table [gy_bh] Record Number：167

OBJECTID	Shape	AREA	PERIMETER	GY_BH_	GY_BH_ID	BCLASS	COUN
1	Polygon	3340000000	1383252. 2	2	1	999999	河北
2	Polygon	550594. 5	2987. 789	3	2	011013	河北
3	Polygon	7604883	11706. 029	4	3	014013	河北
4	Polygon	1745294. 2	10929. 338	5	4	011013	河北

Figure 5.1-6 Change Class of land use type

As shown in figure 5.1-5 and 5.1-6, the attribute table can not only record the information of land use class and land use change code, but also can record various auxiliary information, such as ID, feature type, area number, perimeter number and administrative division ID number. The auxiliary information is useful for the land use data user, as well as provides important information for designing data inconsistency detection algorithm in this study.

Figure 5.1-7 and figure 5.1-8 indicate that the relationship between land classification data and land use change data when they represent the same area. The type code of the target land in land classification data is “13”, while the change code in land use change data is “014013”, which means the target land changed from meadowland to woodland.



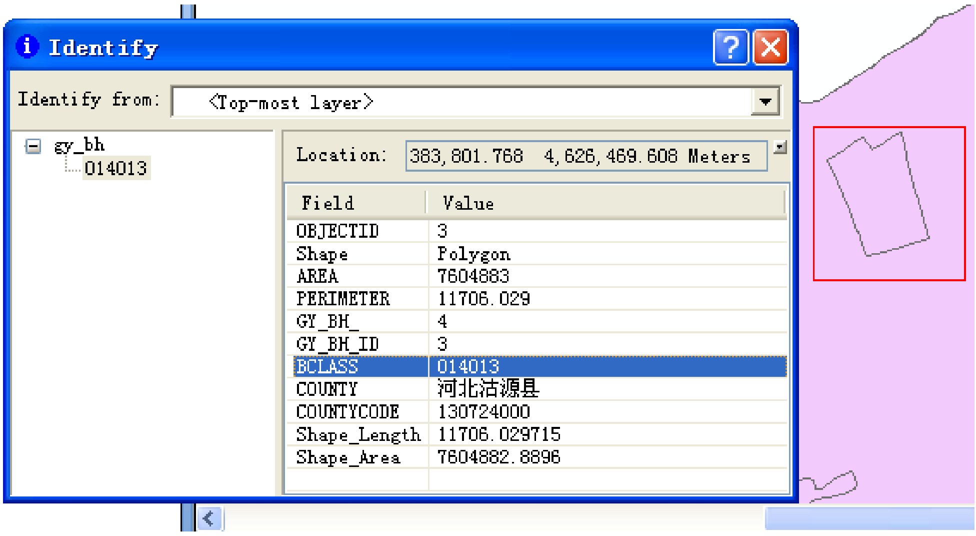
Identify from: <Top-most layer>

gy_fl
13

Location: 383, 040. 199 4, 626, 741. 597 Meters

Field	Value
OBJECTID	104
Shape	Polygon
AREA	7604883
PERIMETER	11706. 029
GY_FL_	105
GY_FL_ID	99
CLASS	13
COUNTY	河北沽源县
Shape_Length	11706. 029715
Shape_Area	7604882. 8896

Figure 5.1-7 Relationship: Classification data



Identify from: <Top-most layer>

gy_bh
014013

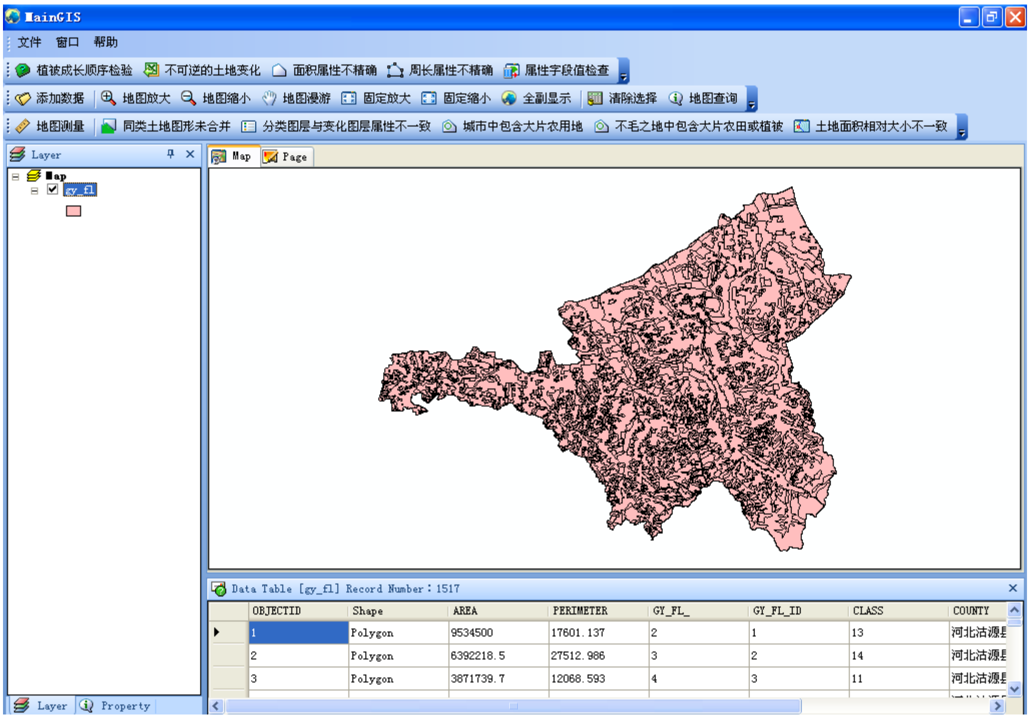
Location: 383, 801. 768 4, 626, 469. 608 Meters

Field	Value
OBJECTID	3
Shape	Polygon
AREA	7604883
PERIMETER	11706. 029
GY_BH_	4
GY_BH_ID	3
BCLASS	014013
COUNTY	河北沽源县
COUNTYCODE	130724000
Shape_Length	11706. 029715
Shape_Area	7604882. 8896

Figure 5.1-8 Relationship: Change data

5.2 Results of Inconsistency Detection

The data used for the experiment are shown as bellow:



Data Table [gy_fl] Record Number：1517

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS	COUNTY
1	Polygon	9534500	17601.137	2	1	13	河北沽源
2	Polygon	6392218.5	27512.986	3	2	14	河北沽源
3	Polygon	3871739.7	12068.593	4	3	11	河北沽源

Figure 5.2-1 Experiment land use map

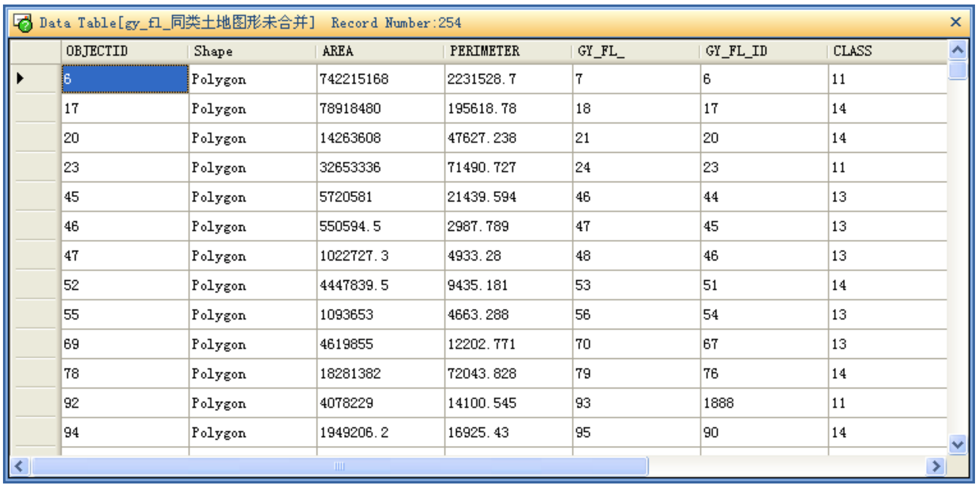
The results of the test are as follows:

- Same type lands without merging: 254 records
- Matching problems between land use change and classification data: 33 records
- City region contains large area of agricultural land: 0 record
- Desert contains large area of agricultural land or vegetation: 7 records
- Relative area number of polygons breaks common senses: 0 record
- Vegetation growing stage sequence detection: 0 record
- Irreversible land type changing: 1 record
- Inaccurate area key value: 1 record
- Inaccurate perimeter key value: 0 record

- Attribute field inspection: 0 record

Since data records containing inconsistency are concerned, results are saved into tables. User can check and query them by click them in table panel. The map will pan to the target polygon. Using data query tool can check the information of polygon.

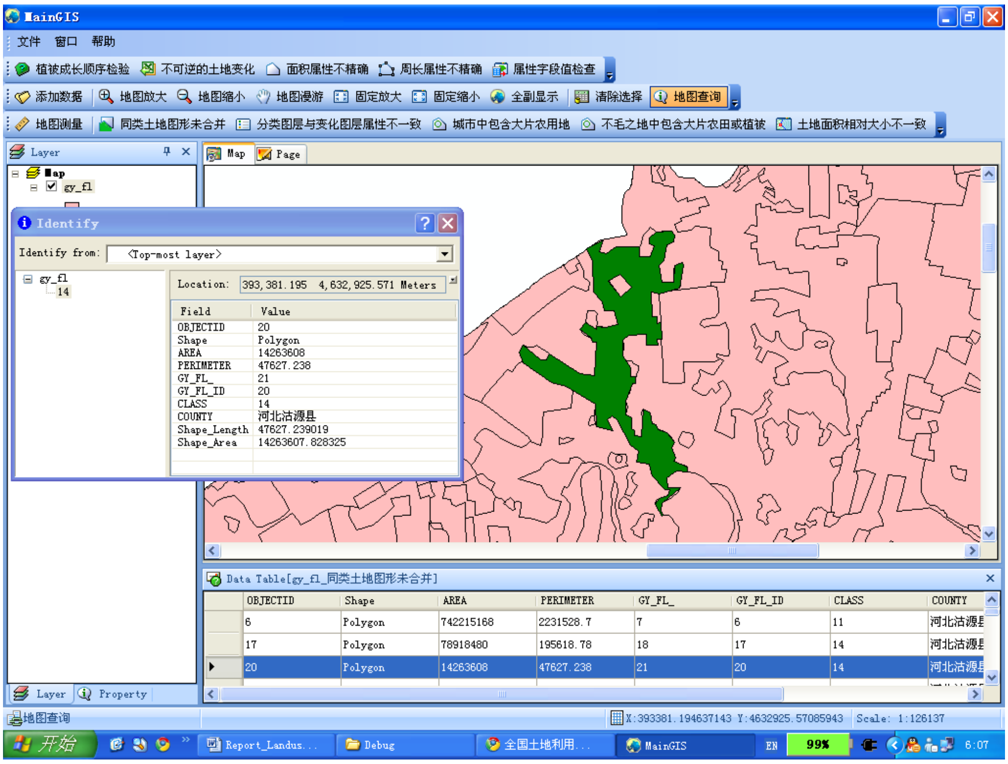
Figure 5.2-2 describes the 254 records which represent "Same type lands without merging". Two records in the result table were randomly chosen, shown in figure 5.2-3 and 5.2-4. In figure 5.2-3, the target polygon has a classification code 14, which is grassland. While in figure 5.2-4, the adjacent land is also grassland, with a same classification code "14".



Data Table[gy_fl_同类土地图形未合并] Record Number:254

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS
6	Polygon	742215168	2231528.7	7	6	11
17	Polygon	78918480	195618.78	18	17	14
20	Polygon	14263608	47627.238	21	20	14
23	Polygon	32653336	71490.727	24	23	11
45	Polygon	5720581	21439.594	46	44	13
46	Polygon	550594.5	2987.789	47	45	13
47	Polygon	1022727.3	4933.28	48	46	13
52	Polygon	4447839.5	9435.181	53	51	14
55	Polygon	1093653	4663.288	56	54	13
69	Polygon	4619855	12202.771	70	67	13
78	Polygon	18281382	72043.828	79	76	14
92	Polygon	4078229	14100.545	93	1888	11
94	Polygon	1949206.2	16925.43	95	90	14

Figure 5.2-2 Result table



Identify

Identify from: <Top-most layer>

gy_fl
14

Location: 393,381.195 4,632,925.571 Meters

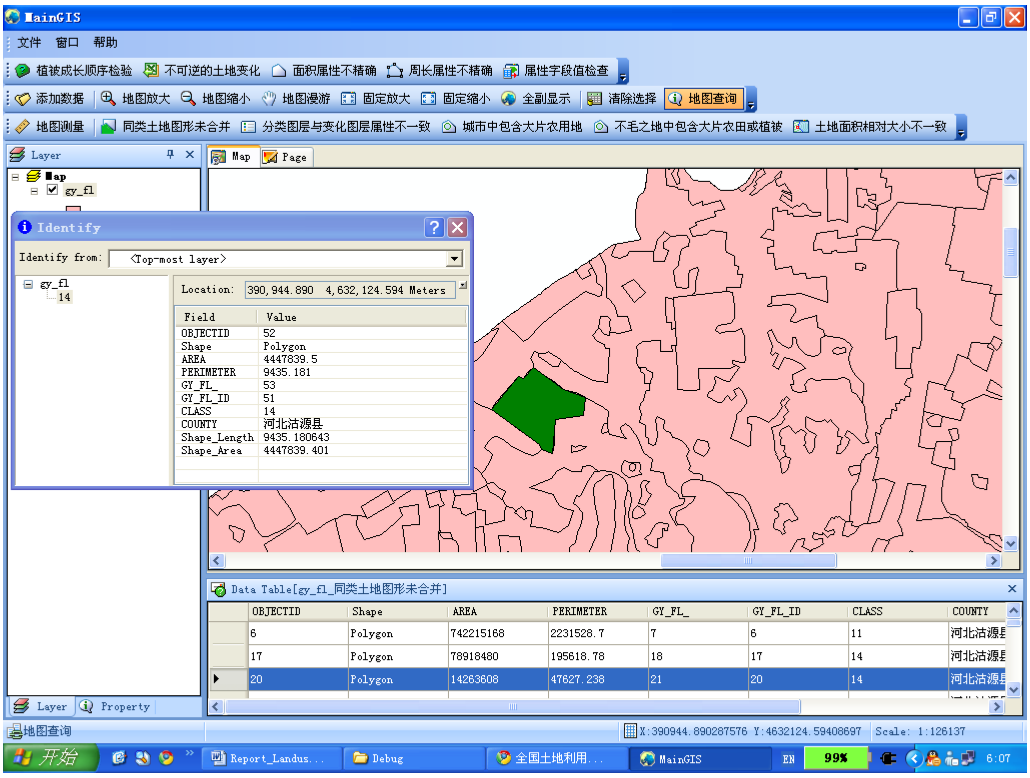
Field	Value
OBJECTID	20
Shape	Polygon
AREA	14263608
PERIMETER	47627.238
GY_FL_	21
GY_FL_ID	20
CLASS	14
COUNTY	河北沽源县
Shape_Length	47627.239019
Shape_Area	14263607.828325

Data Table[gy_fl_同类土地图形未合并]

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS	COUNTY
6	Polygon	742215168	2231528.7	7	6	11	河北沽源县
17	Polygon	78918480	195618.78	18	17	14	河北沽源县
20	Polygon	14263608	47627.238	21	20	14	河北沽源县

X:393381.194637143 Y:4632925.57085943 Scale: 1:126137

Figure 5.2-3 Target polygon



Identify

Identify from: <Top-most layer>

gy_fl
14

Location: 390,944.890 4,632,124.594 Meters

Field	Value
OBJECTID	52
Shape	Polygon
AREA	4447839.5
PERIMETER	9435.181
GY_FL_	53
GY_FL_ID	51
CLASS	14
COUNTY	河北沽源县
Shape_Length	9435.180643
Shape_Area	4447839.401

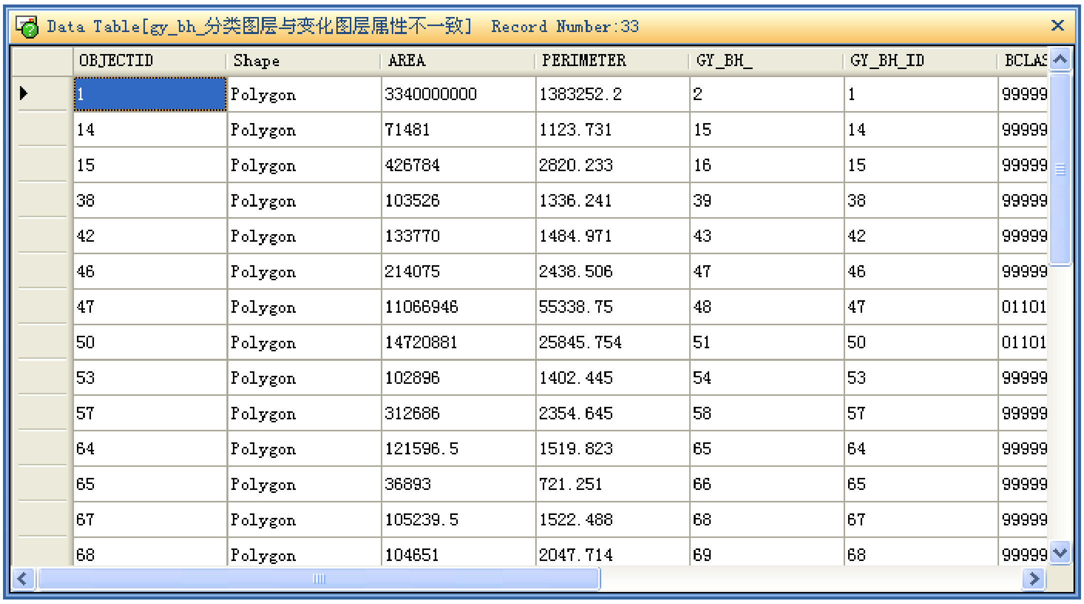
Data Table[gy_fl_同类土地图形未合并]

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS	COUNTY
6	Polygon	742215168	2231528.7	7	6	11	河北沽源县
17	Polygon	78918480	195618.78	18	17	14	河北沽源县
20	Polygon	14263608	47627.238	21	20	14	河北沽源县

X:390944.890287576 Y:4632124.59408697 Scale: 1:126137

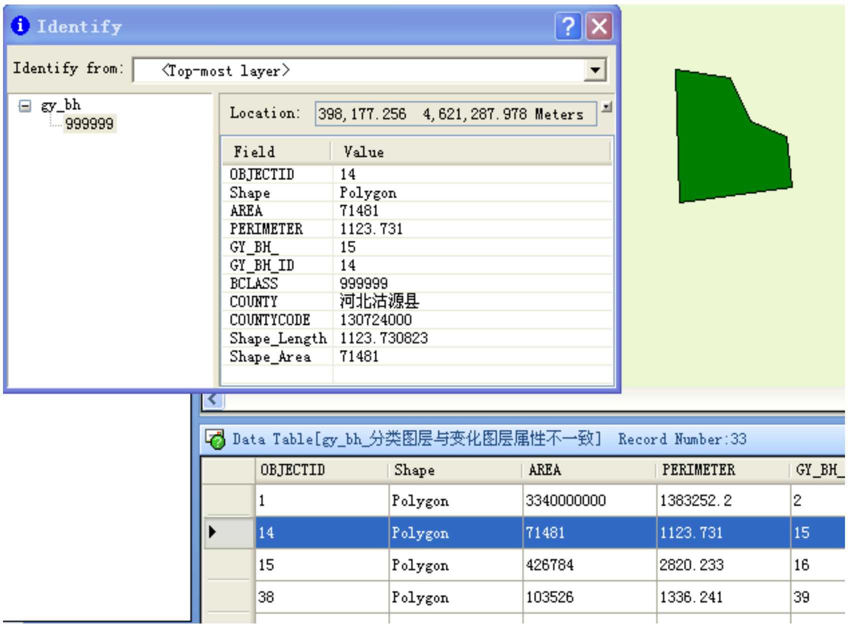
Figure 5.2-4 Adjacent polygons of the target one

Figure 5.2-5 describes the 33 records which represent “Matching problems between land use change and classification data”. Several records were also randomly selected in the result table: One has the land type as “999999” in classification data, while it is “20” in land use change data. “999999” means it comes from the base map or hasn’t been updated. If it has not been updated yet, the corresponding record should be deleted from the data table. Otherwise, if the code comes from the base map, which means the target polygon may miss its land cover information.


Data Table[gy_bh_分类图层与变化图层属性不一致] Record Number:33

OBJECTID	Shape	AREA	PERIMETER	GY_BH_	GY_BH_ID	BCLAS
1	Polygon	3340000000	1383252.2	2	1	99999
14	Polygon	71481	1123.731	15	14	99999
15	Polygon	426784	2820.233	16	15	99999
38	Polygon	103526	1336.241	39	38	99999
42	Polygon	133770	1484.971	43	42	99999
46	Polygon	214075	2438.506	47	46	99999
47	Polygon	11066946	55338.75	48	47	01101
50	Polygon	14720881	25845.754	51	50	01101
53	Polygon	102896	1402.445	54	53	99999
57	Polygon	312686	2354.645	58	57	99999
64	Polygon	121596.5	1519.823	65	64	99999
65	Polygon	36893	721.251	66	65	99999
67	Polygon	105239.5	1522.488	68	67	99999
68	Polygon	104651	2047.714	69	68	99999

Figure 5.2-5 Result table



Identify from: <Top-most layer>

gy_bh
999999

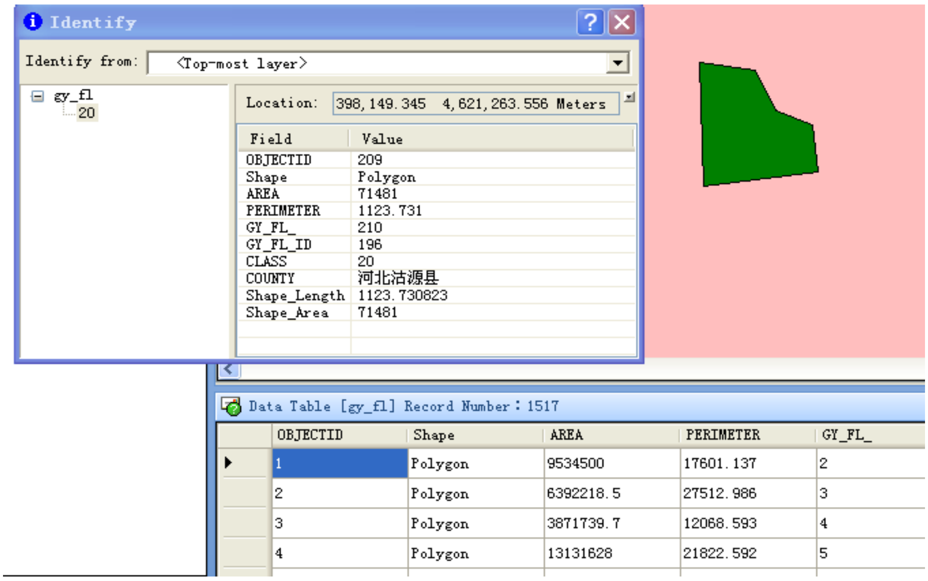
Location: 398,177.256 4,621,287.978 Meters

Field	Value
OBJECTID	14
Shape	Polygon
AREA	71481
PERIMETER	1123.731
GY_BH_	15
GY_BH_ID	14
BCLASS	999999
COUNTY	河北沽源县
COUNTYCODE	130724000
Shape_Length	1123.730823
Shape_Area	71481

Data Table[gy_bh_分类图层与变化图层属性不一致] Record Number:33

OBJECTID	Shape	AREA	PERIMETER	GY_BH_
1	Polygon	3340000000	1383252.2	2
14	Polygon	71481	1123.731	15
15	Polygon	426784	2820.233	16
38	Polygon	103526	1336.241	39

Figure 5.2-6 One polygon in classification data



Location: 398,149.345 4,621,263.556 Meters

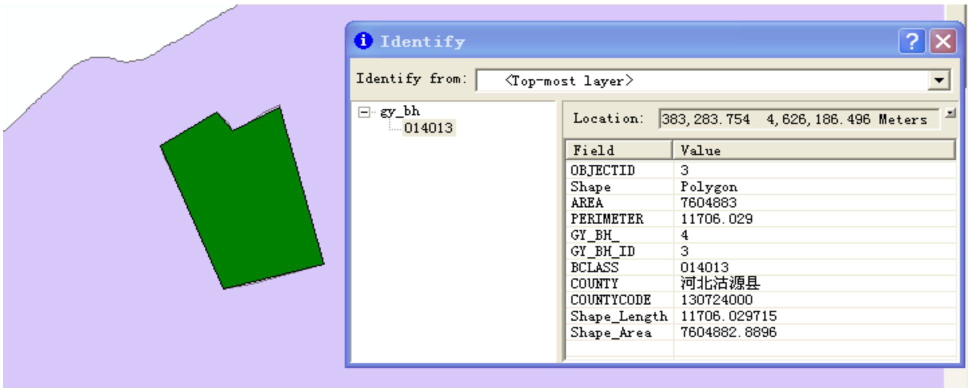
Field	Value
OBJECTID	209
Shape	Polygon
AREA	71481
PERIMETER	1123.731
GY_FL_	210
GY_FL_ID	196
CLASS	20
COUNTY	河北沽源县
Shape_Length	1123.730823
Shape_Area	71481

Data Table [gy_fl] Record Number：1517

OBJECTID	Shape	AREA	PERIMETER	GY_FL_
1	Polygon	9534500	17601.137	2
2	Polygon	6392218.5	27512.986	3
3	Polygon	3871739.7	12068.593	4
4	Polygon	13131628	21822.592	5

Figure 5.2-7 Related polygon in change data

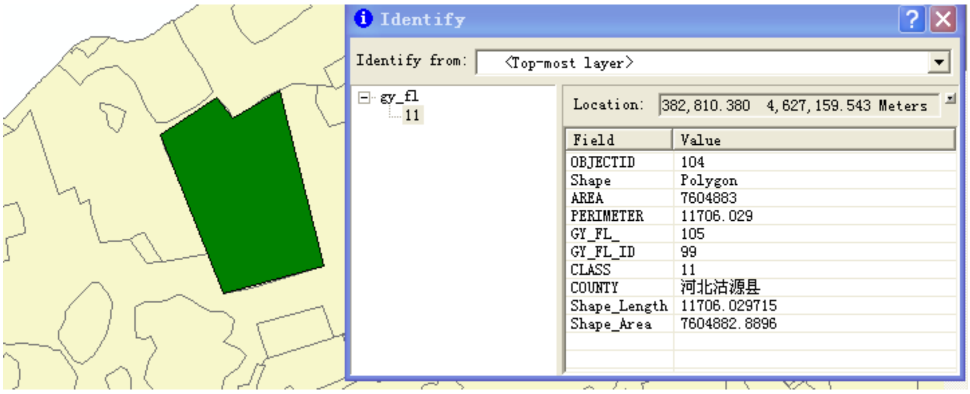
Figure 5.2-8 and 5.2-9 indicate another situation. In land use change data, the target land change from type 14 to type 13, while in figure 5.2-9, the land code is 11, which is inconsistent with 13.



Location: 383,283.754 4,626,186.496 Meters

Field	Value
OBJECTID	3
Shape	Polygon
AREA	7604883
PERIMETER	11706.029
GY_BH_	4
GY_BH_ID	3
BCLASS	014013
COUNTY	河北沽源县
COUNTYCODE	130724000
Shape_Length	11706.029715
Shape_Area	7604882.8896

Figure 5.2-8 A polygon in the land use change data



Location: 382,810.380 4,627,159.543 Meters

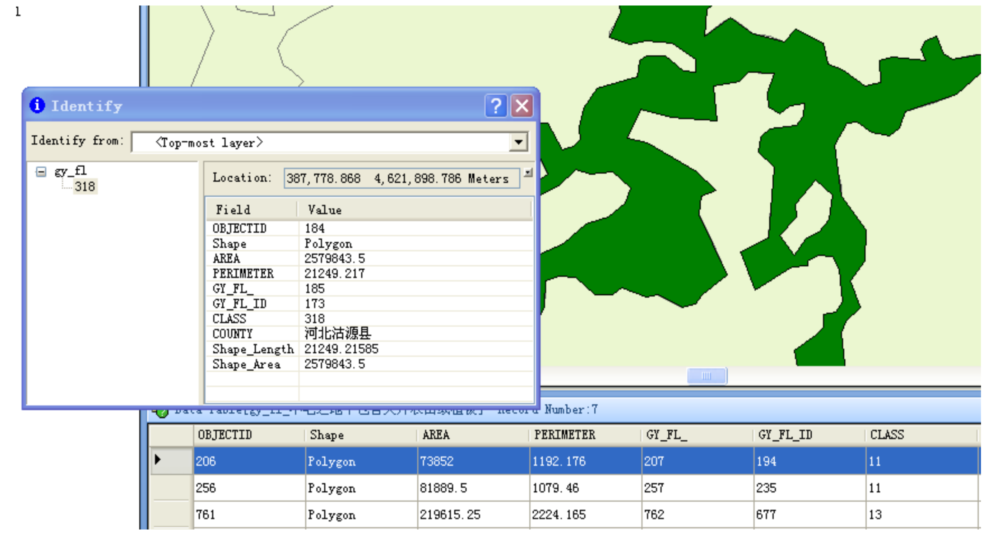
Field	Value
OBJECTID	104
Shape	Polygon
AREA	7604883
PERIMETER	11706.029
GY_FL_	105
GY_FL_ID	99
CLASS	11
COUNTY	河北沽源县
Shape_Length	11706.029715
Shape_Area	7604882.8896

Figure 5.2-9 The corresponding polygon in the classification data

Figure 5.2-10 indicates 7 records of “Desert contains large area of agricultural land or vegetation”. In figure 5.2-11 and 5.2-12, the container polygon is classified as type “31”, which is an unused land, while the inside polygon is farmland because its land type is “11”. This phenomenon is illogical.

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS
206	Polygon	73852	1192. 176	207	194	11
256	Polygon	81889. 5	1079. 46	257	235	11
761	Polygon	219615. 25	2224. 165	762	677	13
1057	Polygon	128874. 5	1504. 807	1058	935	11
1351	Polygon	157896. 48	1656. 04	1352	1492	13
1359	Polygon	636344	3976. 006	1360	1200	13
1392	Polygon	374267. 5	3607. 969	1393	1232	14

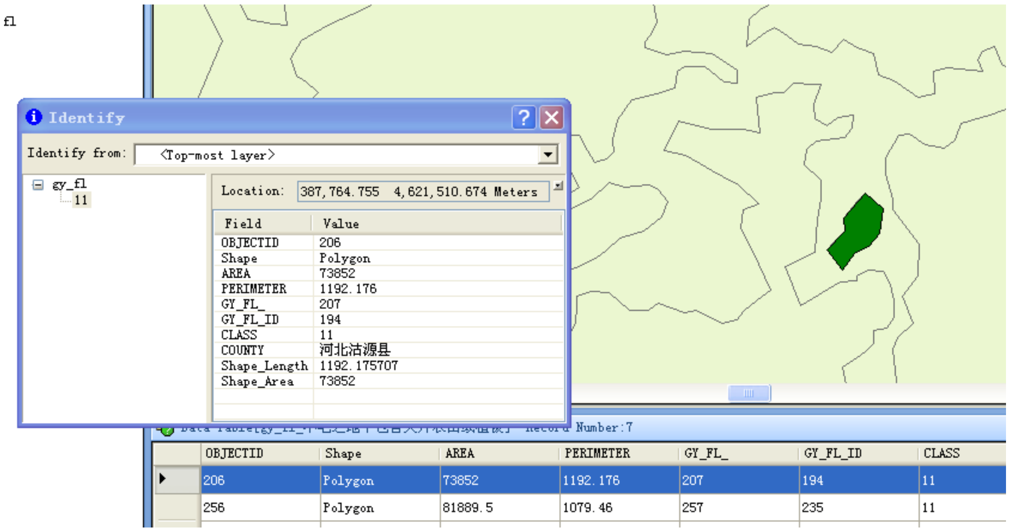
Figure 5.2-10 Result table



Field	Value
OBJECTID	184
Shape	Polygon
AREA	2579843. 5
PERIMETER	21249. 217
GY_FL_	185
GY_FL_ID	173
CLASS	318
COUNTY	河北沽源县
Shape_Length	21249. 21585
Shape_Area	2579843. 5

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS
206	Polygon	73852	1192. 176	207	194	11
256	Polygon	81889. 5	1079. 46	257	235	11
761	Polygon	219615. 25	2224. 165	762	677	13

Figure 5.2-11 Container polygon

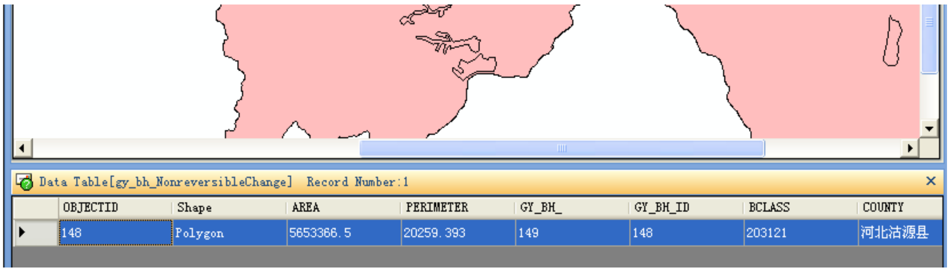


Field	Value
OBJECTID	206
Shape	Polygon
AREA	73852
PERIMETER	1192. 176
GY_FL_	207
GY_FL_ID	194
CLASS	11
COUNTY	河北沽源县
Shape_Length	1192. 175707
Shape_Area	73852

OBJECTID	Shape	AREA	PERIMETER	GY_FL_	GY_FL_ID	CLASS
206	Polygon	73852	1192. 176	207	194	11
256	Polygon	81889. 5	1079. 46	257	235	11

Figure 5.2-12 Inside Polygon

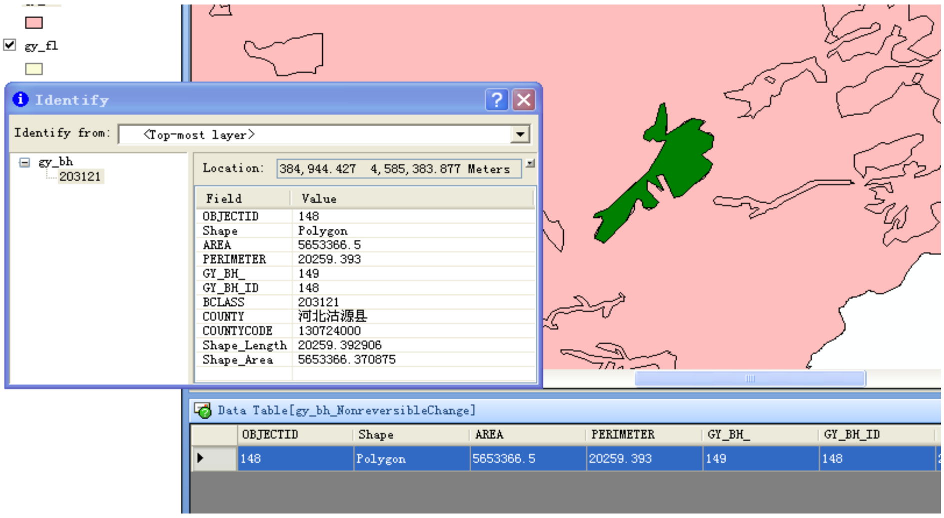
Only one record is identified by rule “Irreversible land type changing”. In figure 5.2-13 and 5.2-14, the land changed from type “203” to type “121”, which means that the land changed from building land to farm land.



Data Table[gy_bh_NonreversibleChange] Record Number:1

OBJECTID	Shape	AREA	PERIMETER	GY_BH_	GY_BH_ID	BCLASS	COUNTY
148	Polygon	5653366.5	20259.393	149	148	203121	河北沽源县

Figure 5.2-13 Result table



Identify from: <Top-most layer>

Location: 384,944.427 4,585,383.877 Meters

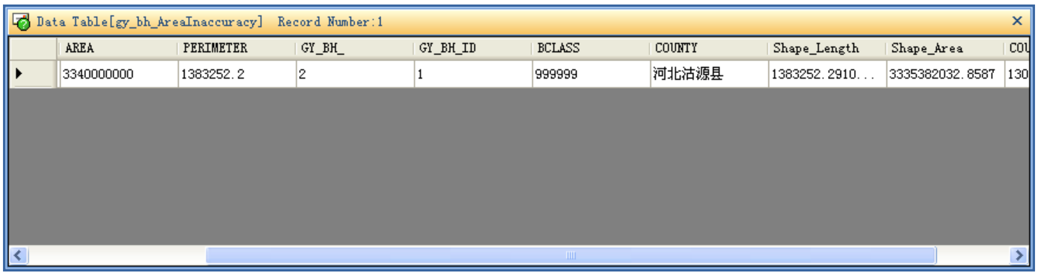
Field	Value
OBJECTID	148
Shape	Polygon
AREA	5653366.5
PERIMETER	20259.393
GY_BH_	149
GY_BH_ID	148
BCLASS	203121
COUNTY	河北沽源县
COUNTYCODE	130724000
Shape_Length	20259.392906
Shape_Area	5653366.370875

Data Table[gy_bh_NonreversibleChange]

OBJECTID	Shape	AREA	PERIMETER	GY_BH_	GY_BH_ID
148	Polygon	5653366.5	20259.393	149	148

Figure 5.2-14 the record contains inconsistency

Rule “Inaccurate area key value” also has only one result. As depicted in figure 5.2-15, area number recorded is 3340000000, while the calculated value is 3335382032.8587. The difference is 4617967.1413, which is much larger than the threshold (Threshold can be set in the system as user’s accuracy requirements).



Data Table[gy_bh_AreaInaccuracy] Record Number:1

AREA	PERIMETER	GY_BH_	GY_BH_ID	BCLASS	COUNTY	Shape_Length	Shape_Area	COU
3340000000	1383252.2	2	1	999999	河北沽源县	1383252.2910...	3335382032.8587	130

Figure 5.2-15 Result table

After applying all 10 data inconsistency detection rules, data with inconsistency have been found out by the rules in the prototype system using KuYuan land use data.

5.3 Analysis of Results

Based on the results of experiment, the data were checked in details to verify whether the results are inconsistency. Details of detection results from ten rules are as follows:

Table 5.3-1: Experiment result

NO.	Detection rules	Inconsistencies detected	Total percentage
1	Same type lands without merging	254	16.74%
2	Matching problems between land use change and classification data	33	2.18%
3	City region contains large area of agricultural land	0	0.00%
4	Desert contains large area of agricultural land or vegetation	7	0.46%
5	Relative area number of polygons breaks common senses	0	0.00%
6	Vegetation growing stage sequence detection	0	0.00%
7	Irreversible land type changing	1	0.07%
8	Inaccurate area key value	1	0.07%
9	Inaccurate perimeter key value	0	0.00%
10	Attribute field inspection	0	0.00%

From the above table, the same class land merging problem is very common, which is about 17% records in the experiment data. Even though the experiment data is the final land use data products obtained from government, which has passed normal quality control, experiment result shows that it still has considerable inconsistencies. Here we try to overlap experiment land use data with image map (issued in 2010 by Bing Map) to roughly model some possible inconsistencies in figure5.3-1.Via this view, we can know the impact of inconsistency more clearly.

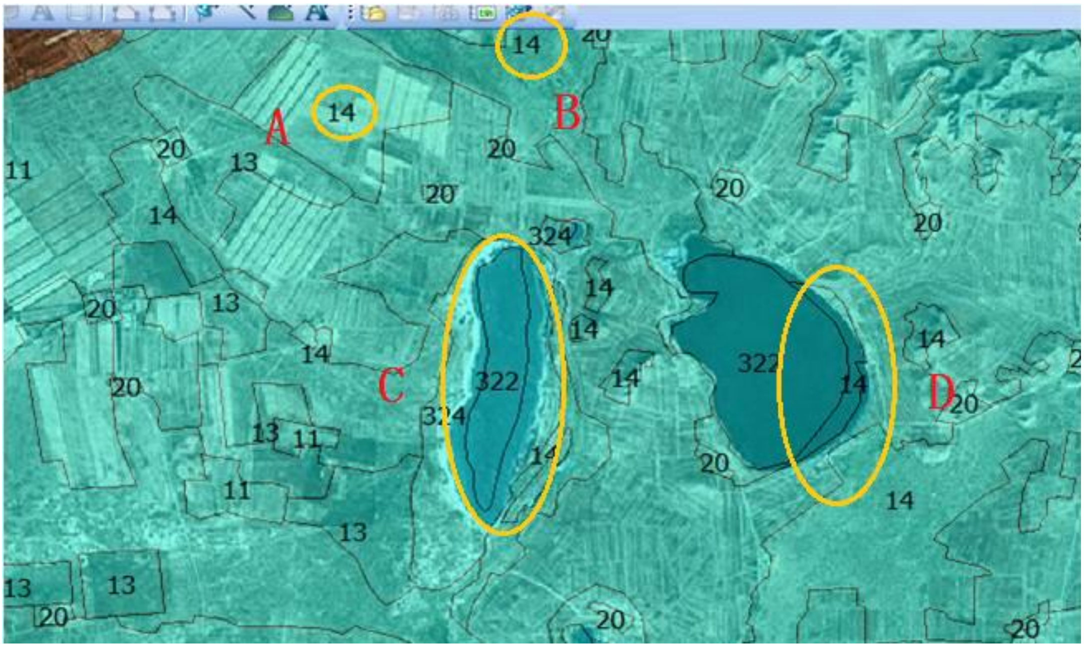


Figure 5.3-1: Model of same inconsistencies

From figure 5.3-1 we noted that we can use image map as reference to find inconsistency manually, but obviously automatic method from this study is much more convenient. However we could still use image to evaluate if the rules can really find out inconsistency and the conflict between data and the real world. The chosen image should be issued in time with the land use data as close as possible. As we do not have the image data issued in similar time with experiment land use map, here we use the oldest image we can find to test and model. Land A and B were both type 14, and should be merged as single one. But obviously, the image was issued about 6 years later than our experiment data, land A has changed much more. Land C told us what area inconsistency might be. When you find a land with area value inconsistency, in true environment its data conflict may like land C and its image. Land D is type 14 now, but obviously it will change to 322 in next round of updating. If land D was updated into type 322, then it will face same class land merging problem. If land D was not merged into the lake, one lake in real world would be divided into two adjacent lakes in land use data.

Every land in experiment map is checked by the 10 rules we listed. For data with inconsistency, we also zoom into the target area to check if the target land really break related rules. If it does not obey current rule, it will be treated as inconsistency. It is relatively easier to check attribute inconsistency, while rules for checking geometric inconsistency need support from spatial overlapping and additional calculation. Therefore, detection of geometric inconsistency can only be feasibly under help of GIS spatial analysis techniques. Detection for logical inconsistency is more complex, because logic of land use data can be various. For instance, for the agricultural land, vegetation growing violates the major logical principle; while for urban area, vegetation growing logic could be useless. Instead, its data logic may be related to ownership of land, demographic meaning of attribute, etc. For some special cases, special logic is needed. For an example, for the land which is close to the airport, the logical height of buildings in these lands must be lower than a limited value.

During the remote sensing image interpretation, for instance, the result is highly relied on the interpreter’s subjective judgment, knowledge and experience if automatic interpretation was not applied. Interpreters may misunderstand the image as human beings or have difference in opinion with other interpreters. Incorrect judgment will lead to false change information or false classification result. On the other hand, if the automatic interpretation method was applied, none of current automatic feature identification and classification algorithm is considered as perfect. During unsupervised classification, features in the shadow or backlight face may be misidentified as other features. Vegetation on a mountain is normally classified into two classes. That is because one side of the mountain is faced to the sun, but the other is in the shadow. During the supervised classification, misidentification and incorrect classification also happened if supervised samples were not selected typically or sufficiently.

Generally, the solution proposed by this study for attribute inconsistency detection performed well in the experiment. Geometric or spatial inconsistency detection needs support form GIS spatial analysis, while the rules can be fixed into a standard technological process. For the most important and complex inconsistency detection, the logical detection rules cannot be fixed into standard steps or aspects. Therefore, it needs manual analysis on the reality of the target lands. Different lands may cause different logical inconsistency, and need different detection rules.

6. Conclusion

6.1 Summary

Some errors or logical inconsistencies in land use data are hard to be identified and may be ignored by normal data quality control. This study developed rules and methods for detecting spatial inconsistency in land use data updating by analysis on geometric, attribute and relationship information of land use data.

Basically, the study is divided into two parts. The first part is to analyze the land use data itself to find the examples of data inconsistency. The second part is to establish a set of rules to detect data inconsistency, based on the analysis which has been done in the previous step.

The first step is to analyze data features and data content to generalize the forms and reasons of inconsistency occurring in land use data, based on analysis over single record and interaction/relations between multiple records.

The second step is to analyze the origin of data inconsistency and to establish a set of rules and method, which could detect the data inconsistency. These works are all based on the result of data analysis in the first step. Moreover, inconsistency detection rules are just concepts, which need some detailed algorithms to be designed to implement the concept of detection rules.

Finally, a prototype system was developed to offer an environment for implementing the designed detection algorithms. The prototype system is a convenience framework using DLL interface technique so that the new rules or detection algorithms can be added into the prototype system to meet the needs of experiments or testing new designed rules. Besides, an expansible system framework is very important for the further study to expand discussion over the influence assessment and elimination of data inconsistency, which are the top two levels of inconsistency handling model.

During the stage of method design, analysis over the land use data is the most important section. Besides, before the development of the prototype system, extracting effective rules for inconsistency detection is also a key point.

This study focuses on inconsistency analysis and handling method in land use data. Land use data is required to be produced and need continuous updating for the most of governments. In updating task, many reasons may cause inconsistency. Additionally, large land use data updating project is often divided into smaller parallel tasks, in which different updating methods were used and different standards were applied. As irregularity updating results from parallel tasks were combined together to establish database, inconsistency is easily to occur. Therefore, appropriate correction process must be taken after land use data updating. This study extracted rules from inconsistency examples, and used those rules to detect inconsistency in another land use data. It was implemented by prototype system and experiment which demonstrated the ability of the inconsistency detection rules.

The proposed rules mainly focus on logic or semantic inconsistency of land use data, as well as the format consistency. This study did detect potential errors of the meanings of data, not just the data itself. The rules can be applied on semantic of data, regardless the format or structure of the data

6.2 Conclusion

Based on experiment results and analysis, following conclusions are stated:

1) It is possible to detect inconsistency in land use updating data, even for some complicated logical inconsistencies.
2) Detection on logical inconsistency is a crucial issue for better quality control on land use data.
3) Detection rules for logical inconsistency of land use data must consider the feature of land. Specific rules should be made for the specific land, while no rules can be designed to detect all kinds of logical inconsistency.

6.3 Limitation

After inconsistency detection has been implemented, the next job should be correcting data inconsistency automatically, which means some method is needed to modify the data to correct status. However, due to the complexity of data inconsistency, it is difficult to implement automatic data correction function. There are two main issues that cause the difficulty. One situation is impossible correction. Some data inconsistencies can be detected but cannot be corrected reasonably without additional information. If some invalid value in attribute table were found during detection, it is impossible to modify to a valid one when the value is unknown. The other situation is that correction can be implemented but may cause some new data inconsistencies. One instance has been illustrated in figure 5.2-4. It shows that a land has been changed from type 14 to type 13, but classification data shown in figure 5.2-3, the corresponding polygon was not marked as 13. If this inconsistency was corrected by modifying to 13, its adjacent polygons may also be marked as type 13, which would cause the new inconsistency, the same type polygons without merging.

6.4 Future work

Inconsistency rules established in this study can only detect the inconsistencies in land use updating. Further analysis of land use data collection, classification, standardization, as well as spatial relations, should be done to formulate the inconsistency detection rules or to improve the rules. In detail, at least a new rule for boundary inconsistency checking should be developed since some available functions can be improved.

Another function should be paid attention to is print- out of analysis report, which can illustrate the inconsistency detection results and give a brief statistic overview of input land use data.

Finally, spatial inconsistency correction functions should be developed; even it is difficult to implement a perfect solution for correction.

References

Barnard, C. (2005). Employment growth, population growth and electronic technologies as determinants of land use. *Land Use Problems and Conflicts: Causes, Consequences and Solutions,* Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 33–49.

Bell, P. (2005). Spatial analysis and applied land use research, *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 151–163.

Bergstrom, C. (2005). Postproductivism and changing rural land use values and preferences. *Land Use Problems and Conflicts: Causes, Consequences and Solutions,* Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 64–66.

Castilla, G. & Hay, G.J. Uncertainties in land use data. *Hydrology Earth Syst. Sciences Discussions*, no. 3, pp. 343–347

Chen, X. W. (2005). Establish and consistency maintain for topological relationship in land use database(土地利用数据库综合中图斑拓扑关系的创建和一致性维护). *Geomatics and Information Science of Wuhan University*, no.30 (4), pp. 370-373.

Cheng, Y., Bao, H. Y. & An, X. F. (2005). Using 3S to update land use map(基于 3S 技术的土地利用现状图更新). *Journal of Liaoning Technical University*, no.24, pp. 50-52.

Cockcroft, S. (1997). A taxonomy of spatial data integrity constraints, *GeoInformatica*, no. 1(4), pp. 327-343.

Cockcroft, S. (1996). Towards the automatic enforcement of integrity constraints in spatial database systems, *The 8th Colloquium of the Spatial Information Research Center*, New Zealand: University of Otago.

Fathi & Alimohammadi. (2010). Inconsistency detection between spatial rules in land use planning application, *Gis Development*, [Online].
Available: http://www.gisdevelopment.net/application/urban/products/ma0793.htm. [Accessed Oct. 11, 2010].

Gaede, V. & Gunther, O. (1995). Constraint-Based query optimization and processing, *First Int. Database Workshop on Constraint Database Syst.*

Geo Way (2008). TD Software Production for land resources profession(TD 国土资源行业应用软件产品系列). [Online]
Available: http://www.geoway.com.cn/Article/ShowArticle.asp?ArticleID=26#2 [Accessed Jan. 14, 2011]

Goetz, J., Shortle, S. & Bergstrom, C. (2005). Contemporary land use problems and conflicts, *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 1–5.

Goetz, J. (2005). Socioeconomic and health outcomes of land use, *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 133.

Hanley. D & Watson, F. (2005). Land use problems: A European perspective. *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 26–32.

Hong, A. L. & Zhao, Z. (2009). Using Walk platform to design land use data quality control system(基于 Walk 平台的国土资源空间数据质量检查系统设计). *Land & Resources*, pp. 46-49.

Jones &, Stokes. (2005). Snohomish county land use consistency draft report, *Snohomish County*, pp. 2–7.

Libby, W. (2005). Rural land use problems and policy options: Overview from a US perspective, *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 9–15.

Li, J. P. (2008). Analysis and discussion over consistency checking between interior mapping and site surveying of land use updating（土地利用更新调查成果内外业一致性检查方法的分析与探讨）. *Modern Surveying and Mapping*, no. 31(5), pp. 34-37.

Li, Z. J., Liang, S. P. & Wong, Z. G. (2007). Application of land use database quality control and checking system(土地利用数据库的质量检查与成果预检系统的应用). *QingHai Technology*, no. 3, pp. 78-79.

Lin, A., Huang, R. & Zuoteng, Y. (2001). Discussion for the New China Land Use Classification System. *Land Resource Technique Management*, pp. 7-10.

Marshall, P. & Shortle, S. (2005). Urban development impacts on ecosystems, *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 77–78.

Plumer, L. & Grogre, G. (1997). Achieving integrity in geographic information systems: Maps and nested maps, *Geoinformatica*, no. 1(4), pp. 345–367.

Seidl, F. (2005). Failing markets and fragile institutions in land use: Colorado's experience, *Land Use Problems and Conflicts: Causes, Consequences and Solutions*, Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 50–63.

Servigne, S. (2000). A Methodology for spatial consistency improvement of geographic databases, *Geoinformatica,* no. 1(4), pp. 7-34.

Stedman, C. (2005). Sense of place as an integrated framework for understanding human impacts of land use change. *Land Use Problems and Conflicts: Causes, Consequences and Solutions,* Goetz, J., Shortle, S. & Bergstrom, C., Eds. New York: Routledge, pp. 120–128.

Tang, Y. (2006). Discussion of land use data updating using 3S technology(应用“3S”技术更新土地利用数据的探索). *GEOSPATIAL INFORMATION,* no. 4 (3), pp. 60-62.

Xiaomei, Y. & Ronqing, L. Q. (1999). Change Detection Based On Remote Sensing Information Model and its Application to Coastal Line of Yellow River Delta. *Earth Observation Center,NASDA,China*.

Zhang, S. H. (2008). Discussion of quality checking of land use database(土地利用数据库质量检查方法的探讨). *QingHai Land Method,* pp. 42-43.

Zhang, Y. M., Liu, P. Y. & Ge, Y. F. (2009). Discussion of land investigation data quality checking(土地调查数据质量检查方法的探讨). *Surveying and Mapping of Geology and Mineral Resources*, no. 25 (3), pp. 31-33.

Zhu, Y. Q., Zhu, S. C. & Liu, H. C. (2008). Research and implementation of quality checking in land use data(土地利用数据质量检查方法研究与实现). *Scientific and Technological Management of Land and Resources*, no. 25(6), pp. 66-70.

CA-Webs. (2009). Melrose drive extension project draft environmental impact report, *City of Oceanside*, pp. Appendix B:1–27.

Available: http://www.ci.oceanside.ca.us/gov/dev/eng/meloseeir.asp [Accessed Jan. 10, 2011]