World Applied Sciences Journal 10 (Special Issue of Tourism & Hospitality): 14-23, 2010
ISSN 1818-4952

A Geographic Information System Based Approach for Mapping Tourist Accommodations in the East Coast States of Malaysia

Abdul Ghapar Othman, Badaruddin Mohamed, Azizi Bahauddin, Ahmad Puad Mat Som and Shida Irwana Omar

School of Housing, Building and Planning,
Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia

Abstract: Geographic Information System (GIS) has been recognized as a useful tool by a wide range of disciplines in managing, storing, analyzing and visualizing the spatial and non-spatial data. This paper addresses the use of GIS technology in developing a database and analyzing data associated with the tourist accommodations in the east coast states of Malaysia, namely; Pahang, Terengganu and Kelantan. Spatial data on lodging establishments, road networks, administrative boundaries, tourism products and water bodies were digitized into ESRI's ArcView system, whilst, textual data on lodging establishments containing contact information, facilities and status of operation were stored in a time-series framework. Both spatial and non-spatial data were merged via a joining process. Selected components of lodging establishments in the study areas were mapped as data layers in GIS. The outcomes of GIS analysis demonstrate the spatial visualization of the lodging evolution and geographical distribution trends within the study areas. This study provides tourism planners, local authorities, hotel entrepreneurs and potential developers with insights into the planning and constructing of new tourist accommodations according to potential development sites/areas.

Key words: Geographic Information System · Tourist accommodation · Spatial data · Textual data · Spatial visualization · East coast states of Malaysia

INTRODUCTION

Accommodation is one of the key supply components in tourism industry. It provides fundamental services to tourists, which is part of essential ingredients of tourism experience [1]. Tourists are attracted to visit destinations that offer appealing tourism products and quality services such as accommodation. It is not surprising that the lack of suitable and sufficient accommodation may well be a deterrent to tourists, especially holidaymakers wishing to visit a destination [2]. As noted by Sharpley [3], the success of tourism destinations is largely dependent upon the appropriate development of the accommodation sector, which influences the ability of destinations to survive in this competitive global tourism market.

Accommodation for tourism developed sporadically around the country. In Penang, hotels were developed around Acheeh Street in late 18th and 19th centuries for the Mecca bound pilgrims. Situated mostly in cities like Kuala Lumpur, Penang and Melaka, urban tourist accommodations were mostly owned and managed by Chinese who mostly resided in urban areas. While hill resorts like those found in Cameron Highlands and Fraser's Hill, were developed by the British as getaways for their residents in Malaya (for Malaysia).

Throughout the 80s and 90s, tourist accommodation development has been based on speculations and not really entirely based on realistic trend analyses. These were due to lack of reliable data to support the feasibility studies and project development proposals. In the absence of proper development guidelines, the previous laissez-faire approach to hotel construction has resulted in hotel stock to be concentrated in a few locations, particularly in the urban settings. As a result, there was an over supply of hotel capacity in some areas. This will give rise to a decline in average occupancy levels and average achieved room rates and will have the inevitable negative impacts on hotel profitability.

Corresponding Author: Abdul Ghapar Othman, School of Housing, Building and Planning, Universiti Sains Malaysia, 11800 Minden, Penang, MALAYSIA. Tel: 604-653 2502; Fax: 604-657 6523. E-mail: ghapar@usm.my

Thus, a reliable and continuous study is important and needed to steer correct supply of tourist accommodation that matches the market demand as to avoid unnecessary surplus of the market. This is especially true when tourism is not only a fragile sector but also a vibrant one, changes through time. So far, there has been no comprehensive study that details up this industry. The existing data is found to be dispersed and lacking. At the same time, there has been little attention on the spatial visualization and mapping results for tourist accommodations. Even though several organizations such as the country's official marketing arm, the Tourism Malaysia and also the Malaysian Association of Hotels (MAH) have undertaken efforts to collect and to compile the directory of accommodation, the listings, however, are not thorough and include mainly official starred lodgings or members of MAH. Furthermore, lodging surveys that often carried out by the Tourism Ministry and hotel association tend to study common aspects of the accommodation sector, such as room occupancy, hotel guests and customer satisfaction. These outputs, however, are unable to demonstrate the real picture of lodging industry, especially in assisting relevant authority and potential developers in planning the hotel sector. Hence, this study put efforts to develop a practical methodology for mapping tourism accommodation establishments. This methodology undoubtedly provides technical support for tourism planners, relevant authorities, hotel entrepreneurs and potential developers to plan and construct new tourist accommodations according to potential development sites/areas. It is the intention of this paper to present a technical framework and key steps for conducting spatial analysis and visualization of tourist accommodation data, illustrated through a case study of the three east coast states of Peninsular Malaysia, namely; Pahang, Terengganu and Kelantan. The Geographic Information System (GIS) is used as a useful tool in data collection, spatial analysis and result mapping.

Research Methodology: The technical framework and four key steps for conducting spatial analysis and visualization of tourist accommodation data are illustrated in Figure 1. Determining the objectives and the scope of this study (Step 1) is the first and the most fundamental step in the methodology process. It helps formulate the tasks and procedures required to achieve the objectives. Step 2 which is data collection from primary and secondary sources provide basis data for database development. The spatial analysis involves joining data (Step 3) and data verification (Step 4) for result mapping.

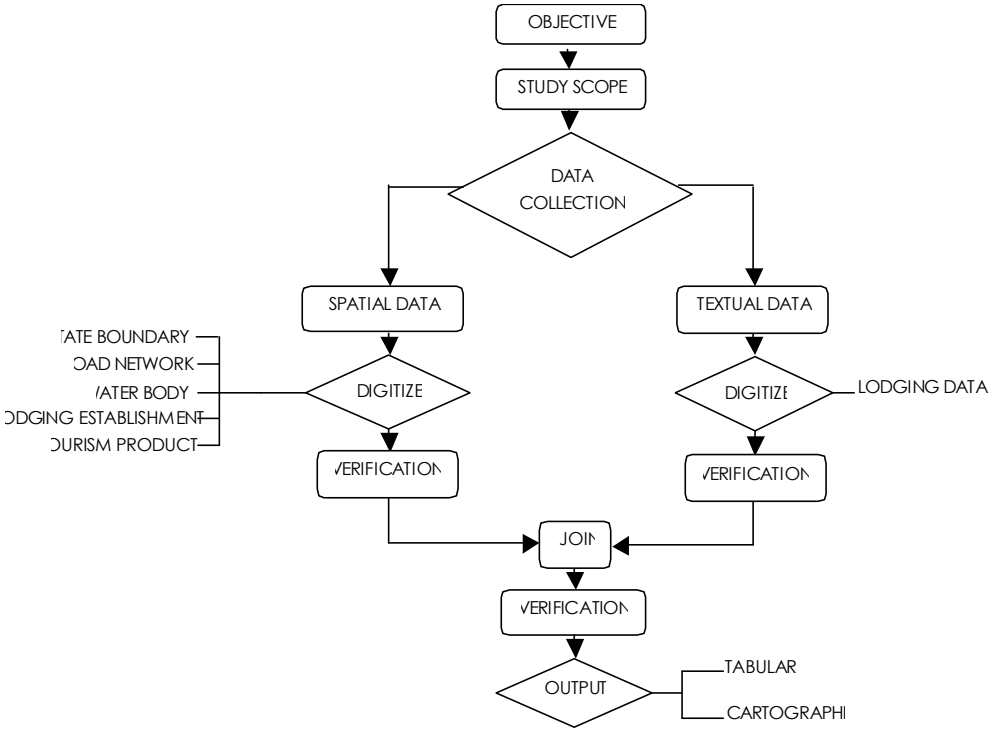

Fig. 1: Technical framework for mapping tourism accommodation establishments

Determining Objective and Scope: As the main objective of this study is to map tourist accommodation establishments in the east coast Malaysia, GIS was utilized in this study as a useful tool for managing, storing, analyzing and visualizing the data associated with the tourist accommodations which include geographical position (spatial data) of lodging establishments and its attributes (textual data). This study focuses on all types of tourist accommodation establishments in the study areas which also include both ranked and unranked establishments at all locations such as urban areas, rural, islands, highlands and etc. The data inventory consists of lodging establishments established since 1970s to year 2008. However, the unavailability of comprehensive yearly lodging information in this country has limited this study to concentrate only to the years 1971, 1975, 1993, 1998, 2001, 2004 and 2008.

Data Collection: This study embarks on an extensive review and comprehensive data collection on lodging industry in study areas. The process of gaining multi-source information involves primary data as well as secondary data. The primary data were collected through series of field works and interviews with hotel managers at selected lodging establishments. During the field works, data on the position of lodging establishments were collected using handheld Global Positioning System (GPS) whereby the positional coordinates was recorded in reference to the West Malaysia Rectified Skew Orthomorphic (RSO) coordinate projection system. The accuracy and reliability of the GPS reading varies from 1 - 5 meter from its actual position. According to Dana [4], GPS accuracy is affected by a number of factors, including satellite positions, noise in the radio signal, atmospheric conditions and natural barriers to the signal. For example, noise can create an error of between 1 to 10 meters due to interference from something near the receiver or object that produces similar frequency. For this study, however, the accuracy of between 1 to 5 meters is acceptable.

On the other hand, the secondary data was obtained from publications or reports by the government agencies and tourism-related organizations such as Ministry of Tourism, Malaysia Tourism Promotion Board, Ministry of Finance Malaysia, Malaysian Association of Hotels and Ministry of Culture, Arts and Heritage. Examples of secondary resource materials include established government reports (e.g. Structure Plans, Local Plans, City Master Plans and tourism property stock in Malaysia), tourist accommodation directory, census data on hotels, tourist/travel maps (papery map), newspaper articles, biographies, travel guide books or pamphlets and budget accommodation guide. In addition, websites of lodging establishments are another good source of secondary data. For establishments which data could not be acquired from the secondary sources, the interview technique was applied either on-site or by telephone. Data acquired from primary or secondary sources were then separated and categorized into two different components, namely; spatial data and textual (attribute) data. Textual data contains information pertaining to lodging's name, address and contact information (website address, phone/fax numbers), year of establishment, rating, support facilities and other related information.

Development of Spatial Data: Spatial data or also known as geospatial data are information on geographic feature in relation to its location and shape, usually stored in the form of coordinates and topology 5]. The data consists of information about the relationships of entities in space; facts about the real world organized geographically; the location, shape of and relationships among geographic features which are usually stored as coordinates and topology [6]. Since locations of most lodging establishments in this study are closely linked to roads, the development of road network spatial data is crucial in determining the correct locations of each establishment.
The road network is used as the main base map for this study. Other spatial data such as the administrative boundaries (state boundary) and water bodies (rivers and lakes) are nevertheless important information to assist in positioning the correct locations of lodging establishments. For the purpose of this study, five spatial data has been identified as follow:

- Road network (polyline feature)
- Lodging establishment (point feature)
- Tourism product (polyline and polygon features)
- State boundary (polygon feature)
- Water bodies (polyline and polygon features)

Preparing the Base Map: Maps of the road network in the study area were obtained from several sources including the Survey and Mapping Department Malaysia, Public Works Department of Malaysia, Department of Town and Country Planning Peninsular Malaysia, local authorities and Google Maps. Most of these maps have state boundaries which are very useful when preparing a base map. They were cross-checked with each other to maximize accuracy and to generate detailed state boundary and road network for the study area.

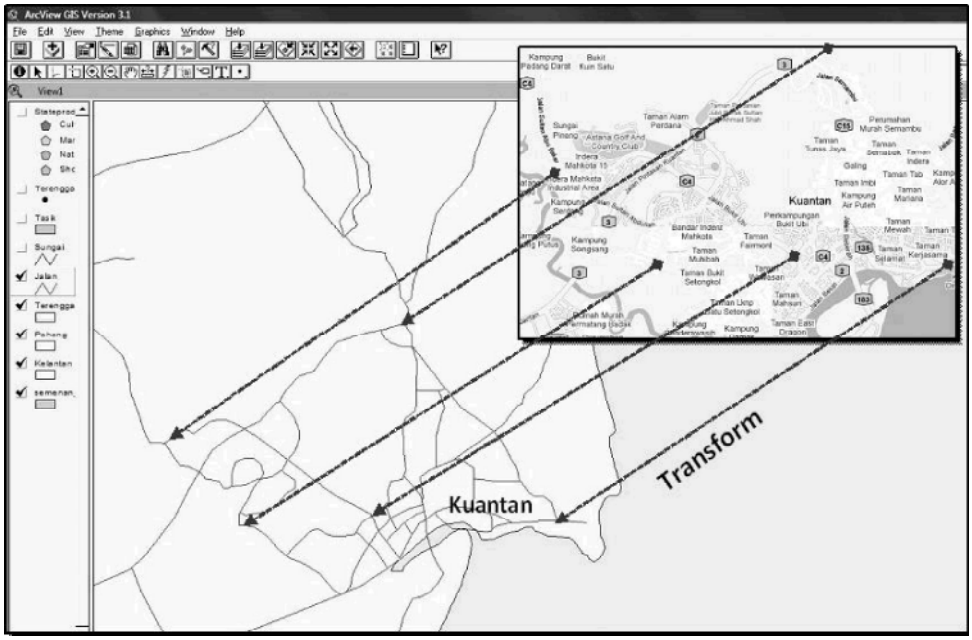

Fig. 2: Georeferencing process with five control points

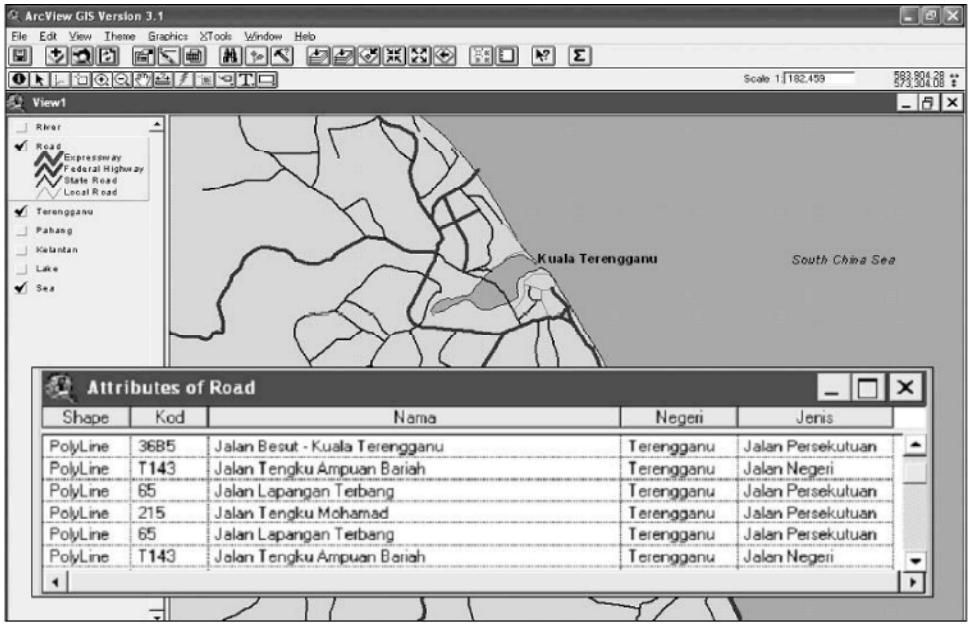

Fig. 3: Map and attributes of road network in Kuala Terengganu, Terengganu

Except for Google Maps whose data is available in digital image (raster data), the others are in the form of papery maps with the West Malaysia RSO coordinate grids on them. These grids are very important for the georeferencing process which will be discussed later in the paper. The papery maps were scanned; converted to digital image format (.JPEG) and imported into ArcView GIS 3.1 for georeferencing and digitizing processes. Georeferencing is a process that defines the existence of a feature in physical space. According to Green 7], this process involves scaling, rotating, translating and deskewing graphical information, particularly raster images to match the geographical coordinate system. It is also referred as a process of registering an image in file coordinates or page coordinates to a file in real-world coordinates particularly in a specific coordinate system, map projection and datum.

The digital image containing data of the state boundary and the road network with the coordinate grids were loaded into ArcView GIS 3.1 which were then georeferenced. Since they are in page coordinates, they need to be georeferenced to the West Malaysia RSO coordinate projection system. Using the coordinate grids on the digital image maps, at least four known control points (x,y coordinates) from both the image page coordinates and the West Malaysia RSO coordinates are registered using a third party application extension in ArcView GIS 3.1. Once the control points were registered with acceptable Root Mean Square (RMS) error, the image was transformed to its correct position based on the West Malaysia RSO coordinate system (Figure 2). Digitizing of the state outline and road network were done by tracing line features on the image and saved on separate files (themes or layers). Attributes of the road such as Identifier (ID), Road Name, Road Category (local, state, federal, expressway) and State (Kelantan, Terengganu, Pahang) were keyed in using the table menu in ArcView GIS 3.1 (Figure 3).

Using the same digital image and overlay with the completed state boundary and road network layers, similar steps were also applied for the development of the water body map. Rivers were digitized as polyline features while lakes were represented by polygons. Due to these two different features, the water body map was separated into two different layers (polyline and polygon features) with different attributes.

Digitizing the Spatial Data-lodging Establishments and Tourism Products: The location of lodging establishments is the most important data in this study. The data was collected by using GPS through field work or information derived from tourist maps, brochures and tourist accommodation directory. Each lodging establishment was assigned a unique identifier (ID) and was categorized according to the state it was located. This unique ID will later be used to merge the spatial and textual data. The GPS coordinate system was set to West Malaysia RSO projection before reading on lodging locations were recorded in the form of x and y coordinate (x,y). This coordinate was processed using the GPS software and converted to a format readable by ArcView GIS. Upon loading into ArcView GIS, this coordinate (x,y) was converted into graphic form which was represented by a point feature. Each lodging establishment will be shown as point on the map. In addition to this, some of the lodging establishments were digitized directly into ArcView GIS using accurate and reliable tourist maps and brochures. The completed road network map was used as a guide when digitizing these locations. In such situation, the use of GPS was not required. The tourist maps and brochures are a good source of references for cross checking data acquired from the GPS to ensure every lodging establishment was recorded. Each lodging establishment was assigned with a unique ID and its name was keyed in as attribute. This spatial data of lodging establishment will later be merged with its textual data developed using Microsoft Access. Figure 4 shows the distribution of lodging establishment in the state of Pahang for the year 2008.

Similar steps above were repeated to develop the spatial data for tourism product which was also represented by point feature. After each location of the product was digitized, its attributes were keyed in using the attribute table menu in ArcView GIS. These attributes include:

- ID (Identifier)
- Product Name
- Category (nature, man-made attraction, culture and heritage, shopping and leisure)
- Location (closest city or town)
- State (Kelantan, Terengganu, Pahang)

Verification of the completed spatial data was done by comparing the completed digital maps with the original papery maps. This was to ensure that every single data was captured and its attributes were recorded accordingly. The purpose of this process was to minimize error and maximize reliability and accuracy.

Development of Textual Data: Textual data, sometimes known as attribute data in GIS, can be defined as any readable information including qualitative data that can be counted for recording and analysis purposes. It is also descriptive information of geographic characteristics of feature and their relationships to each other. As discussed earlier, the attributes of lodging establishment in this study were gathered from several sources in the form of hard copy documents. These attributes were checked before being keyed in as textual data into Microsoft Access (Figure 5). Since the scope of this study only focused on lodging establishments existed in the year 1971, 1975, 1993, 1998, 2001, 2004 and 2008, a file was created for the respective years. The attributes of lodging establishment include:

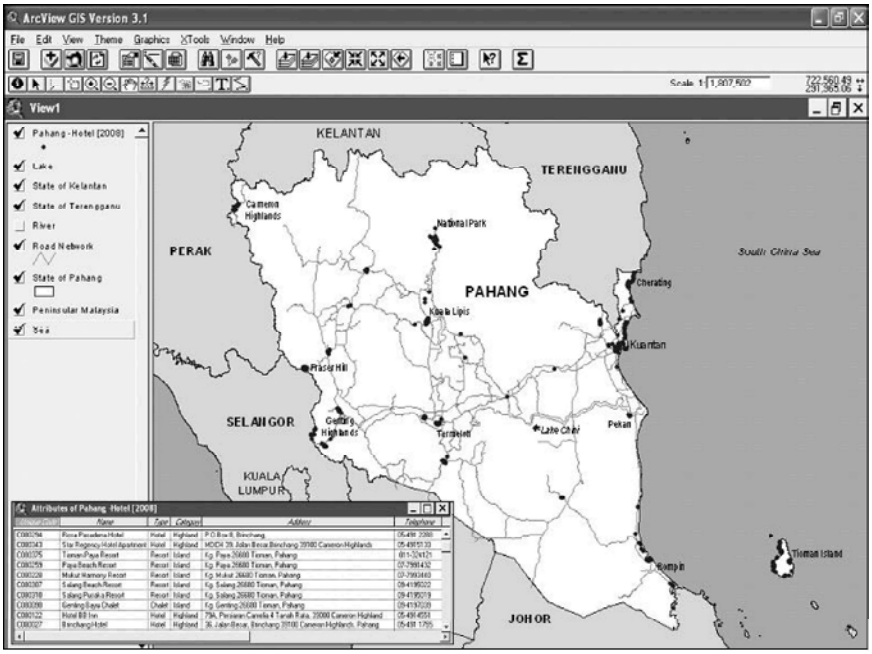

Fig. 4: Distribution of lodging establishments in the state of Pahang (2008)

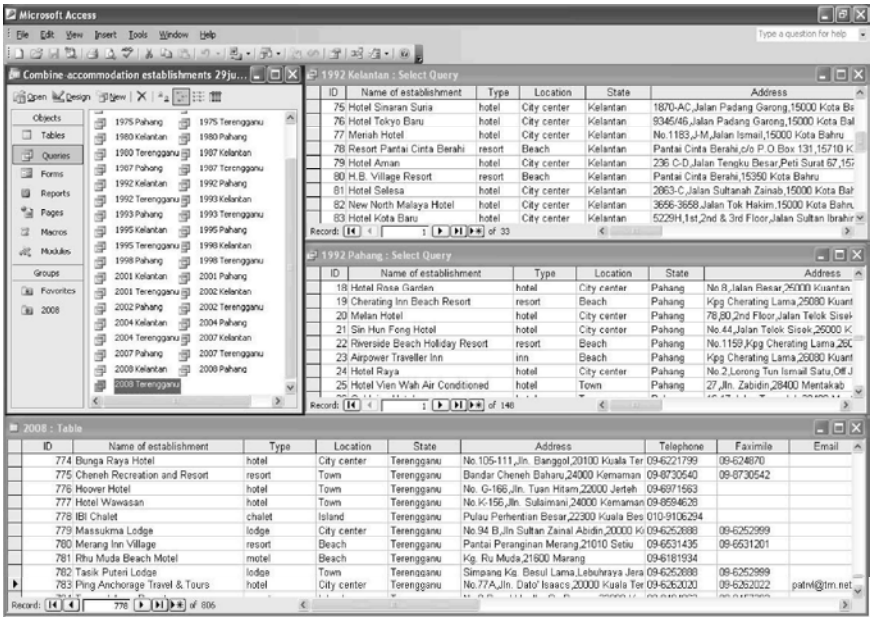

Fig. 5: Database of lodging establishments in Microsoft Access

- ID (Unique identifier)
- Name
- Type (apartment, bungalow, chalet, condominium, cottage, guest house, homestay, hostel, hotel, inn, lodge, motel, resort, rest house, villa)
- State
- Address
- Telephone number
- Email address
- Number of rooms
- Year established
- Facilities
- Hotel rating
- Number of bed
- Status of operation
- Member of Malaysian Association of Hotels (MAH)
- Remarks

It was important that the unique identifier (ID) in this textual data to be the same as those in the spatial data. The ID in both the spatial and textual data will be used to merge these two data sets together. The attributes of the lodging establishment were verified by checking and comparing them to the original sources from where they were acquired.

Joining Data: Data on the lodging establishment was collected and digitized into two types, the spatial and textual data. Spatial data provides geographic information (location) of the establishment while textual data handles its attributes (listed above). Since ArcView GIS utilizes the dbf format for its database file, the textual data of the lodging establishment was converted from the Microsoft Access format to the .dbf format to make it readable in ArcView GIS. The .dbf textual data was then loaded into ArcView GIS. The next step was to open the attribute table file of the spatial data for lodging establishment. Both the attribute table of the spatial data and the textual data will have a common field called ID. The joining process in ArcView GIS will match ID from both dataset and for every matched IDs, it will append records from the textual data into the spatial data attribute table. Before joining these two datasets, a simple count check was done to ensure that both datasets have the same number of records. Once the joining process was completed, the records in the lodging establishment database were checked to ensure they were correct and accurate.

This joining process enables researcher to effectively manage large database that were kept in different format. Textual data for example, was easily updated in computer software like Microsoft Excel or Access as compared to ArcView GIS. However, this software lack of the graphic capabilities which allow data to be published in map form. The joining process in GIS allows attributes of the textual data to be viewed, queried and analyzed by creating links between the spatial and textual datasets. This study allows researchers to study the distribution and evolution of lodging establishment for the past 30 years.

Data Verification: In any GIS project, data verification is an important task to maintain accuracy and to reduce data errors 8]. Although the process is tedious and can be time consuming, it increases the confidence level of researchers when applying the verified data in further analyses. In this study, each record in the lodging establishment database was checked to ensure no missing attribute from the joining process. Missing attribute was displayed as blank which indicates that the attributes from the textual data were not appended to the spatial data during the joining process. Most of the errors found are commonly caused by mismatching of the unique identifier during the joining process. This problem can be solved by removing the join link between the spatial and textual datasets in ArcView GIS. For each record with missing attributes, its ID is cross checked in both the spatial and textual datasets to ensure that they are the same.

Table 1: Summary data of total lodging establishments by state for year 1975, 1998 and 2008

	Kelantan			Terengganu			Pahang		
Lodging Type	1975	1998	2008	1975	1998	2008	1975	1998	2008
Hotel	33	30	46	27	29	64	117	116	203
Resort	-	3	14	-	44	80	3	55	103
Inn	-	8	33	1	5	13	-	10	23
Motel	2	2	5	2	3	4	3	1	2
Lodge	-	2	12	1	5	8	2	6	13
Chalet	-	-	40	2	15	30	1	5	114
Guest House	5	2	19	2	8	9	14	3	13
Rest House	-	-	4	-	2	4	-	4	15
Cottage	-	-	-	-	1	1	4	-	4
Condominium	-	-	-	-	-	-	-	-	2
Bungalow	-	-	-	-	-	-	41	-	63
Homestay	-	-	10	-	-	43	-	-	-
Villa	-	-	3	-	-	-	-	-	3
Hostel	-	2	12	-	-	-	-	-	4
Apartment	-	-	-	-	-	1	-	-	3
Total	40	49	198	35	112	257	185	200	565

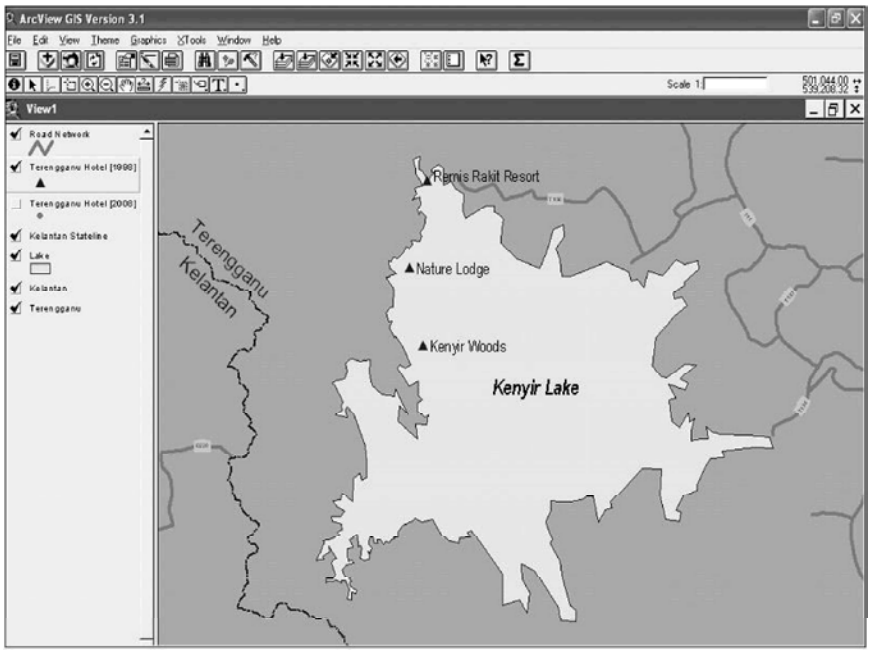

Fig. 6: Lodging establishments within Kenyir Lake in 1998

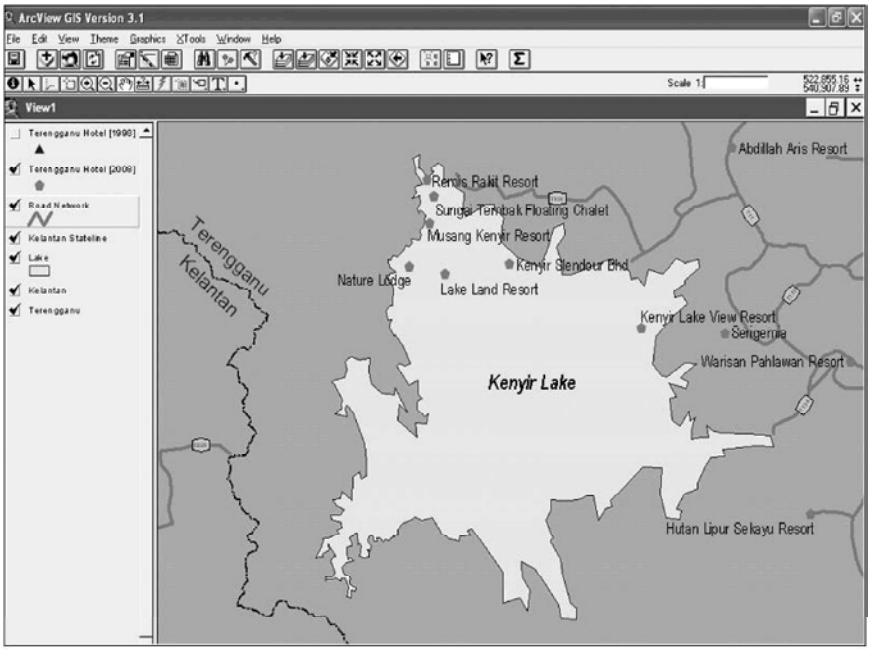

Fig. 7: Lodging establishment within Kenyir Lake in 2008

Once all the error checks were done, both datasets were joined again and this process was repeated until all lodging establishments have no missing attribute.

Research Outcomes: The output of any GIS analysis can be classified into two types: tabular and cartographic output. Tabular output is normally generated from simple GIS procedures such as data manipulation and management. Importing and exporting of data, updating, single and multiple querying are among procedures that generally produce tabular output [9]. In this study, tabular output was generated in the form of summary data and statistics. For example, Table 1 shows a summary data of types of lodging establishments by

state for the year 1975, 1998 and 2008. By comparing the results, a researcher will be able to establish a general understanding of the trend and evolution of lodging establishment in the region.

On the other hand, cartographic output is normally in the form of maps generated from geographical (map) analysis such as data retrieval and classification, overlay, proximity, buffering and other related GIS analyses. A map is represented by cartographic features such as points, polylines, polygons or a combination of these features [10]. It helps a researcher understand the distribution and the relationship of the subject being studied. Overlaying lodging establishment layer with tourism product layer provides researchers with an understanding of types of lodging establishments commonly found within a lake (natural tourism product). By establishing similar overlays but for two different years, it helps researchers study the trend and evolution of lodging establishments located within the vicinity of a lake for a certain period of time. For example, Figures 6 and 7 show the lodging establishments located within the vicinity of Kenyir Lake, Terengganu in 1998 and 2008. In 1998, there were only 2 resorts and one chalet located in the northwestern part of the lake. The number increased to 7 resorts, 2 chalets, one lodge and one homestay in 2008. They were located mainly in the eastern and northwestern parts of the lake; areas with good accessibility and suitable topography for tourist accommodations.

The outcome generated from this study whether it is in the form of tabular or cartographic output, provides researchers with an overview of the distribution of lodging establishments in the study area. The output from the above overlay analysis for example, helps researchers understand the relationships between types of lodging establishments and tourism product. This will enable researchers to formulate improvement strategies for the next phase of the research which was to analyze evolution and future trends of the lodging industry in the study area. In addition, data manipulation and applications of geographic analysis have the potential to generate added-value data to this study.

CONCLUSION

The global growing of interest in tourism sector will definitely demand for an advance and complex application of GIS in this sector. The wide array of GIS application has proven to be a valuable tool for improving the analysis of the evolution of the lodging industry in Malaysia. This paper presented a GIS-based approach for the spatial visualization of tourist accommodations at the state level. The approach and methodology can also be used for mapping tourist attractions, tourist amenities and tourism services. The technical framework involves four key steps. Step 1 (determining objective and scope) is the fundamental step that helped formulate the tasks and procedures required to achieve the objectives. Step 2 (data collection) provide basis data for database development. Step 3 (joining data) and Step 4 (data verification) are a part of spatial analysis processes for results mapping.

Providing map linked to useful information creates an easy way for the researchers to understand and allow them to study the trend of the lodging industry in a spatial context. The use of GIS also allows more profound geographical analysis of lodging data. Therefore, continuous updating and monitoring data is an important action in any GIS application to ensure the reliability and accuracy of data. Nevertheless, acquisition of data, primary or secondary, is a very timely and costly process. When digital data is not free or readily available, the conversion from paper to digital data is a pain staking process.

ACKNOWLEDGMENT

The authors would like to extend appreciation to Universiti Sains Malaysia for granting the Research University Grant [Grant No. 1001/PPBGN/816031] which makes this study and paper possible. We also would like to thank our research assistants, Mr. Mohd Fauzi Sukiman and Ms. Noor ‘Arbaayah Hashim for their kind support and digitization works.

REFERENCE

1. Goss-Turner, S., 1996. The accommodation sector. In P. Jones (Ed.), Introduction to Hospitality Operations Cassell, London. pp: 21-35.
2. Goodall, B., 1989. Tourism Accommodation: A Destination Area Perspective. Built Environ., 15(1): 78-91.
3. Sharpley, R., 2000. The Influence of the accommodation sector on tourism development: lesson from Cyprus. International J. Hospitality Manage., 19(3): 275-293.
4. Dana, P.H., 1997. Global positioning system (GPS) time dissemination for real-time applications. Real-Time Systems: The International J. Time Critical Computing Systems, 12(1): 9-40.

5. Environmental Systems Research Institute (ESRI) 2006. The Geodatabase: Modeling and Managing Spatial Data.
6. DeMers, M.N., 2005. Fundamentals of Geographic Information Systems. USA: John Wiley and Sons Inc.,
7. Green, J., 2006. ArcView How to do it Manual No. 1: Georeferencing for Beginners (Report No. 219). Australia: Maritime Archaeology Department of the Western Australian Museum.
8. Mayberry, M., 2002, September. Data Quality: Before the Map is Produced. Directions Magazine.
9. Chou, Y.H., 1997. Exploring spatial analysis in geographic information systems. Santa Fe, NM, USA: OnWord Press.
10. Falbo, D.L., L.P. Queen and C.R. Blinn, 2002. *Introduction to Data Analysis Using Geographic Information Systems*. St. Paul, Minnesota, US: University of Minnesota.