INTRODUCTION

Electronic data sources have developed considerably in recent years; many electronic platforms are able to provide valuable information regarding engineering topics. One of the most important data sources is the open street map (OSM) platform, providing editable geographic information for most of the world, with different levels of accuracy and at different points in time. Road network mapping requires a high level of effort and accuracy, due to the complexity of the modelling and the amount of information that needs to be included in the feature class. OSM can support road network modelling by providing a different kind of data. In this paper, a systematic procedure was investigated for the production of an automated road network for Basrah city, as a case study for the use of OSM in Geographic Information System (GIS) 10.8 software. Specific spatial analysis tools such as road density and network analysis were also implemented. This study validated a computerised procedure to extract OSM data via two methods of validation and demonstrated the immediate applicability of this data for density and network analysis. The research results show a significant reduction in time and effort required to produce an accurate Basrah city road network using OSM data sources. Density analysis and network analysis show the importance and validity of the produced road network.
Open Street Map (OSM), as a collaborative electronic geographic data source, was first made available online in 2004 in the United Kingdom by Steve Coast [5]. This non-commercial organization specialised in producing different types of information with georeferencing coordinates which can be used to produce paper or electronic maps [5,6]. This platform depends both on data gathered from high-resolution satellite images (see Figure 1) and on publicly-available government data. It offers four different types of data point feature class: polyline feature class, polygon feature class, relations (member), and metadata [7]. Despite being widely used around the world, OSM reliability is still variable from country to country, depending on the accuracy of the data gathered for certain places [8,9].

Zhao et al. (2014) studied the development of OSM road networks in the city of Beijing [10], describing OSM data as an important and qualified source of information for road networks in Beijing. Funke et al. (2015) proposed a method for automatic recognition of gaps in the road network using the OSM database [11]. Brovelli et al. (2017) developed an automated technique to assess the geometric accuracy of OSM data, in comparison to that of an authoritative road database [12]. Despite the success of this technique, its lengthy computational time means that its use is not practical for the slower internet speeds of Iraq. Keller et al. (2020) established and assessed an estimation framework for average speed in rural road networks constructed on a typical ML workflow and OSM road network data [13]. The proposed framework provided good evidence for the importance of extracting information from the OSM platform, however, its sole focus was speed data; OSM provides several forms of georeferenced data for Iraq (and, more specifically, Basrah city), such as roads, buildings, rivers, and green lands. Several inferences can be made from the extensive attempts of previous researchers to work with OSM data, including the fact that expensive software is required for data extraction in some forms of research; additionally, researchers working in an area where road network data is already available (e.g. Canada), would make no attempt to establish automated street map data in Basrah city and examine its validity. Geographic Information System (GIS) 10.8 has been implemented in this study as a key software to explore, investigate, analyse, and visualise georeferenced data. GIS provide great benefits via producing services in location scale and its simplicity empowers its users to implement spatial information and expressive data to make maps, tables, and diagrams. This software appropriately offers several examination tools such as topology, network analysis, route editing, data analysis tools and results display on high resolution updated map [14,15]. For all aforementioned benefits, GIS was implemented in this research as a tool for analysis and visualization.

As part of this research, an automated approach was investigated for the extraction of road data from OSM and the providing of an offline road feature class that can be used in GIS via statistical and network analysis. Actual measurements have been implemented to validate the information extracted from the OSM platform.

![Figure 1 Data gathered from high-resolution satellite images](image-url)
2. BASRAH ROADS DATA

Basrah city suffers from a lack of georeferenced information, such as infrastructure, and buildings. Despite the existence of a GIS department in each local authority in Basrah city, an accurate GIS map for the entire city still not available.

Modern development occurring in Iraq and in Basrah city has produced significant pressures for the road network, such that a traffic jam is a normal phenomenon for the local city road network (see Figure 2) [15]. Therefore, traffic management has become crucial for the solution of this problem. One of the key elements for traffic management is road data; this data is incomplete, thereby preventing any successful traffic management [16,17]. This research will extract most of the road network data using the information provided by OSM via the automated procedure.

3. AUTOMATED EXTRACTION PROCEDURE

Due to the availability of valuable georeferenced data regarding Basrah city in the OSM platform, it is crucial to investigate an automated procedure to extract these data to a standalone feature class that is used in the GIS analysis tool. This study implemented the computerised procedure explained in Figure 3.

It can be seen from the diagram that the extraction process contains several steps as follows:

1. Activate GIS online server via ARC Catalog and load OSM platform, taking care to highlight the study area boundaries to avoid loading unnecessary data and slowing the process down.

2. Build the extension and connection between GIS software and the OSM platform. This connection will enable GIS to identify the georeferenced data in terms of point, line, and polygon.

3. Download the required data from the OSM platform, after recognising exactly what kind of data needs to be downloaded. This step is very important to avoid downloading data without a coordinate system.

4. After the download is complete, extract the data into the required feature class or shapefiles via automated commands.

5. View the extracted data on the ARC Map to start the validation process. If the extracted data meet the validation requirements, then it can be stored as a feature class. Otherwise, the extracted data are useless and it is better to find other data sources.

6. Create the target feature class and start implementing statistical and network analysis.
4. STUDY AREA

This research will focus only on the Basrah City centre, as a case study. The area is located between latitudes 30°24´ N and 30°36´ N and longitudes 47° 43´ E and 47°53´ E as shown in Figure 4.

![Figure 4 study area boundaries.](image-url)
5. RESULTS AND DISCUSSION

5.1. Downloading Barash roads from OSM

After implementing the aforementioned computerised procedure, the Basrah Road network can be downloaded from the OSM platform for specified boundaries. The results of downloading the road network are shown in Figures 5 and 6.

From the Figures above its can be seen that the first data extracted contain all types of feature classes (Point, Line and Polygon) which represent different types of facilities and land use in Basrah city. However, the road network is the sole focus of this study, therefore is crucial to recognise only roads and separate them into a single layer. Due to the similarity of the shape of the roads to other structures like canals, road characteristics such as direction and speed limit were implemented to establish an inquiry that illustrated only road features.

![Figure 5 Row data extraction](image1)

![Figure 6 Recognised Road network only.](image2)
5.2. Validating the extracted information.

As mentioned previously it is vital to validate the extracted information before running any kind of analysis. This study implemented two types of validation as follows:

a) Manual Modelling

The researcher manually modelled small district samples and compared the result with the extraction information from the OSM platform as shown in Figure 7. The green line represents manual modeling and the purple line signifies computerised modelling. It can be concluded that the matching is nearly 99% between the two methods. However, some road extracted from OSM does not exist in a real road network, as illustrated with red circles. This emphasises the importance of validating the extracted network.

![Figure 7 Manual vs computerised modelling.](image)

b) Measurement validation

The second step of validating data is comparing the road properties that are extracted automatically with the actual road characters such as length, width, direction, road name. Length and width properties are implemented as a unit of justification in this research. It has been found that the measurements are quite accurate and can be used to run the analysis. The main reason for this is that the OSM platform depends on high-resolution satellite images to extract this physical information, therefore it can be considered to be reliable information. However, it has been found that other data, such as road names, need to be validated one-by-one according to local information.

c) Topology validation

Topology is one of the GIS software tools that is used to validate the feature class information. It can be seen in Figure 8 that the topology analysis has returned zero errors regarding the ‘Must Not Overlap’ rule.

All validation methods mentioned above illustrate that the extracted information is reliable and can be used for network work analysis and any further traffic management analysis.
Figure 8 Topology validation results

5.3. GIS analysis

Given that the data extracted from OSM is considered to be highly accurate, statistical analysis can be done with a significant level of confidence. Three kinds of analysis have been performed on the extracted data inside GIS software, as follows:

A) Density analysis

This kind of analysis considers identified quantities of some singularity and spreads them across the landscape, grounded by the quantity that is calculated at each position and the unique connection of the positions of the measured quantities [18]. This kind of analysis shows where the road network is concentrated. This information will be highly important to decision-makers in any future road planning [19]. Figure 9 shows the density analysis for Basrah city centre. From the analysis, it can be seen that network length from 3800 m to 4300 m is coloured blue and located in the middle of Barash city; this area will require special attention in future planning. The low-density area with a network reach of 500 m is coloured with yellow, representing most of the city center boundary. These areas are expected to be able to absorb any future road plans, including ring roads.
B) Network Analysis

Network analysis is a GIS analysis tools that is especially applicable to the road network. Network analysis studies the characteristics of the actual and virtual networks, to observe the performance of flows within and around such networks and locational analysis [20]. It implements the methodology of edge-node topology to demonstrate the actual live networks of information. It can be classified into five types of analysis: new route, closest facility, shortest route, services area, and cost matrix [21]. These types of analysis depend on the availability of different types of information in the road feature class. The extracted road feature class from OSM has only length data that can be used directly without any correction or modification. Therefore, only one type of analysis can be performed using network analysis as shown in Figure 10. This analysis can only distinguish between roads and junctions and classify them in the different feature classes. Despite the limitation of using network analysis on extracted information, the production of junctions and roads in separated feature classes can save a considerable amount of manual modelling work.
6. CONCLUSIONS

Based on the data collection and analysis by using the investigated computerised approach, several key points can be concluded as follows:

1. The OSM provides valuable data that can be used for traffic management in Basrah City.
2. The computerised approach achieved its aim by extracting online information from the OSM platform to an independent feature class that can be used offline in GIS software.
3. The physical information gathered from OSM such as road length and width show a high level of accuracy and can be used for to run further analysis. Other information related to the name and direction needs to be viewed again for manual validation.
4. A limited type of analysis such as density and road and junction analysis can be done on the data extracted from OSM due to limitations in information reliability.
5. Despite the limitation in analysis capacity, the information extracted can save significant time and effort by reducing manual modelling.

FUTURE WORK

Further studies need to be done to use OSM data sources to extract other facilities such as buildings and urban areas.

ACKNOWLEDGEMENTS

The author received no kind of fund to produce this research.

REFERENCES


